ORIGINAL RESEARCH PAPER

Medical image fusion and noise suppression with fractional-order total variation and multi-scale decomposition

Xuefeng Zhang | Hui Yan

College of Sciences, Northeastern University,
Shenyang, China

Correspondence
Xuefeng Zhang, College of Sciences, Northeastern University, 110819, Shenyang, China.
Email: zhangxuefeng@mail.neu.edu.cn

Abstract
Fusion and noise suppression of medical images are becoming increasingly difficult to be ignored in image processing, and this technique provides abundant information for the clinical diagnosis and treatment. This paper proposes a medical image fusion and noise suppression model in pixel level. This model decomposes the original image into a noiseless base layer, a large-scale noiseless detail layer and a small-scale detail layer which contains details and noise information. The fractional-order derivative and saliency detection are used to construct the weight functions to fuse the base layers. The proposed total variation model combines the fractional-order derivative to fuse the small-scale detail layers. The mathematical properties and time complexity of the total variation model are also analysed. And choose-max method is used to fuse the large-scale detail medical layers simply. Our approach is based on fractional-order derivative, which enables keep more information and decrease blocky effects more effectively compared with the integer-order derivative. To verify the validity, the proposed method is compared with some fusion methods in the subjective and objective aspects. Experiments show that the proposed model fuses the source information fully and decreases noise cleanly.

1 INTRODUCTION

Image fusion is significant in image processing, owing to the increase of image acquisition models [1–5]. Image fusion refers to the integration of image information acquired by one or different kinds of sensors. The useful information obtained from multiple sensors is contained in the fused image, which is significant for the machine vision and human operations. In recent years, image fusion has been widely applied in the field of remote sensing, clinical medicine, machine vision and more [4–6]. The medical images fusion [7, 8] is the research focus in the image fusion recently. In the clinical medicine, different imaging devices provide different kinds of medical information whose characteristics and functions are different. Since the information provided by a single device is limited, it is necessary to fuse various medical information acquired by different devices into an image. In general, the medical images are divided into two categories: the anatomical images and the functional images. The anatomical images provide detailed anatomical information. The functional images provide metabolic information of human body. Refs. [9, 10] analyse the significant brain information and classify the brain image with different effective image features, which are beneficial to doctors’ diagnosis by processing the medical image information. However, the resolution of medical image is usually poor. Due to the influence of equipment and external environment, the medical images provided by medical imaging equipment contain noise, which is not conducive to doctors’ diagnosis. Thus, the research of the medical image fusion and noise suppression technique is of great significance for providing reliable and comprehensive disease information, and able to remove the interference of noise on doctors’ diagnosis [11, 12].

The multi-scale transform method divides images into different layers, and fuses layers to ensure the fusion result completely. Popular multi-scale transforms are mainly wavelet, curvelet, pyramid and some effective edge-preserving filters, such as cartoon and texture image decomposition, relative total variation (RTV) structure extraction method and $L_0$ gradient minimization smoothing method [13–15]. The edge-preserving filter is widely used in image fusion because of its simple operation, easy
implementation and effective decomposition. The guided filter [16] is a local linear filter which has effective edge preservation performance. Zhu et al. use guided filter to construct a novel hybrid multi-scale decomposition method which decomposes image at different scales. Then various layers are fused according to the corresponding fusion [17]. Compared with the classical mean filter, the guided filter is able to preserve significant edges under decomposing images. Although cartoon and texture image decomposition filter, RTV and smoothing method based on $L_0$ keep edges effectively, these filters create serious halos and artefacts which result in the poor performance of image. Since the guided filter keeps edges, reduces halos and artefacts and is insensitive to noise, the guided filter is used to decompose images in this paper. Fusion rule directly affects the quality of fusion result, and it also plays a decisive role in the model of image fusion and noise suppression. Image fusion is divided into three levels: pixel level, feature level and decision level. Ref. [18] is based on the optimal feature level fusion to classify the brain image effectively. The feature level fusion method extracts the obvious features to analyse image information, but it loses the detail information of the source image. Image fusion based on the pixel level directly analyses the pre-processed pixel information and fuses it. Compared with the feature level fusion method, the pixel level fusion method keeps the full source information and provide more detail information. Recently, fuzzy logic, deep learning, saliency detection and entropy are applied in pixel level decomposition [19, 20].

Specially, the saliency detection which extracts the significant region is often used in the fusion rules to preserve important regions simply [21–23]. Durga and Ravindra [24] propose an image fusion method based on two-scale image decomposition, and they fuse images using saliency detection fully. The saliency detection extracts some significant regions effectively, but it loses some important details and textures. The weight functions with the image gradient features help to fuse the gradient information of multi-mode medical images into a single image [25]. If only the gradient information is considered in the process of image fusion, the fusion result is incomplete and ignores significant pixel information. Therefore, the saliency detection and gradient features are utilized to build fusion rules in this paper.

With the development of fractional derivative, many practical problems are gradually solved by fractional derivative, such as image restoration, image enhancement, image super-resolution and image fusion [26–29]. Compared with the integer derivative, the fractional derivative has the property of non-locality. It means that the fractional derivative does not rely on the limited previous step information, and depends on all the information in the scope. Because of the non-locality, the fractional derivative keeps more textures compared with integer derivative, and has memorability in the image processing. Especially, Pu et al. [30] use fractional-order differential to enhance images effectively, and their enhancement result contains lots of detail information. Total variation (TV) methods [31–33] contribute to the pixel level fusion which keeps source images information and provides detail information effectively. TV is a kind of energy functional which is robust to the noise and is used to suppress noise. Recently, some literatures apply fractional-order variational model to image fusion. In [34], the authors propose a new fusion model, which is based on fractional differential and variational method. This model includes three terms: fusion, up-resolution, edge enhancement and noise suppression, and it fuses images feasibly. In [35], Liu et al. build a unified variational framework which includes fidelity term and fractional prior term to fuse the low-resolution multi-spectral image and the high-resolution panchromatic image effectively. Fractional-order variational model eliminates the staircase, preserves edges and keeps weak textures and smoothness, but few researchers consider it into the image fusion and denoising.

Motivated by the above observations, the guided filter is regarded as the edge-preserving filter of the multi-scale transform in this paper. It divides an image into a smooth base layer, a smooth large-scale detail layer and a noisy small-scale detail layer. In order to keep textures comprehensively, the fractional gradient and saliency detection are combined to construct the fusion rule of the large-scale detail layers. This fusion rule do not only keep gradient features, but also makes up for the lack of neglecting pixel information. A fractional TV model is built to fuse noisy small-scale detail layers. And its convexity and the existence and uniqueness of the solution are studied mathematically to prove that the model is meaningful to fuse details and remove noise. Then, choose-max method is applied to fuse the base layers. Finally, the fractional order is adaptively selected according to the noise standard deviation of the source images. The novelty of medical image fusion and noise suppression with fractional-order TV and multi-scale decomposition proposed in this paper consists in:

- The guided filter decomposes image into three layers which keep details and edges effectively. This decomposition method is beneficial to fuse information comprehensively and suppress noise effectively.
- The fractional-order derivative and the saliency detection are used to construct the fused weight of the base layer. These weight functions contain the pixel and gradient information fully.
- A novel variational model for the medical images fusion and noise suppression is proposed. It combines the knowledge of fractional-order derivative to fuse the small-scale details and remove noise.
- The convexity of the model and the existence and uniqueness of the solution for the fractional-order TV model are proved. They demonstrate the feasibility of the proposed model from a mathematical view.
- The fractional order is selected adaptively relying on the noise standard deviation of the source images, which improves the fusion results directly.

The rest of the paper is organized as follows. Section 2 describes the proposed medical image fusion and noise suppression method. In Section 3, the experiments results are proposed. Finally, a conclusion is shown in Section 4.

## 2 PROPOSED METHOD

In this part, the technique about medical image fusion and noise suppression is introduced. This model focuses on how to fuse
the useful medical information completely and suppress the source noise effectively. As shown in Figure 1, the framework of the proposed model needs three steps to fuse and denoise medical images: multi-scale decomposition, layers fusion and reconstruction. The noisy medical image is divided into a smooth base layer, a large-scale detail layer and a noisy small-scale detail layer by the guided filter. Fractional-order TV is utilized to integrate the small-scale detail layers. The fractional-order derivative and the saliency detection are used to construct the weight functions to fuse the base layers. And the large-scale detail layers are fused by choose-max method. The fusion result is obtained by the base layer, the large-scale detail and the small-scale detail layer reconstructing. Each step is shown in the following subsections.

2.1 Multi-scale decomposition

In recent years, multi-scale transform is widely used in image fusion. It decomposes original images into various scales according to different techniques, so that the fusion result contains adequate information. The guided filter was proposed by He et al. in 2010 [36]. The guided filter is a local linear filter to remove noise and decompose images into different scales. Assuming that the input image and the guiding image are $I$ and $G$, respectively, and the output image is $O$. The local window with centre point $k$ is represented as $\omega_r(k)$, and its radius is $r$. $O$ is a linear transformation of $G$ at $k$:

$$O = a_k G + b_k, \forall i \in \omega_r(k),$$  

where $(a_k, b_k)$ is a set of constants in the local window $\omega_r(k)$. The cost function in the window $\omega_r(k)$ is minimized:

$$E(a_k, b_k) = \sum_{i \in \omega_r(k)} ((a_k G_i + b_k - I_k)^2 + \lambda a_k^2),$$  

where $\lambda$ is the regular parameter. The optimal $a_k$ and $b_k$ are obtained by linear regression methods:

$$a_k = \frac{1}{|\omega_r|} \sum_{i \in \omega_r(k)} (G_i I_i - \mu_k I_k) \sigma_k^2 + \lambda,$$

$$b_k = \bar{I}_k - a_k \mu_k,$$

where $\mu_k$ and $\sigma_k$ are the mean and variance of $G$ in $\omega_r(k)$, respectively, $\omega$ is the pixels number of $\omega_r(k)$, and $\bar{I}_k$ is the mean of $I$ in $\omega_r(k)$. The final $O_i$ is calculated as

$$Q_i = \bar{a}_i G_i + \bar{b}_i,$$

where $\bar{a}_i$ and $\bar{b}_i$ are the mean values of $a_k$ and $b_k$ in the window. The guided filter is expressed as

$$O = GF(G, I, r, \lambda).$$

Then, the guided filter is used to construct the various layers. The guided image in our decomposition step based on the guided filter is the same as the input image. Two filtering layers
are obtained as
\[ B^1_n = GF(u_i, u_j, r, \lambda), \quad n = 1, 2, \]
\[ B^2_n = GF(B^1_n, B^2_n, r, 3 \times \lambda), \]  
where \( u_i, \quad n = 1, 2 \) are the original noisy medical images, \( B^1_n \), and \( B^2_n, \quad i = 1, 2 \) are smoothing layers with different levels. The small-scale detail layers \( B^1_n, \quad n = 1, 2 \) are obtained
\[ D^1_n = u_n - B^1_n, \quad n = 1, 2. \]  
The large-scale detail layers \( B^2_n, \quad n = 1, 2 \) are obtained
\[ D^2_n = B^1_n - B^2_n, \quad n = 1, 2. \]  
And the base layers \( D^2_n, \quad n = 1, 2 \) are obtained
\[ D^2_n = B^2_n, \quad n = 1, 2. \]

Multi-scale decomposition decomposes original images into three scales which represent different component. And the guided filter is used to obtain different scales layers which preserve significant edges and overcome the block effect effectively. The base layers in this paper are smooth, and they include background changes. The small-scale detail layers contain the noise and a part of details. And the large-scale layers contain the other details. The fusion method based on multi-scale decomposition fuse images in three scales, and different scales need special fusion rules to fuse information comprehensively.

### 2.2 Base layers fusion

The base layers contain background information of the original images, and the key in medical base layers fusion lies in the fusion rules. The medical base layers fusion aims to integrate the anatomical and metabolic information of human body. The classical choose-max method is based on the pixels, and it loses some important metabolic information when the brightness difference between the source medical images is large. Gradient-based fusion rules have obvious advantages in medical image processing. The integer gradient operator is defined as
\[ G(x, y) = \sqrt{u^2_x + u^2_y}, \]
where \( u_x \) and \( u_y \) are the difference operators, and \( u \) is the input image.

Fractional-order derivative, an important branch of mathematics, was born in 1695. In recent years, the theory of fractional derivative has been successfully applied to various fields, and it describes some non-classical phenomena in the fields of natural science and engineering applications [37, 38]. The popular definitions of the fractional derivative are Riemann–Liouville (R–L), Caputo and Grünwald–Letnikov (G–L), and the G–L definition is widely used in image processing [39]. The definition of the discrete G–L derivative at point \((i, j)\) along the \(x\) and the \(y\) directions are as follows:
\[ D^\alpha_{x} u_{ij} = \sum_{k=0}^{\alpha+1} \Gamma(\alpha+1) \frac{\Gamma(\alpha+1-k+1)}{\Gamma(\alpha-k+1)} u_{i,j-k+1}, \quad i, j = 1, 2, \ldots, n, \]
\[ D^\alpha_{y} u_{ij} = \sum_{k=0}^{\alpha+1} \Gamma(\alpha+1) \frac{\Gamma(\alpha+1-k+1)}{\Gamma(\alpha-k+1)} u_{i-j+k+1}, \quad i, j = 1, 2, \ldots, n, \]
where \( u^\alpha_k = (-1)^k \frac{\Gamma(\alpha+1)}{\Gamma(\alpha-k+1)} \) and \( k = 0, 1, \ldots, n + 1 \). The coefficients \( u^\alpha_k \) has the following recursive relation:
\[ u^\alpha_0 = 1, \quad u^\alpha_k = \left(1 - \frac{\alpha + 1}{k} \right) u^\alpha_{k-1}. \]

According to the above definition, the discrete \( \alpha \)-order derivative is written in the matrix form
\[ D^\alpha u = \begin{pmatrix} D^\alpha_{x} u \\ D^\alpha_{y} u \end{pmatrix} = \begin{pmatrix} Mu \\ uM^T \end{pmatrix} \in R^{2 \times n}, \]
where
\[ M = \begin{pmatrix} w^\alpha_1 & 0 & \cdots & 0 \\ w^\alpha_1 & w^\alpha_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ w^\alpha_1 & \cdots & w^\alpha_2 & w^\alpha_1 \end{pmatrix} \in R^{2 \times n}, \]
and \( w = w^\alpha_0 + w^\alpha_2 \).

The fractional-order derivative has the property of non-locality, which causes fractional-order derivative to keep more detail information in image processing compared with integer one. Thus, the fractional-order derivative is used in image gradient
\[ G^\alpha(x, y) = \sqrt{(Mu)^2 + (uM^T)^2}. \]

To simplify the calculation, the fractional operator is truncated in a proper manner and the fractional-order derivative is defined as
\[ G^\alpha(x, y) = |Mu| + |uM^T|. \]

The gradient of a medical image with different orders is shown in Figure 2. Figure 2(a) is the original image. Figure 2(b) shows the integer gradient of Figure 2(a), and Figures 2(c) and (d) are gradient figures with order 1.3 and 1.7, respectively. In Figure 2, the figures based on fractional-order derivative contain more information compared with Figure 2(b). The functional information in the middle of Figures 2(c) and (d) are more abundant and obvious than that of Figure 2(b), and the gradient
of the anatomical information is vividly shown in Figures 2(c) and (d).

The saliency detection automatically detect the significant regions, and it is widely used in image fusion. It designs fusion rules from the perspective of image pixels. Achanta et al. proposed a classical saliency detection method which is called frequency tuned (FT) method [40]. FT is easy to implement, and it analyses and detects the image from the perspective of frequency. FT extracts image as

\[ S(x, y) = \| u_{\alpha}(x, y) - u_{	ext{subc}}(x, y) \|, \]  

where \( u_{\alpha}(x, y) \) is the mean value of the input image \( u \), \( u_{	ext{subc}}(x, y) \) is the pixel blurred by the Gaussian filter and \( \| \cdot \| \) is the Euclidean distance.

In the base layer fusion step, the base functional layer is completely reserved to keep full functional information. And the anatomical information is injected into the functional layer to integrate the saliency textures. Combined with the fractional-order gradient and FT, the weight functions of the base layers are as follow:

\[ u^\alpha = GF(G_1^\alpha + S_1, C_1^\alpha + S_1, r, \tilde{\lambda}), \]  

where

\[ C_1^\alpha(x, y) = \| MB_1 | + | B_1 M^T |, \]

\[ S_1(x, y) = \| B_1 u_{\alpha}(x, y) - B_{\text{subc}}(x, y) \|, \]

and \( B_1 \) is the anatomical base layer.

Equation (18) combines the fractional-order gradient and the saliency detection. It considers the image gradient and the saliency information to construct the fusion weight functions, which overcomes the singularity of classical pixel-based or gradient-based fusion method. The proposed function is simple and has fast computing speed. Due to the application of saliency detection, the weight function outlines the detailed anatomical features. The guided filter is used to smooth \( u^\alpha \) to decrease noise easily. Then, the base layers are fused according to (18)

\[ B(x, y) = u^\alpha(x, y)B_1(x, y) + B_2(x, y), \]  

where \( B \) is the fused base layer, and \( B_2 \) denotes the functional base layer.

2.3 Small-scale detail layers fusion

What we concern here is how to integrate two noisy small-scale detail layers \( D_1^1(x, y) \) and \( D_2^1(x, y) \) into a noiseless detail layer image \( D(x, y) \) such that edge details of the two input medical layers are preserved and the noise are suppressed effectively (the given layers \( D_1(x, y) \), \( D_2(x, y) \) and \( D(x, y) \) has \( N_1 \times N_2 \) pixels). Denote \((x, y) \in \Omega \subseteq \mathbb{R}^2 \) where \( \Omega \) is bounded domain of the image with two space dimensions and has a Lipschitz boundary. Based on the model in [41], the fractional-order variational model for the detail layers fusion and noise suppression is as follows:

\[ \min_{D_i \in B^{1-\alpha}(\Omega)} E(D^i) := F(D^i) + \lambda TV^{\alpha}(D^i), \]  

\[ F(D^i) = \int_{\Omega} (\omega_1(D^1_i - D^1) + \omega_2(D^1_i - D^2_i)^2)dx\,dy, \]  

\[ TV^{\alpha}(D^i) = \int_{\Omega} \left( \frac{\partial^\alpha \nu}{\partial \xi^\alpha} \right)^2 + \left( \frac{\partial^\alpha \nu}{\partial \eta^\alpha} \right)^2 dx\,dy, \]  

where \( D^1 \), \( D^1_i \) and \( D^2 \) are the vector forms of matrix \( D^1 \), \( D^1_i \) and \( D^2 \), respectively, \( D^1(x, y) \), \( D^1_i(x, y) \) and \( D^2(x, y) \) \( \in \mathbb{R}^{1 \times N_1 N_2} \), \( \omega_1 \) and \( \omega_2 \) are diagonal weight matrices which satisfy \( \omega_1(i, i) \geq 0, \omega_2(i, i) \geq 0, \omega_1(i, i) + \omega_2(i, i) = 1, \) \( i = 1, 2, \ldots, N_1 N_2 \), \( TV^{\alpha}(D^i) \) represents the fractional-order variation whose order is \( \alpha (\alpha \in (1, 2)) \), \( \lambda = 2\alpha \) is a scalar which controls the balance between the two terms in the model, and \( \nu \in \mathbb{R}^{1 \times N_1 N_2} \) is the diagonal smoothness weight.

The small-scale detail layers contain small scale changes such as edges, sharp textures and some details. In order to integrate the detail information fully, the choose-max method is used to build the weights functions \( \omega_1 \) and \( \omega_2 \). As for the smoothness weight \( \omega_2 \), the four-directional Laplacian energy is used based on the two-direction Laplacian energy equation.

\[ L_i = \frac{\partial^2 D_i}{\partial x^2} + \frac{\partial^2 D_i}{\partial y^2} + \frac{\partial^2 D_i}{\partial i^2} + \frac{\partial^2 D_i}{\partial j^2}, \]  

\( i = 1, 2. \)
The smoothness weight in the regular term is to keep details while denoising the small-scale detail layer. The Laplacian energy detect textures and noise at the same time which leads to removing noise incompletely. Thus, it is necessary and important to combine the Laplacian energy and other techniques to construct the smoothness weight. First, the Laplacian energies of two layers are normalized $L_i = \frac{L_i}{\max L_i}$, $i = 1, 2$. The max($L_1, L_2$) reflects layer details and noise, and $\min(L_1, L_2)$ is the noise information. Then, the weight of the detail information is calculated simply $d = \max(L_1, L_2) - \min(L_1, L_2)$. Finally, the smoothness weight is defined $\omega = \text{diag}(1 - d)$.

In model (21), the first term (22) is called fusion term which serves for image fusion. The second term (23) is regularity term which serves for noise suppression. Equation (23) keeps smoothness and preserve edges while reducing noise. When $\alpha = 1$ and $\omega = I$, model (21) is

$$
\min_{D^1 \in BV(\Omega)} E(D^1) := \int_{\Omega} \left\{ w_1(D^1 - D_1^1)^2 + w_2(D^1 - D_2^1)^2 \right\} \, dx \, dy
+ \lambda \int_{\Omega} \left\{ \frac{\partial D^1}{\partial x} \right\}^2 + \left\{ \frac{\partial D^1}{\partial y} \right\}^2 \, dx \, dy,
$$

which is a classical variational model for fusion and denoising [41]. When the input images are noise-free, the fractional-order variational term in model (21) vanishes. The novel model is used to fuse different small-scale detail layers in the pixel-wise. Furthermore, model (21) reduces to fractional-order TV denoising model when the source noisy layer is a single, and the simplified model [42] is described by

$$
\min_{D^1 \in BV(\Omega)} E(D^1) := \int_{\Omega} \left\{ (D^1 - D_1^1)^2 \right\} \, dx \, dy
+ \lambda \int_{\Omega} TV^\alpha(D^1) \, dx \, dy.
$$

In order to solve the practical medical layers fusion problem meaningfully, it is important and necessary to prove the research significance of model in mathematics. However, this step is ignored by many researchers. The convexity and the existence and uniqueness of the solution for the variational model is proved to support the rationality of this technique and the completeness of this paper.

The definition of the fractional TV is the basis of the relevant proof of the model. Thus, based on the fractional derivative and Refs. [43, 44], the fractional TV definition is given. Let $C_0^\alpha(\Omega, \mathbb{R}^2)$ denote the $\ell$-order compactly supported continuous-integrable functions space, where $\Omega \subset \mathbb{R}^2$. The space of special test functions is

$$
K := \{ \varphi \in C_0^\alpha(\Omega, \mathbb{R}^2) \mid |\varphi(x,y)| \leq 1, \forall x \in \Omega \},
$$

$$
|\varphi| = \sqrt{\varphi_1^2 + \varphi_2^2}.
$$

The $\alpha$-order TV of $\bar{D}^1(x,y)$ is defined as

$$
\sup_{\Omega} \int (\bar{D}^1 \nabla^\alpha \varphi) \, dx \, dy, \quad \varphi \in K,
$$

where $\nabla^\alpha \varphi = D^\alpha_x \varphi_1 + D^\alpha_y \varphi_2$. Then, the $\alpha - BV$ norm is denoted as

$$
\| D^1 \|_{BV^\alpha} := \| D^1 \|_{L^1} + \sup_{\Omega} \int (\bar{D}^1 \nabla^\alpha \varphi) \, dx \, dy,
$$

and the space of functions of $\alpha$-bounded variation on $\Omega$ (Banach space $BV^\alpha(\Omega)$) is defined by

$$
BV^\alpha(\Omega) := \{ D^1 \in L^1(\Omega) \mid \sup_{\Omega} \int (\bar{D}^1 \nabla^\alpha \varphi) \, dx \, dy < +\infty \}.
$$

Assume $D^1 \in W^\alpha_1 = \{ D^1 \in L^1(\Omega) \mid \| D^1 \|_{BV^\alpha(\Omega)} < +\infty \}$, and proposition $\int_{\Omega} |\nabla^\alpha D^1| \, dx \, dy$ holds.

**Property 1** (Convexity). When $\lambda \geq 0$, the functional $E(D^1)$ in $BV^\alpha(\Omega)$ is convex. And it is strictly convex if $\lambda > 0$.

**Proof.** The fractional-order derivative $D^\alpha$ has the property of linearity. For any fractional differentiable functions $f_1(x)$, $f_2(x)$ and $a_1, a_2 \in \mathbb{R}$,

$$
D^\alpha(a_1 f_1(x) + a_2 f_2(x)) = a_1 D^\alpha f_1(x) + a_2 D^\alpha f_2(x).
$$

The $\int_{\Omega} |\nabla^\alpha D^1| \, dx \, dy$ is positively homogeneous and sub-additive. Due to the properties of above functional, $TV^\alpha(D^1) = \int_{\Omega} |\nabla^\alpha D^1| \, dx \, dy$ is convex. And $F(D^1) = \int_{\Omega} (\omega_1 D^1 - D_1^1)^2 + \omega_2(D^1 - D_2^1)^2 \, dx \, dy$ is strictly convex.

Since the $TV^\alpha(D^1)$ and $F(D^1)$ are convex, the functional $E(D^1)$ has convexity.

**Property 2** (Existence and uniqueness). When $\lambda > 0$, there is a unique minimum for the functional $E(D^1)$ in $BV^\alpha(\Omega)$.

**Proof.** To prove the existence and uniqueness of the solution of model (7), we set a minimizing sequence $D^1(i) \in BV^\alpha(\Omega) \subset L^1(\Omega)$ of (7). The minimum of $E(D^1)$ is finite because $E(0)$ is finite. Thus, there is a positive constant $C$ such that the following formula is true:

$$
E(D^1(i)) := \int_{\Omega} (\omega_1(D^1(i) - D_1^1)^2 + \omega_2(D^1(i) - D_2^1)^2 \, dx \, dy + \lambda TV^\alpha(D^1(i)) \leq C.
$$

Since $E(D^1) \leq C$, we have

$$
TV^\alpha = \int_{\Omega} |\nabla^\alpha D^1(i)| \, dx \, dy \leq C \text{ and } F(D^1) \leq C.
$$


Owing to $D_1^1, D_2^1 ∈ L^1(Ω)$,
\[ \int \int_Ω \omega |D^1(i)| \, dx \, dy ≤ C \]  \tag{34}
holds. According to formulas (33) and (34), $D^1(i)$ is bounded in $BV^α$. And in $BV^α(Ω)$, a weak $BV^α(Ω) - ω^*$ topology is
\[ D^1(i) \xrightarrow{BV^α(Ω)} D^1 \xrightarrow{L^1(Ω)} D^1. \]  \tag{35}
It means that there exists a subsequence $D^1(i)$ which converges to $D^1 ∈ BV^α ⊂ L^1(Ω)$. Based on the semi-continuity of the fractional-order TV, we deduce
\[ \int \int_Ω \omega |D^α D^1| \, dx \, dy ≤ \lim_{i→∞} \inf i \int \int_Ω \omega |D^1(i)| \, dx \, dy. \]  \tag{36}
Then $E(D^1) ≤ \lim_{i→∞} \inf E(D^1(i))$ and $D^1(i)$ is a minimum of model (21). Finally, the minimum is unique according to Property 1.

Using matrix notation, Equation (21) is rewritten as
\[ \min_{D^1} (\hat{D}^1 - D^1) \omega_1(\hat{D}^1 - D^1) + (\hat{D}^1 - D^1) \omega_2(\hat{D}^1 - D^1) + \lambda(D^1 D^α D^1 + D^1 D^α D^1 \omega D^α D^1). \]  \tag{37}
Because Equation (21) has a unique solution, the vector $\hat{D}^1$ is defined
\[ (I + \lambda(D^α D^1 + D^α D^1))D = \omega_1 \hat{D}^1 + \omega_2 \hat{D}^1. \]  \tag{38}
Then, we have
\[ \hat{D}^1 = (I + \lambda(D^α D^1 + D^α D^1))^{-1}(\omega_1 \hat{D}^1 + \omega_2 \hat{D}^1). \]  \tag{39}

Through the above proofs and analyses, proposed optimization-based model (21) has a unique solution which is calculated as (39). Thus, the fractional TV model is solved by the Laplacian matrix directly, resulting in an algorithm with $O(1)$ computational complexity.

2.4 Large-scale detail layers fusion

The large-scale detail layers contain detail information which is different from that of the small-scale detail layers. In order to fuse the details fully, the classical choose-max method is used in this subsection to fuse details
\[ D^2 = \max(D^1, D^2), \]  \tag{40}
where $D^2$ denotes the fused large-scale detail layer, $D^1$ and $D^2$ are the large-scale detail information of $u_1$ and $u_2$, respectively.

2.5 Image reconstruction

The fused image $F$ is obtained by multi-scale image reconstruction which adding the fused base layer $B$, the denoised final detail layer $D^1$ and the fused large-scale detail layer $D^2$
\[ F = B + D^1 + D^2. \]  \tag{41}

2.6 Adaptive fractional order

The fractional-order derivative is used in the fusion of the base layers and small-scale detail layers, and the selection of fractional order plays an important part in the medical fusion and noise suppression model. The order $α$ is related to the medical images fusion and denoising effect directly. Adaptive algorithm is a research hotspot in many fields [45–47]. In a sequel, $α (α ∈ [1, 2])$ is adaptively selected according to the original noise. The classical estimation method for the standard deviation of additive zero mean Gaussian noise in an input image is utilized [48]. The two kinds of Laplacian filters $H_1$ and $H_2$ process original images first,
\[ H_i = \frac{(H_i * u_i - H_2 * u_i)}{2}, \quad i = 1, 2, \]  \tag{42}
where $H_i = \begin{pmatrix} 0.5 & 0 & 0.5 \\ 0 & -2 & 0 \\ 0.5 & 0 & 0.5 \end{pmatrix}$, $H_2 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & -4 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, $u_i, \quad i = 1, 2$ is the input image, and $*$ is the convolution. The standard deviation $σ$ is calculated:
\[ σ_i = \sqrt{\frac{1}{2} \sum_{i,j} |u_i(x,y)|}, \quad i = 1, 2, \]  \tag{43}
where $(N_1, N_2)$ is the size of $u_i, \quad i = 1, 2$. Then, the standard deviation is normalized as $σ_i = \frac{255 σ_i}{\sqrt{(N_1-1)(N_2-1)}}$, $i = 1, 2$, and $\sqrt{σ_i}$ means the noise density of $u_i$. The difficulty of image fusion and denoising is increasing with the noise of the original images increasing. Thus, the value of fractional order increases with the noise density. In order to keep details vividly and suppress noise cleanly, $\sqrt{σ_1}, σ_2$ larger than $\sqrt{σ}$ is used to calculate the approximate noise density within a reasonable range. Based on the noise standard deviation of the original noisy
images, the fractional order is defined as
\[
\alpha = \sqrt{\max(\sigma_1, \sigma_2)} + 1, \tag{44}
\]
where \(\alpha \in [1, 2]\).

Equation (44) shows that the order is directly related to the noise density. When the noise level is large, the fractional order is large so that the gradient information of the base layers is abundant and fusion effect of the bases layers is complete. In the processing of details layers fusion, the fractional-order model \(\alpha\) removes more noise than integer case. When the noise standard deviation is small, the smoothness of the fractional-order TV model is small to keep rich details while denoising. On the contrary, a bigger noise standard deviation results in bigger \(\alpha\) to remove noise fully.

3 | EXPERIMENTAL RESULTS

Medical image fusion and noise suppression are interdiscipli- nary subjects of modern computer technique and medical imaging. It has become an indispensable part of modern medical treatment. In this section, the experiments of the medical images are executed by mathematical software MATLAB. The medical fusion image data sets provided by Bavirisetti are used in the experiment [49]. The data sets are supported by Bavirisetti and widely used in the image fusion. They contain various representative aligned images, such as medical images, multi-focus images and remote sensing images.

Five fusion methods are compared with the proposed algorithm. These are: the discrete wavelet transform (DWT) [50], the deep convolutional neural network (CNN)-based fusion method [51], the multi-spectral image fusion method with fractional-order differentiation (AFM) proposed in [52], the medical image fusion and denoising with alternating sequential filter and adaptive fractional-order total variation (ASAFTV) [53] and variational models for fusion and denoising of multi-focus image (WTV) [41]. The choose-max method is used to serve as the fusion rule of DWT which fuses information conservatively and classically. However, choose-max method is based on pixel information, and it leads to losing some important fusion gradients. CNN uses deep convolutional neural network which is popular. CNN is able to fuse information fully, but its calculation time is longer than those of other methods. AFM and ASAFTV are based on fractional-order derivative which is helpful to denoise sufficiently and reduce blocks and artefact effectively. However, AFM is not suitable for fusing medical images due to its fusion rule. Due to the using of the mean filter, ASAFTV blurs significant edges when fusing features and suppressing noise. WTV fuse and denoise images with TV model which creates serious blocks. Our method uses the guided filter to decompose images, so it keeps more edges and textures compared with ASAFTV. What is more, the fractional TV, saliency detection and gradient information are utilized to build fusion rules, which fuse images and suppress noise fully.

3.1 | Performance index

Five objective performance indices, i.e. normalized mutual information (NMI) index [54], gradient-based \(Q^{GF}\) index [55], feature similarity (FSIM) index [56], multi-scale structural similarity (MS-SSIM) index [57] and visual information fidelity for fusion (VIFF) [58] are used to evaluate the quality of fusion and denoising results. Uniformly, \(n_1\) and \(n_2\) denote the original images, and \(F\) denotes the fused image, respectively.

NMI is a normalized index based on mutual information, which is used to calculate how much source information has been transferred to the fused image. The larger the NMI, the better the image fusion effect. Its formula is given by
\[
NMI = 2 \times \frac{\text{MI}(u_1, F) + \text{MI}(u_2, F)}{H(u_1) + H(u_2) + H(F)}, \tag{45}
\]
where \(H(u_1), H(u_2)\) and \(H(F)\) are the marginal entropy of \(u_1, u_2\) and \(F\), respectively, \(\text{MI}(u_1, F) = H(u_1) + H(F) - H(u_1, F)\), \(i = 1, 2\) is the mutual information between \(u_i\) and \(F\), and \(H(u_i, F)\), \(i = 1, 2\) is the joint entropy between \(u_i\) and \(F\).

The gradient-based index \(Q^{GF}\) calculates the gradient information of the fused image transferred from the source images. A higher \(Q^{GF}\) means a better fusion result. It is given by
\[
Q^{GF} = \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \left( Q^{GF_1}(i,j) \omega_1(i,j) + Q^{GF_2}(i,j) \omega_2(i,j) \right) / \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \left( \omega_1(i,j) + \omega_2(i,j) \right), \tag{46}
\]
where \((N_1, N_2)\) is the image size, \(Q^{GF_1} = Q^{GF_1}(i,j)\) and \(Q^{GF_2}\) and \(Q^{GF_2}\), \(i = 1, 2\) denote the edge strength and orientation information preserved in \(F\) from \(u_i\), respectively, and \(\omega_i, i = 1, 2\) is weighting factor of \(Q^{GF}\).

FSIM measures the feature similarity degree between the source images and the fused image. The higher the value of FSIM is, the better fusion result is achieved. FSIM is given by
\[
\text{FSIM} = \frac{1}{2} \left( \text{FSIM}(u_1, F) + \text{FSIM}(u_2, F) \right),
\]
\[
\text{FSIM}(u_i, F) = \frac{\sum_{j \in \Omega} S_{FC}(j) S_{GC}(j) IC(j)}{\sum_{j \in \Omega} IC(j)}, \tag{47}
\]
\[
IC(j) = \max(I_{GF_1}(j), I_{GF_2}(j)),
\]
\[
S_{FC}(j) = \frac{2 I_{GF_1}(j) I_{GF_2}(j) + T_1}{I_{GF_1}(j)^2 + I_{GF_2}(j)^2 + T_1},
\]
\[
S_{GC}(j) = \frac{2 G_{GF_1}(j) G_{GF_2}(j) + T_2}{G_{GF_1}(j)^2 + G_{GF_2}(j)^2 + T_2},
\]
where \(\Omega\) is the image spatial domain, \(j\) denotes the spatial coordinate, \(T_1, T_2\) are constants, \(IC\) is the importance of local structure and \(G\) measures the gradient magnitude.
**FIGURE 3** The original images

**TABLE 1** Quantitative assessment of fusion methods for noise-free images

| Image set   | 1         | 2         | 3         | 4         | 5         | 6         | 7         | 8         | 9         | 10        |
|-------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| Index       | NMI       | QC        | FSIM      | MS-SSIM   | VIFF      | NMI       | QC        | FSIM      | MS-SSIM   | VIFF      |
| DWT         | 0.9749    | 0.4067    | 0.9606    | 0.9999    | 0.4798    | 0.8339    | 0.4388    | 0.9622    | 0.9998    | 0.3698    |
| CNN         | 0.7697    | 0.6212    | 0.9516    | 0.9999    | 0.7118    | 0.6888    | 0.6754    | 0.9618    | 0.9997    | 0.7058    |
| AFM         | 0.8909    | 0.3932    | 0.9264    | 0.9999    | 0.3873    | 0.8555    | 0.4393    | 0.9536    | 0.9989    | 0.1472    |
| ASAFTV      | 0.7690    | 0.2810    | 0.9562    | 0.9999    | 0.8240    | 0.7828    | 0.4281    | 0.9605    | 0.9991    | 0.5897    |
| WTV         | 0.8814    | 0.4268    | 0.9595    | 0.9999    | 0.5887    | 0.8777    | 0.5661    | 0.9578    | 0.9994    | 0.5873    |
| Ours        | 0.7680    | 0.4849    | 0.9585    | 0.9998    | 0.7991    | 0.8200    | 0.5717    | 0.9602    | 0.9995    | 0.9392    |

**Note:** (In the sequel, the boldened value means the best result for each group of experiments.

MS-SSIM is an extension of the structural similarity. Compared with the structural similarity, MS-SSIM is more consistent with the visual perception of human visual system, and its evaluation is more effective than those of the structural similarity. The higher SSIM is, the higher similarity between the fused image and the original images is obtained. The MS-SSIM is defined as

\[
MS - SSIM = \frac{1}{2}(MS - SSIM(u_1, F) + MS - SSIM(u_2, F)),
\]

\[
MS - SSIM(u_i, F) = \frac{1}{|W|^2} \sum_{\mu=1}^{|W|} SSIM(u_{iw}, F_w),
\]

\[
SSIM(u_{iw}, F_w) = \frac{(2\mu_1\mu_2 + C_1)(2\sigma_{w_1w_2} + C_2)(\sigma_{w_1w_2}^2 + C_3) + C_2 \sigma_{w_1}^2 + C_3 \sigma_{w_2}^2)}{(\mu_1^2 + \mu_2^2 + C_1)(\sigma_{w_1}^2 + \sigma_{w_2}^2 + C_2)(\sigma_{w_1w_2}^2 + C_3)},
\]

where \(\mu\) and \(\sigma_w\) are the mean and the variance of the region \(w\), respectively, and \(\sigma_{w_iw_j}\) is the covariance of the regions \(w_i\) and \(w_j\), \(i = 1, 2\). In general, \(C_1 = 0, C_2 = 0\) and \(C_3 = 0\).

VIFF based on the visual information fidelity evaluates the fusion quality in a perceptual manner. The higher VIFF is, the better fusion quality is obtained. Due to its complexity, VIFF is shown [58].

### 3.2 Fusion results

Figure 3 shows three sets of noise-free images. Table 1 and Figures 4–6 show fusion experiments for original images in Figure 3 with six fusion methods. In Figures 4–6, the first column is the fused images proposed by the classical DWT. The second column shows the results of CNN. The third method is AFM which is based on fractional-order derivative. The fourth method is ASAFTV adopted the fractional-order TV model. The fifth line uses WTV which fuses and denoises images with TV model, and the final method is our method.
To prove the proposed method fuse the anatomical textures and the metabolic details fully and compare the fusion results of different methods, the subjective evaluation of Figures 4–6 is analysed. In Figure 3(a), the difference between the original images is large and the fusion of Figure 3(a) is difficult. Figure 4 is the fusion result of Figure 3(a). In Figure 4(d), the image fused by DWT contains too much anatomical information and ignores the metabolic textures. The detail information of Figure 4(d) is smoothed by ASAFTV. In Figures 4(b) and (c), the images fused by DWT and AFM lose important skeletal structures unfortunately. Figure 4 shows our method and CNN integrate the anatomical and metabolic information of original images more fully than other methods, especially the metabolic details and textures are fused richly by our method and CNN.

Figure 5 shows the fusion images of image set 2. In Figure 4, ASAFTV and WTV smooth medical images seriously and result in blurring the fusion results. In Figures 5(a) and (c), it is concluded DWT and AFM lose some source information in different degrees when fusing images, such as bones of Figure 5(a), and metabolic details of Figure 5(c). The above situation does not appear in Figures 5(b) and (f) proposed by our method and CNN, respectively. Figure 6 shows the fused images of image set 3. In Figure 6(c), AFM displays the anatomic textures of source image well, but it loses the most of metabolic information. So AFM is suitable for fusing multi-spectral images. Since the Gaussian low-pass filter and the feature extraction are used in ASAFTV, Figure 6(d) fused by ASAFTV are not ideal. The proposed method keeps the source information completely in the processing of fusing noise-free medical images.

Table 1 shows NMI, $Q^c$, FSIM, MS-SSIM and VIFF of Figures 4–6. The objective evaluation and statistical analysis of different methods are analysed. NMI and $Q^c$ evaluate the amount of fused information. FSIM and MS-SSIM measure the similarity between source images and fused images in the featural and structural aspects, respectively. VIFF analyses the quality of fused images in a visual manner. The large value of the above indices shows that the quality of the fused image is good. The boldened value in Table 1 means the best result for each group of experiments. Although the NMI of images fused by DWT is the biggest in Table 1, the saliency and important functional
FIGURE 7 Comparison of different methods for denoising, where white Gaussian noise $\sigma = 20$. Panels (a) and (b) Original images, (c) WTV, (d) ASAFTV and (e) Ours method.

information is not vividly shown in images fused by DWT. In Table 1, the NMI and $Q^G$ of fused images proposed by our method are in the middle, since some anatomical background is ignored to highlight and keep full functional information. MS-SSIM measures the average similarity between two inputs and an output. The image fusion is an information selection. In order to keep more important and detail information, some unimpor-
tant background information is lost in the fused image. Thus, MS-SSIM of images fused by our method is smaller than DWT. The proposed method achieves the best visual fusion effect for statistical analysis of VIFF. For image sets 1–3, five indices of our method achieve good results. Table 1 shows the proposed method fuses medical images fully and perceptually. And it is necessary to analyse and compare the fusion results scientifically and fairly from the subjective and objective perspectives.

3.3 Noise suppression results

In this subsection, image of “Brain” is used in the experiment. In order to outstand the effect of noise suppression with different methods, the images which added additive white Gaussian noise are chosen as inputs of the model. The noise levels of the input images are $\sigma = 20$ and $\sigma = 30$. The results are shown in Figures 7 and 8. Noise suppression and fusion are considered at the same time in our model. In Figures 7 and 8, WTV has a strong ability of suppressing noise. However, WTV leads to the serious blocky effect, which is obvious with the increase of noise density. Especially when the noise level is 30, the block in Figure 8(c) seriously affects the readability of the image. From Figures 7(d) and 8(d), ASAFTV avoids blocky effect effectively when removing images noise, but the brightness of Figures 7(d) and 8(d) is dark. After the analysis and experimental comparison of ASAFTV, the large smoothness degree of TV model and the Gaussian filter using in the fusion rules lead to losing some important fusion information. Compared with WTV and

FIGURE 8 Comparison of different methods for denoising, where white Gaussian noise $\sigma = 30$. Panels (a) and (b) Original images, (c) WTV, (d) ASAFTV and (e) Ours method.
ASAFTV, our method avoids the block effect effectively when suppressing noise, and it integrates the useful information of the source image completely.

Figure 9 shows the comparison of zoomed fused images in Figures 7 and 8, where the white Gaussian noise is $\sigma = 20$ and $\sigma = 30$. In Figures 9(a) and (d), the block effect of WTV based on integer-order TV model is serious. Because the fractional-order TV model is used in ASAFTV and our method, the block is avoided in Figures 9(b)–(c) and (e)–(f). The results proposed by ASAFTV are still affected by the fusion rules. ASAFTV reduces the image contrast in zooms images, and it provides dark fused images. From Figure 9, we conclude that our method fuses rich features and salient information from the noisy input images.

### 3.4 Comparison of advantages between fractional model and integer model

In order to outstand the effect of fusion and noise suppression of the proposed model with integer order and adaptive fractional order, the images blocks ($100 \times 100$) where the white Gaussian noise level is $\sigma = 30$ are chosen as the inputs of the models. The original images are shown in Figures 10(a) and (b). The results of integer-order model are shown in Figure 10(c), and the adaptive fractional-order results are shown in Figure 10(d). In Figure 10(c), the image fused by integer-order model contains obviously noise. Figure 10(c) is affected by the noise and fuses some noise in the metabolic details. The

![Figure 10](image_url)

**FIGURE 10** Comparison of advantages between fractional order model and integer order model, where white Gaussian noise $\sigma = 30$. Panels (a) and (b) Original images; (c) integer order approach and (d) adaptive fractional order approach

results processed by integer-order model are affected by the noise and artefacts. Figure 10(d) fused by the proposed adaptive fractional-order model fuses details and textures fully and suppresses noise effectively. Table 2 shows NMI, FSIM and
TABLE 2  Comparison between the adaptive fractional-order model and the integer-order model with different noise

| Index order | NMI | FSIM | VIFF |
|-------------|-----|------|------|
| 1           | 0.8358 | 0.8396 | 0.9631 | 0.9618 | 0.8428 | 0.9916 |
| 20%         | 0.7906 | 0.7974 | 0.9617 | 0.9634 | 0.7748 | 0.9478 |
| 30%         | 0.7697 | 0.7680 | 0.9609 | 0.9616 | 0.7155 | 0.9466 |
| 40%         | 0.7563 | 0.7540 | 0.9572 | 0.9580 | 0.6716 | 0.8015 |

VIFF of the images fused by the integer-order model and the adaptive fractional-order model, in which the source images contain different noise levels. NMI, FSIM and VIFF analyze the fused image in three aspects, containing the amount of information, the similarity with source images and the visual perception. It is shown in Table 2 that three indices of the adaptive fractional-order model are mostly larger than that of integer-order model. Table 2 shows that the adaptive fractional-order model can be used to fuse more details, textures and keep the fused image more perceptual than integer-order model in the numerical analysis. According to the analysis of Figure 10, we conclude that the proposed adaptive fractional TV model fuses noisy images fully and removes noise effectively. Compared with the proposed fusion model with integer order, the adaptive fractional-order model utilizes fractional-order derivative to fuse rich information and suppress noise, and the noise detection of it selects order adaptively to balance the degree of denoising and fusing. From the perspective of human vision, the fusion effects of the fractional model are better than those of integer order in the medical fusion and denoising.

Remark 1. Compared with Refs. [42, 50–53], the advantages of our fusion method include three folds. First, the guided filter decomposes image into the noisy small-scale detail, the noiseless large-scale detail and the smooth base layer. It preserves important edges and textures to overcome the blurring of the fused image compared with the mean filter. Second, the adaptive fractional variation model is used to fuse and denoise detail layers sufficiently. The fractional-order derivative of this model is able to reduce blocks and keep textures. And the basic mathematical properties are proofed to solve the model directly. Third, the saliency map and the fractional gradient information are useful for fusing base layers completely. The results fused by our algorithm suppress noise effectively and fuse significant information fully. The decomposition step in this method directly affects the quality of fused image. Although the guided filter keeps edge information effectively, the suitable selection of \( r \) and \( \lambda \) in (5) has some difficulties in the real-word application. When source images contain a lot of noise, the improper selection of \( r \) and \( \lambda \) in (5) may bring artefacts in fused images. The selection of parameters in the image decomposition step brings some restraints in our method. And the truncation of fractional-order derivative operator may affects the fusion result. Table 3 provides the advantages and disadvantages of the different medical fusion methods.

TABLE 3  The advantages and disadvantages of the different medical fusion methods

| Method | Suppressing noise | Keeping edges | Reducing blocks | Suppressing blurring | Overcoming parameters |
|--------|-------------------|---------------|----------------|----------------------|-----------------------|
| DWT    | ×                 | ×             | ×              | ✓✓✓✓                | ✓✓✓✓                  |
| CNN    | ×                 | ✓✓            | ✓✓             | ✓✓✓✓                | ✓✓✓✓                  |
| AFM    | ×                 | ✓✓            | ✓✓             | ✓✓✓✓                | ✓✓✓✓                  |
| ASAFTV | ✓✓                | ✓✓            | ×              | ✓✓✓✓                | ✓✓✓✓                  |
| WTV    | ✓✓                | ✓✓            | ×              | ✓✓✓✓                | ✓✓✓✓                  |
| Ours   | ✓✓✓✓              | ✓✓            | ✓✓             | ✓✓✓✓                | ✓✓✓✓                  |

Note. ✓ and × represent with and without corresponding advantages, respectively.

4  | CONCLUSION

This paper proposes a novel model for medical images fusion and noise suppression. This model is based on multi-scale decomposition which is beneficial for fusing medical images. The fractional-order TV model is built to process the noisy small-scale details layers, and it suppresses noise and fuses detail information at the same time. The proposed model overcomes the shortcoming of traditional inter-order TV to fuse images. It preserves details of the source noisy images and makes full use of the information of different multi-modal images to obtain a comprehensive image. The experiments show that our method fuses the functional and anatomical information fully and suppress noise effectively according to quantitative measures and visual analysis. And the adaptive fractional-order model keeps more textures and decreases blocky effects than integer one.

ORCID

Xuefeng Zhang  https://orcid.org/0000-0002-2831-5747

REFERENCES

1. Vishwakarma, A., Bhuyan, M.K.: Image fusion using adjustable nonsubsampled shearlet transform. IEEE Trans. Instrum. Meas. 68(9), 3367–3378 (2019)
2. Yang, Y., et al.: Multiple visual features measurement with gradient domain guided filtering for multisensor image fusion. IEEE Trans. Instrum. Meas. 66(4), 691–703 (2017)
3. Niu, Y., et al.: Multi-modal multi-scale deep learning for large-scale image annotation. IEEE Trans. Image Process. 28(4), 1720–1731 (2019)
4. Yin, M., et al.: Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain. IEEE Trans. Instrum. Meas. 68(1), 49–64 (2019)
5. Aglika, S., et al.: Infrared and visible image fusion for face recognition. Proc. SPIE Int. Soc. Opt. Eng. 5404, 585–596 (2004)
6. Simone, G., et al.: Image fusion techniques for remote sensing applications. Inf. Fusion 3(1), 3–15 (2002)
7. Xu, Z.P.: Medical image fusion using multi-level local extrema. Inf. Fusion 19, 38–48 (2014)
8. Kaur, G., Singh, S., Vig, R.: Medical fusion framework using discrete fractional wavelets and non-subsampled directional filter banks. IET Image Proc. 14(4), 658–667 (2020)
9. Ke, Q., et al.: Adaptive independent subspace analysis of brain magnetic resonance imaging data. IEEE Access 7, 12252–12261 (2019)
10. Kim, M., Han, D.K., Ko, H.: Multimodal brain tumor classification using deep learning and robust feature selection: A machine learning application for radiologists. Diagnostics 10, 565 (2020)

11. Li, H.F., et al.: Multifocus image fusion and denoising scheme based on homogeneity similarity. Opt. Commun. 285, 91–100 (2012)

12. Li, H.F., et al.: Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning. Pattern Recognit. 79, 130–146 (2018)

13. Li, S.T., et al.: Pixel-level image fusion: A survey of the state of the art. Inf. Fusion 33, 100–112 (2017)

14. Shabanazadeh, F., Khateri, M., Liu, Z.: MR and PET image fusion using non-parametric Bayesian joint dictionary learning. IEEE Sens. Lett. 3(7), 1–4 (2019)

15. Kim, M., Han, D.K., Ko, H.: Joint patch clustering-based dictionary learning for multimodal image fusion. Inf. Fusion 27, 198–214 (2016)

16. Li, S., Kang, X., Hu, J.: Image fusion with guided filtering. IEEE Trans. Image Process. 22(7), 2864–2875 (2013)

17. Zha, J., et al.: Multiscale infrared and visible image fusion using gradient domain guided image filtering. Infrared Phys. Technol. 89, 8–19 (2018)

18. Shankar, K., et al.: Optimal feature level fusion based ANFIS classifier for brain MRI image classification. Concurr. Comput. Pract. Exper. 32(1), e4887 (2020)

19. Ullah, H., et al.: Multi-modality medical images fusion based on local-features fuzzy sets and novel sum-modified-Laplacian in non-subsampled shearlet transform domain. Biomed. Signal Process. Control. 57, 101724 (2020)

20. Liu, Y., et al.: Deep learning for pixel-level image fusion: Recent advances and future prospects. Inf. Fusion 42, 158–173 (2018)

21. Greenberg, S., et al.: Fusion of visual salience maps for object acquisition. IET Comput. Vision 14(4), 113–121 (2020)

22. Liu, C.H., Qi, Y., Ding, W.R.: Infrared and visible image fusion method based on saliency detection in sparse domain. Infrared Phys. Technol. 83, 94–102 (2017)

23. Cheng, B., Jin, L., Li, G.: Adaptive fusion framework of infrared and visual image using saliency detection and improved dual-channel PCNN in the LNSIT domain. Infrared Phys. Technol. 92, 30–43 (2018)

24. Durga, D.B., Ravindra, D.: Two-scale image fusion of visible and infrared images using saliency detection. Infrared Phys. Technol. 76, 52–64 (2016)

25. Liu, X.B., Mei, W.B., Du, H.Q.: Structure tensor and nonsubsampled shearlet transform based algorithm for CT and MRI image fusion. Neurocomputing 235, 121–139 (2017)

26. Pu, Y.F., Wang, W.X.: Fractional differential masks of digital image and their numerical implementation algorithms. Acta Autom. Sin. 33(11), 1128–1135 (2007)

27. Tao, R. Deng, B., Wang, Y.: Research progress of the fractional Fourier in signal processing. Sci. China Ser. F Inf. Sci. 49, 1–25 (2006)

28. Xu, L., Zhang, F., Tao, R.: Fractional spectral analysis of randomly sampled signals and applications. IEEE Trans. Instrum. Meas. 66(11), 2860–2881 (2017)

29. Geng, L.L., et al.: Fractional-order sparse representation for image denoising. IEEE/CAA J. Autom. Sin. 5(2), 555–563 (2018)

30. Pu, Y.F., et al.: A fractional-order variational framework for retinex: Fractional-order partial differential equation-based formulation for multiscale nonlocal contrast enhancement with texture preserving. IEEE Trans. Image Process. 27(3), 1214–1229 (2017)

31. Chmollne, A.: An algorithm for total variation minimization and applications. J. Math. Imaging Vis. 20, 89–97 (2004)

32. Wang, Y., et al.: A new alternating minimization algorithm for total variation image reconstruction. SIAM J. Imag. Sci. 1(3), 248–272 (2008)

33. Ludusan, C., Laviale, O.: Multifocus image fusion and denoising: A variational approach. Pattern Recognit. Lett. 33(10), 1388–1396 (2012)

34. Li, H.F., Yu, Z.T., Mao, C.L.: Fractional differential and variational method for image fusion and super-resolution. Neurocomputing 171, 138–148 (2015)

35. Liu, P.F., Xiao, L., Li, T.: A variational pan-sharpening method based on spatial fractional-order geometry and spectral-spatial low-rank priors. IEEE Trans. Geosci. Remote Sens. 56(3), 1788–1802 (2018)

36. He, K., Sun, J., Tang, X.: Guided image filtering. In: European Conference on Computer Vision, vol. 6311, pp. 1–14 (2010)

37. Zhang X.F., Chen, Y.Q.: Admissibility and robust stabilization of continuous linear fractional order systems with the fractional order < α < 1 case. ISA Trans. 82, 42–50 (2018)

38. Li, D.Z., et al.: Adaptive fractional-order total variation image restoration with split Bregman iteration. ISA Trans. 82, 210–222 (2017)

39. Podlubny, I.: Fractional differential equations. Mathematics in Science and Engineering. Academic Press, New York (1999)

40. Achanta, R., et al.: Frequency-tuned salient region detection. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1579–1604 (2009)

41. Wang, W.W., Shui, P.L., Feng, X.C.: Variational models for fusion and denoising of multifocus images. IEEE Signal Process. Lett. 15, 65–68 (2018)

42. Chen, D.L., Chen, Y.Q., Xue, D.Y.: Fractional-order total variation image denoising based on proximity algorithm. Appl. Math. Comput. 257, 537–545 (2015)

43. Zhang J.P., Chen, K.: A total fractional-order variation model for image restoration with non-homogeneous boundary conditions and its numerical solution. SIAM J. Imag. Sci. 8(4), 1–26 (2015)

44. Mei, J.J., Dong, Y.Q., Huang, T.Z.: Simultaneous image fusion and denoising by using fractional-order gradient information. J. Comput. Appl. Math. 351, 212–227 (2019)

45. Zhang, J.X., Yang, G.H.: Prescribed performance fault-tolerant control of uncertain nonlinear systems with unknown control directions. IEEE Trans. Autom. Control 62(12), 6529–6535 (2017)

46. Zhang, J.X., Yang, G.H.: Low-complexity tracking control of strict-feedback systems with unknown control directions. IEEE Trans. Autom. Control 64(12), 5175–5182 (2019)

47. Zhang, J.X., Yang, G.H.: Fault-tolerant output-constrained control of unknown Euler-Lagrange systems with prescribed tracking accuracy. Automatica 111, 108606 (2020)

48. Immerkar, J.J.: Fast noise variance estimation. Comput. Vision Image Understanding 64(2), 300–302 (1996)

49. Image Fusion Datasets. [Online]. https://sites.google.com/view/durgaprasadbavirisetti/datasets/ (2020)

50. Pu, T., Ni, G.G.: Contrast based image fusion using discrete wavelet transform. Opt. Eng. 39(8), 2075–2082 (2000)

51. Liu, Y., et al.: Multi-focus image fusion with a deep convolutional neural network. Inf. Fusion 36, 191–207 (2017)

52. Azarang, A., Ghassemian, H.: Application of fractional-order differentiation in multispectral image fusion. Remote Sens. Lett. 9(1), 91–100 (2017)

53. Zhao, W.D., Lu, H.C.: Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation. IEEE Trans. Instrum. Meas. 66(9), 2283–2294 (2017)

54. Estevez, P., et al.: Normalized mutual information feature selection. IEEE Trans. Neural Networks 20(2), 189–201 (2009)

55. Xydeas, C.S., Petrovic, V.S.: Objective image fusion performance measure. Electron. Lett. 36(4), 308–309 (2000)

56. Zhang L., et al.: FSIM: A feature similarity index for image quality assessment. IEEE Trans. Image Process. 20(8), 2378–2386 (2011)

57. Wang, Z., Simoncelli, E.P., Bovik, A.C.: Multiscale structural similarity for image quality assessment. In: Proc IEEE Asilomar Conference on Signals, pp. 1398–1402 (2003)

58. Han, Y., et al.: A new image fusion performance metric based on visual information fidelity. Inf. Fusion 14, 127–135 (2013)