ORIGINAL RESEARCH

Multi-modality medical image fusion based on guided filter and image statistics in multidirectional shearlet transform domain

Ayush Dogra[1] · Sanjeev Kumar[1]

Abstract

Due to varying imaging principles and complexity of human organ structures, single-modality image can only provide limited information. Multimodality image fusion is the technique which integrates multimodal images into a single image which improves the quality of images by retaining significant features and helps diagnostic imaging practitioners for accurate treatment and evaluation of medical problems. In the current prevailing image fusion techniques presents numerous challenges including the prevelance of fusion artefacts, design complexity and high computational cost. In this paper, a novel multimodal medical image fusion method has been presented to address these problems. The proposed approach is based on combination of guided filter and image statistics in shearlet transform domain. The multimodal images are subjugated to image decomposition using shearlet transform that captures textures information of original images in multidirectional orientations and then decompose these paired images in low-and high-frequency coefficients (i.e. base and detail layers). Then guided filter with high epsilon value is used to obtain weights of original paired images. These weights are then added to the base layer to obtained unified base layers. A guided image filter and image statistics fusion rule is used to fuse base layers to obtain a fused base layer in covariance matrix and Eigen values are computed to figure out the significant pixels in the neighborhood. Similarly, a choose max fusion rule is used to fuse the detail layers for reconstruction. A unified fused base and detail layers are merged together to obtain final fusion result using inverse shearlet transform. The proposed method is evaluated using medical image datasets. Experimental result demonstrates that our proposed algorithm exhibits promising results and outperforms other prevailing fusion techniques.

Keywords Shearlet transform · Guided filter · Binary weight maps · Image fusion · Performance metrics · Medical imaging

1 Introduction

Numerous medical imaging modalities are available nowadays which capture images via different sensory systems that target mainly on tissues or organ details. Computed tomography (CT) imaging captures skeletal structure and other third-party implant, whereas magnetic resonance imaging (MRI) identifies internal anatomy structural details such as pancreas, liver, belly and other soft tissues, but is less competent to capture dense structures than CT imaging.

Furthermore, operational imaging system such as positron emission tomography (PET) and single photon emission computerised tomography (SPECT) are generally employs to capture a metabolism details of organism that aids for detection of tumours and diagnosis of vascular diseases. A functional image is expressed in pseudo colour and their low spatial resolution often makes analysis difficult. Therefore, radiologists or physicians must examine multimodal medical images independently for proper diagnosis which is inconvenient in terms of accuracy and time. Numerous image processing approaches are implemented in recent years to enhance the quality of source medical images (Połap 2019; Połap and Srivastava 2021; Amin et al. 2020a, b; Rajinikanth et al. 2021; Albahli et al. 2021). Nowadays multi-modality medical image fusion is plying a pivotal role for improving image quality obtained from medical imaging system to have adequate information about numerous tissues and human organs. Image fusion can be used to solve this problem

✉ Ayush Dogra
  ayush123456789@gmail.com; ayush.dogra@csio.res.in

  Sanjeev Kumar
  virdi205@csio.res.in
  https://icsio.csio.res.in/csio.ems/eprofile.aspx

[1] CSIR-Central Scientific Instruments Organisation (CSIO), Chandigarh, India

because it gathers necessary details from multimodal image and fuses it so that the information received by humans or machines are superior to the original images. It can be used to diagnose diseases, plan treatments and perform operation effectively.

Numerous fusion techniques are discussed in current surveys in terms of fusion of medical images (Tawfik et al. 2021; Faragallah et al. 2021). Fusion of medical images is currently playing a pivotal role recently as it helps physicians and radiologists to accurately identify the medical problems of the patients and save time as well. Traditional image fusion techniques are implemented in spatial domains, frequency domains and fuzzy logic. The majority of image fusion techniques performed in multiscale domains with source images first being convert into multiscale elements. Then multiscale components are combined using various fusion process. Finally, inverse transform is employed to obtain the fusion results (Khare et al. 2021; Nair and Singh 2021; Kurban 2021; Ullah et al. 2020). The fundamental steps of multiscale transform analysis are: (a) translate the input images into multiscale transform domain to obtain alternative multiscale feature representation (b) fuse these multiscale features using different techniques and (c) restore the result using inverse transform. These multiscale transform-based algorithms often produce quality visual performance than other techniques and also more efficient in case of other traditional methods. But these algorithms still need to explore to a great extant to obtain best quality results. Motivated by the often adaptation of these methods, we tried to use the same technique in our method to perform fusion process. The detail study of prevailing techniques has been discussed in Sect. 2.

This paper introduces a novel multimodal medical image fusion approach based on combination of guided filtering and image statistics in shearlet transform domain. Shearlet was introduced with the stated intent of providing a highly effective image representation with edges. In fact, Shearlet representations is made up of a combination of well localized waveforms, high anisotropic shapes, range at many locations, scales and orientations. Thus, the Shearlet representation is particularly well suited to represent the edges and other anisotropic objects that are common features in natural images. Further guided filter is a verity of smoothing filter that preserve edges of the images. This filter can also be used to remove noise or texture while keeping the sharp edges. We first decompose input paired images into base and detail using Shearlet transform. A guided image filter and image statistics (GFS) fusion technique is employed to fuse base layers to obtain a unified base layer in which covariance matrix and then eigen values are calculated to figure out the significant pixels in the neighborhood. The guided filter with high epsilon value is used to generate weights of input paired images. These weights are then applied to base layers.

Similarly, to fuse detail layers, a choose max fusion rule is utilized so that the detail layer can be reconstructed. Lastly, an inverse shearlet transform is used to sum of unified based and detail layer to obtain final fusion result. Experimental results on medical image datasets shows that the strategies using Shearlet transform with guided filtering benefits image fusion method more effectively as compares to other algorithms to recover subsequent intracranial details, tissues and combining structural details into a final fusion result.

The key technical contributions of the proposed work consist of following three aspects:

- This method utilizes shearlet transform which is multiscale and multidirectional configuration which has anisotropic properties. Therefore, it exhibits an ability to detect directionality which is an advantage over another traditional wavelet transform.
- The proposed method uses guided filter with high epsilon value that generate optimal approximations of the source images which is later utilised to generates weights.
- We introduce two fusion technique to fuse low- and high-frequency components separately. The fusion rule is used where covariance matrix and eigen values are computed to figure out the important pixels in the neighbourhood. The computed weights are then later added to low frequency coefficients (base layers) and obtained unified base layers.

The remaining of our study is structured as follows: a Sect. 2 illustrates the benchmark to recent prevailing image fusion algorithms. Our proposed method for medical image fusion is illustrated in Sect. 3. In Sect. 4 briefly discusses the performance evaluation metrics used in this paper. Experimental results and detailed analysis have been addresses in Sect. 4. In section discusses the conclusion of our study.

2 Related work

Image fusion has recently gained popularity in a variety of fields i.e. medical, multiple exposure, visible and infrared image fusion and so forth (Tawfik et al. 2021; Liu et al. 2020). There are significant benchmarks fusion strategies that are available in the literature. Naidu (2010) which uses multiresolution discrete cosine transform (m-DCT) algorithm to implement an image fusion that generate a single composite image having adequate information than the multiple source images. The effectiveness of this method is compares with other benchmark fusion techniques using wavelet. It is computationally easy to implement and could be useful for real time applications. This method improves the fused image making them high noise resistant but technique possess high computational cost. Naidu (2011) further

proposes an improved version of fusion technique known as multiresolution single value decomposition (*m*-SVD) to enhance fusion results. It is moreover similar to *m*-DCT but this method does not possess fixed set of basic vectors like fast fourier transform (FFT), discrete wavelet transform (*DCT*) and wavelets etc. and its basic vectors rely on the data set. Rodriguez-Sánchez et al. (2011) introduced the fusion approach known as "Attention fusion" (ATF). This multiresolution technique employs the attention maps for determining the level of activities of each one of the coefficients and to construct the fusion rules. The multiresolution decomposition is performed using dual tree complex wavelet transform. The method demonstrates best performance across several s of images. Wang and Chang (2011) introduced an easy and effective multifocus fusion of images system based on a multiresolution signal decomposition method known as Laplacian pyramid technique. A fusion result is retrieved using inverse Laplacian pyramid transform. The set of images are used to validate the fusion method and shows that proposed method generates better results and produces good performance. However, this method is weak on capturing image details and fails to maintain local properties of input images. Kumar (2013) presents a discreet cosine harmonic wavelet transform (DCHWT) based fusion on by keeping an optimal visual result of fusion image and better quality while reduces the mathematical complexness. An effectiveness of DCHWT is compares with convolutional and lifting related fusion methods. The result of DCHWT is shown equivalent to convolution related wavelets and better than or same as lifting related wavelets.

Li et al. (2013) proposed a faster and efficient fusion technique where two-level decomposition is performed to generates base and detail layers where base layer contains large scale variation in magnitude and detail layer captures small-scale information of the images. To fully utilise spatial frequency to fuse base and detail layers, a unique guided filtering (GF) related weighted average method has been introduced. An experimental output indicate that this technique obtains better performance for fusing multispectral, multifocus, multimodal and multi-exposure datasets. Hui et al. proposes a fusion technique related to wavelet and block dividing (WBD) (Liu and Wang 2013). It performs discrete wavelet transform (DWT) on each input images in initial stage. This method proposes to use block process instead of pixel process of low frequency coefficients. Since traditional image fusion techniques are mostly centred on the fusion process of high frequency coefficients related to single pixel. Leads to serious ringing effect and decreases the visual effect of fused image. After wavelet Transform, an energy in image is focused in low frequency section and a multifocus image has the feature that the most of adjacent pixels are either the clear region or blur region. Vijayarajan and Muttan (2015) proposes a fusion technique based on

principal component averaging (PCA) in discrete wavelet transform (DWT) for fusing of CT-MRI and MRI slices. This technique does not loose popularity due to its conceptual simplicity. This technique includes the benefits of wavelet transform in PCA fusion in the form of eigen values of multiscale representation. Kumar (2015) proposed another method to fuse multisensor images using weighted average technique by utilizing the weights calculated from the detail images which is obtained from the input images utilizing cross bilateral filter (CBF). The outputs of this technique have been using different sets of multisensor and multifocus images visually and quantitatively. Liu et al. (2015) proposes a dense scale invariant features transform (DSIFT) image fusion technique for multifocus images. The key contribution of their work is that they demonstrate the enormous capabilities of image local features. In addition, the local feature descriptor can be utilised to not only quantify activity levels but it also to match misregistered pixel from different input images to improve the performance of fusion result. For multifocus fusion, Bai et al. (2015) introduces a quadtree- based technique where input images are disintegrated in block with ideal sizes in a quadtree structure using a proposed technique. The source images are decomposing into blocks with proper size in the quadtree framework. The concentrated regions are detected in this tree structure utilizing a weighted focus measure named as the sum of weighted modified Laplacian. Experimental results reveal that this technique produces good performance.

In the case of visible and infrared sensor images, Bavirisetti and Dhuli (2016a) proposes an edge preserving fusion approach where the input image is initially decomposes into detail and base layer utilizing anisotropic diffusion. Further, Korhonen Leove transform with weighted linear super-position are employed to compute the resultant detail and base layers. The linear combination of final base and detail yields a fused image. Further in order to fuse medical images, another important structure extraction method via structure preserving filtering was proposed by Bavirisetti and Dhuli (2016b) related to saliency detection (SD) and image decomposition on two scales. This approach is advantageous since the proposed visual saliency extractions procedure effectively highlights a saliency details of input images. Image fusion for colour images was introduced by Paul et al. (2016) related to blending the gradients of luminance component of input image utilizing highest gradient intensity at each pixel position and generate fused luminance utilizing Haar-wavelet related recontraction approach. Further, Ma et al. (2016) proposes a gradient transfer fusion (GTF), a fusion method related to gradient transfer and total variation (TV) to maintain the thermal radiation and appearance information simultaneously. This method generalizes a technique to fuse image pairings without preregistration considerably expands its applications since high-precision registration of

multisensor data is extremely difficult. Bavirisetti and Xiao (2017) proposed a fourth-order partial differential equation (FPDE) and PCA related fusion algorithm. A FPDE is used for the first time in context of fusion of images. Standard fusion datasets are utilized in the experiments. Further, this method is ideal for real time applications because of its reasonable processing time, simple and effective implementation procedures. On the basis of guided image filtering and image statistic (GFS), Bavirisetti et al. (2017) presented a weighted average fusion rule to fuse brain CT and MRI datasets. The guided image filter id used to extract detail layers from each source image. Image statistics are used to generate weights corresponds to each input image from the detail layers. A weighted average fusion method is employed to combine input image details into a single image. Ma et al. (2017) proposes a multiscale fusion method related to visual saliency map with weighted least squares optimisation with the goal of overcoming some of the shortcomings of traditional methods. In comparison to traditional multiscale decompositions, this decomposition has the special property of keeping scale-specific details while minimising halos near edges. The fusion result information seems to be more natural and appropriate for visual perception.

Zhan et al. (2017) proposes a fusion technique for various types of multimodal images with fast filtering image fusion (FFIF) in spatial domain. The magnitude of the image gradient is employed to identify the contrasts and sharpness of images and to bridge the gaps and fill holes, a quick morphological closing operation is conducted on image gradient magnitude. Further a quick structure-preserving filter is used to filter the weight map which is derived from the multimodal image gradient magnitude and a weighed-sum rule is employed to reconstruct the fusion result. The result indicate that this approach performs fast fusion operation and generate high performance than other previous algorithms. Li et al. (2018a) proposes a fusion technique related to structure aware (SA) which utilizes a low-complexity technique for solving the issue of multimodal fusion of images in spatial domain. To fuse medical images, a prominent structure extraction approach and a structure preserving filtering have been introduced. The proposed structure preserving filtering have the property of recovering small-scale information of the guidance images in the surrounding of large-scale structure of source images. Th fused image obtained by merging result of structure preserving filtering with original image related to property of structure preserving filtering. However, image fusion approaches related to multiscale decomposition demonstrate poor contrast and energy loss. In order to over these problems, deep learning is popular tool in the domain of fusion of images in recent years, Wang and Ma (2008) proposed a multichannel pulse coupled neural network (*m*-PCNN) for medical image fusion. The computational model of *m*-PCNN is first explained followed by

a detailed introduction of dual channel model as a special case. This method uses four sets of medical images with several techniques for experimental fields to demonstrate that the *m*-PCNN to cope with multimodality medical images. In order to achieve better fusion results, Liu et al. (2017) uses a deep learning technique attempting to learn a direct mapping between input images and focus map. With that, the mapping is encoded using a deep convolutional neural network (DCNN) trained on superior quality patches and their blurred versions. Li et al. (2018b) proposed an efficient fusion technique that uses a deep learning method to produce a single image that incorporates the attributes of visible and infrared images. The raw images are first decomposed into detail and base parts. A base part is then fused together using weighted averaging. This method employs a deep learning network for extracting multilayer features the detail part. Parvathy et al. (2020) recently proposes a technique based on optimal Shearlet and deep learning. The adequate threshold of fusion rule the shearlet transform (ST) is determined using an enhanced monarch butterfly optimization (EMBO). The extraction element of the deep learning method was then utilised to fuse base and detail layers related to feature maps. The fusion technique was carried out using a restricted Boltzmann machine (RBM). This proposed method was shown to be effective in terms of performance for both visually and quantitatively. Lepcha et al. (2020) recently proposes a medical fusion approach for enhancing medical images using a CBF and rolling guidance filtering (RGF). For scale aware operation, the detail images acquired by subtracting the CBF outputs from source images are processed through the rolling guiding filter. Where it eliminates small scale details while it preserves the other essential features. And the weighted averaged rule and weight normalisation is employed to obtained final fusion result.

Wang et al. (2021) proposes a fusion technique related to non-subsampled Shearlet transform (NSST) and convolutional representation (CSR). The alternating direction product approach is employed to decomposed source images into multiscale and multidirectional sub-images which is subsequently trained to produce several sub-dictionaries. An experiment is conducted on different multimodal brain images such as CT, MRI, PECT and SPECT for validation. However, the sparse based fusion techniques exhibit weak expression capability caused by the single dictionary and the spatial inconsistency. Due to under exposure and bad atmospheric circumstances, night mode visible images are vulnerable to noise and artifacts resulting in a deterred level of details processing and extraction. Dogra et al. (2020) proposes an effective image fusion approach for infrared and visible images for night mode that produces high quality output which largely focuses on the object of interest and is preferable as compares of current existing methods. This technique has a wide scope of applications in the domain
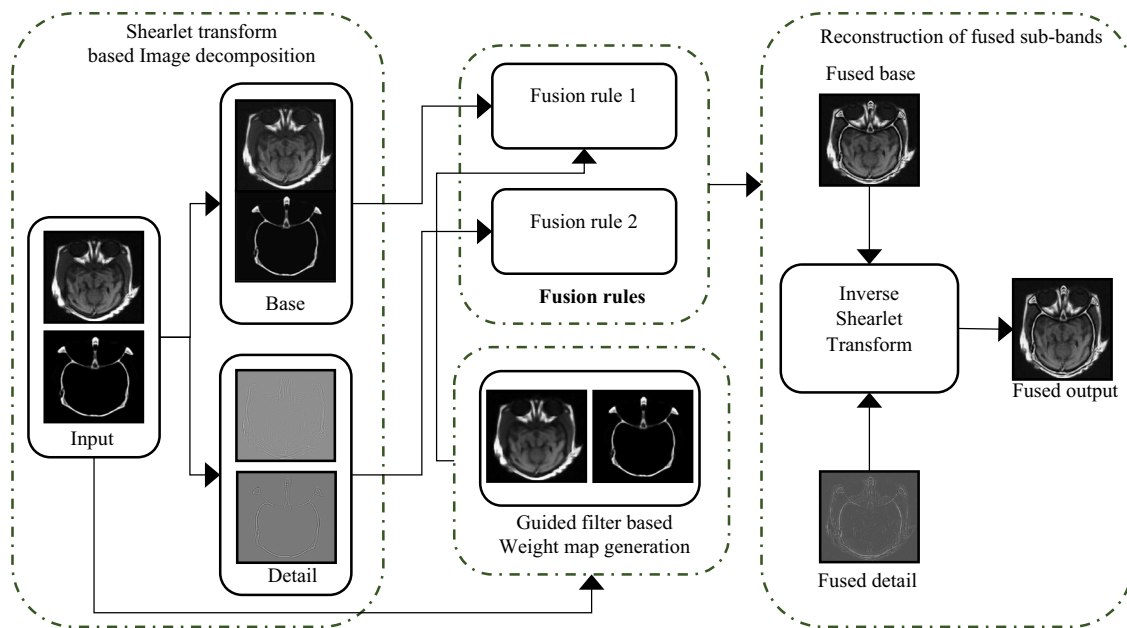
Fig. 1 Framework of the proposed method

of armed forces and surveillance. Nair and Singh (2021) proposes a method known as denoised optimum B-Spline shearlet image fusion (DOBSIF), a unique multimodal fusion of medical images technique related to NSST on real time and standard radiological images. For improving the fusion process, pre-fusion approach is executed with the aid of whale optimization algorithm (WOA) by employing ideal B-spline based registration technique and then a weighted energy fusion process is used to obtain prominent details from the original images. Goyal et al. (2021) proposed a fusion algorithm which fuses low quality multimodal images having low computational complexity to increase target recognition reliability and provide a foundation for clinical applications. Similarly, Jose et al. (2021) proposes a multimodal fusion technique based on to identity search for the NSST for achieving an image optimisation while reducing computational cost and time at the same time. Li et al. (2021) proposes a deep-learning related multimodal fusion algorithm. This approach intended to build a fusion algorithm concept based on supervised deep learning. This approach is appropriate for image fusion to enhance the performance and efficiency of resultant images. Kaur and Singh (2021) have proposed a fusion algorithm using non-subsampling counterlet transform (NSCT) domain. This method first decomposes medical images in the sub-bands. Then, an extreme variant of inception (Xception) is utilised for extracting features from the original images. The appropriate features are selected using multiobjective differential evolution. Experimental results have shown that this method outperforms various recent multimodal fusion algorithms.

3 The proposed fusion algorithm

The proposed medical image fusion technique uses shearlet transform to decomposes input images into low and high frequency coefficients (i.e. base and detail layers). This method further generates weight maps of both source images which are then apply base layers. The fusion rules are used separately to both base and detail layers. Lastly, inverse shearlet transform is employed to obtain final fusion result. The flow-chart of the proposed approach is demonstrated in Fig. 1. In order to understand in better manner, we have demonstrated this algorithm with the help of images through detailed continuity in step by step process. This method contains of mainly four steps; (a) initial step: decomposes input images to generate base and detail layers using shearlet transform (b) second step: obtain weight map of source images using guided filtering and add them to the base layers. (c) third step: fuse base images using GFS fusion rule and detail images using choose max fusion rule (d) final step: use inverse transform to obtained final fusion result. The method is described as follows.

3.1 Decomposition

The original source images CT and MRI, respectively are decompose using Shearlet transform (ST) (Ji 2016) and obtain corresponding low- and high frequency coefficients (i.e. base and detail layers). This approach employs both horizontal and vertical cones. There are two aspects to image

decomposition i.e. decomposition of multi direction (*Kth* directions) and *J* level multiscale wavelet packets.

3.2 Weight map generation

Visual saliency indicates the physical, bottom up distinctness of image information. It is a subjective feature that rely on how visually distinct a detail information is from its surroundings. Saliency is a metric that measures the visual relevance of image features. In the weighted step of multiscale image fusion methods, saliency maps are widely used. Bottom up saliency is computed as local multiscale luminance contrast using frequency tuned filtering (Toet and Hogervorst 2016). For an image *I*. A saliency map *S* is calculated by

$$S(x, y) = \left\| \left| I_\mu - I_f(x, y) \right| \right\| \tag{1}$$

where I_μ is a arithmetic mean image feature vector and I_f is a correspond image pixel vector value a Gaussian blurred variant (utilizing a 5×5 separable binomial kernel), $\| \ \|$ is a L_2 norm (i.e. Euclidian distance and x, y is the pixel coordinates in the Gaussian blurred (utilizing a 5×5 separable binomial kernel) and is a L_2 norm (i.e. Euclidian distance). The above equation allows us to meet all of the conditions for detecting salient regions.

For each source layers X_i and Y_i, $i \in \{0, 1, 2\}$, we construct saliency maps S_{X_i} and S_{Y_i}. The pixelwise maximum of related saliency maps S_{X_i} and S_{Y_i} is then used to generate binary weight maps BW_{X_i} and BW_{Y_i}:

$$BW_{X_i}(x, y) = \begin{cases} 1 & if \quad S_{X_i}(x, y) > S_{Y_i}(x, y) \\ 0 & \quad Otherwise \end{cases} \tag{2}$$

$$BW_{Y_i}(x, y) = \begin{cases} 1 & if \quad S_{Y_i}(x, y) > S_{X_i}(x, y) \\ 0 & \quad Otherwise \end{cases}$$

The resultant binary weight map is noisy and frequently misaligned with object boundaries, resulting in artefacts in output image. The guided filtering (Toet and Hogervorst 2016) of these binary weigh map with associated input layer as guidance image restores spatial consistency:

$$\begin{aligned} W_{X_i} &= GF(BW_{X_i}, X_i) \\ W_{Y_i} &= GF(BW_{Y_i}, Y_i) \end{aligned} \tag{3}$$

As discussed earlier, before guided filter integrates noise reduction and edge preservation, with a result being a scaled variant of the guidance image locally. These qualities are used in the current system to guide the transformation of binary weight map in the smooth continuous weight map using the matching source images as guidance images.

Simultaneously, then these obtained weight maps are applied to the base layers.

3.3 Fusion rules

In image fusion algorithms, the fusion rule plays a vital role. A formation of fused multiscale representation of source images is determined by the fusion rule, which is an important processing step. Since low-frequency coefficients incorporate majority of the data content, a GFS fusion rule (Bavirisetti et al. 2017) has been utilized to fuse base layers as mentioned in Sect. 3.3.1 and to fuse detail layers, a max fusion rule (Panguluri and Mohan 2020) has been utilized as mentioned in Sect. 3.3.2 where detail layers incorporates information about edges.

3.3.1 Fusion strategy for low-frequency coefficients

The GFS fusion rule employs weighted average technique for fusion process. Applying the statistical features, this approach determines optical weight adaptively. The basic proposal is to determine weights corresponding to the pixels in the images based upon its horizontal and vertical edge strengths. However, to determine the weights corresponding to the pixels at the location (x, y) in image takes a square window w of size mxm around its neighbourhood. Considering Z as the matrix and determine its covariance matrix $(cov(Z))$ by considering row as an observation and column as a variable;

$$(cov(Z) = E[(Z - E[Z])(Z - E[Z]^T] \tag{4}$$

Compute unbiased estimate $C_H^{x,y}(Z)$ of the covariance matrix at a pixel (x, y) by

$$C_H^{x,y}(Z) = \frac{1}{m-1} \sum_{j=1}^{m} \left(Z_j - \overline{Z} \right)(Z_j - \overline{Z})^T \tag{5}$$

where Z_j is the *jth* observation of *m*-dimensional variable and $\overline{Z}$ is an average of the observation. $C_H^{x,y}(Z)$ has an interesting diagonal that is the variance vector. Calculate eigen value λ_H^j of $C_H^{x,y}(Z)$. As the matrix is *m*x*m* in size, the number of Eigen value that could be found in *m*. To find the horizontal edge strength α_H. Sum all these eigen values as.

$$\alpha_H(x, y) = \sum_{j=1}^{m} \lambda_H^j \tag{6}$$

Similarly, consider each column as observation and each row as variable to capture vertical edge strength into account. Compute an unbiased estimate $C_V^{x,y}$ and then calculate the eigen value λ_V^j of $C_V^{x,y}$. To find the vertical edge strength α_V, sum all these eigen values as

$$\alpha_V(x, y) = \sum_{j=1}^{m} \lambda_V^j \tag{7}$$

To determine the weight $W(x, y)$ of the pixel at location (x, y), take the sum of $\alpha_H(x, y)$ and $\alpha_V(x, y)$.

$$W(x, y) = \alpha_H(x, y) + \alpha_V(x, y) \tag{8}$$

To assign weights adaptively, repeat this procedure for each and every pixel present in the images. The weight of the pixels is determined on its edge strength rather than its magnitude values in this case. The input image $X(i, j)$ and $Y(i, j)$ are applied to guided filtering as shown in Fig. 1. The procedure is applied until detail layers image are obtained from input images. The source images $X(i, j)$ and $Y(i, j)$ behave as source image and guidance image for guided filtering respectively. In the guidance of image $Y(i, j)$, this filter carries out edge preserving smoothing operation on input image $X(i, j)$. This filter perform structure transferring property to smoothes the source image if two input images are different. $GF_{r,\varepsilon}(X, Y)$ is the Guided filtering operation. The output of $GF_{r,\varepsilon}(X, Y)$ yields the base layer $X_B(i, j)$. Finally, the base layer $X_B(i, j)$ is subtracted from an input image $X(i, j)$ to produce the detail layer $X_D(i, j)$. Further, weights $W_X(i, j)$

and $W_Y(i, j)$ are computed from detail images utilising image statistics. The simple weighted average procedure is used to generate fused image $F(i, j)$ after finding weights of the corresponding input images as given in Eq. 9 as follows,

$$F(i, j) = \frac{X(i, j) * W_X(i, j) + Y(i, j) * W_Y(i, j)}{W_X(i, j) + W_Y(i, j)} \tag{9}$$

3.3.2 Fusion strategy for high-frequency coefficient

The choose max fusion rule is utilized to fuse detail layers as illustrated in Sect. 3.3. The edge details of the images mostly represented by detail layers. A texture of the original images is reflected in the edge information. Let $F_{H1}(X, Y)$ denote detail coefficients obtained from the Shearlet transform of CT image and $F_{H2}(X, Y)$ denote detail coefficients obtained from Shearlet transform of the MRI image. A formula for the max fusion rule is as follows:

$$F_{Max}(X, Y) = \max\left[(F_{H1}(X, Y), F_{H2}(X, Y))\right] \tag{10}$$

The key idea of the max fusion rule is to draw attention to the edge information in the fusion result. Thus, it helps to improve the texture contents of the fused image.

Algorithm: (Pseudocode)

Input: Source images of two modalities $\{X, Y\}$

Output: Fusion result Z

> - Apply the Shearlet transform to decompose both source images into base and detail layers; f_a^{i-1} and f_a^{j-1} into base layers f_a^i, f_a^j and detail layers f_d^j, f_d^j (Ref 47)
> - Compute saliency maps; Saliency maps S_{X_i} and S_{Y_i} for the above base and detail layers by eq. (1)
> - Weights map generation; Generate continuous weight maps W_{X_i} and W_{Y_i} for above saliency maps using guided filtering by eq. (3)
> - Fused base layers; Obtain fused image $F(i, j)$ using simple weighted average technique after finding weights for the corresponding base layers by eq. (9)
> - Fused detail layers; Obtain fused image $F_{Max}(X, Y)$ by choose max fusion rule of detail layers $F_{H1}(X, Y)$ and $F_{H2}(X, Y)$ by eq. (10)
> - Apply weight maps to base layers; Weights maps W_{X_i} and W_{Y_i} are applied to base fused image $F(i, j)$ to obtain unified base layer
> - Combine fused unified base and detail layers; Apply linear combination to $F(i, j)$ and $F_{Max}(X, Y)$
> - Inverse Shearlet transform; Apply inverse ST to obtain final fused image Z

3.4 Inverse shearlet transform

The fused low and high coefficients represented by unified base and detail layers are combined together and reconstructed using inverse Shearlet transform to obtained final fused image

4 Fusion evaluation metrics

The main aim of the image fusion is to preserve uniform details in the images corresponding to the original images as possible and the prevalence of artefacts should be minimum. Certain quantitative metrics are required to demonstrate the effectiveness of our fusion technique. In the literature, numerous evaluation metrics have been proposed including Petrovic Metrics (Petrovic and Xydeas 2005). Following is a quick description of performance evaluation metrics (Ji 2016) utilised in this paper. Consider an input image $f(m, n)$ of size $p * q$.

4.1 Fusion information score

The weighted sum of edge details value quantified for both source images i.e. Q^{AF} and Q^{BF} is used to evaluate total fusion performance $Q^{AB/F}$, in which the weight parameters w^A and w^B indicate perceptual relevance of each source image pixels. The range for $Q^{AB/F}$ is 0 to 1 with 0 indicating absolute source information loss and $Q^{AB/F} = 1$ indicates ideal fusion with no input information loss. The perceptual weight w^A and w^B take the value of respective gradient strength factors g_A and g_B in simplest form.

$$Q^{AB/F} = \frac{\sum_{\forall n,m} Q_{n,m}^{AF} w_{n,m}^A + Q_{n,m}^{BF} w_{n,m}^B}{\sum_{\forall n,m} w_{n,m}^A + w_{n,m}^B} \quad (11)$$

4.2 Fusion loss

The details lost during the fusion procedure is measured by the fusion loss $L^{AB/F}$. Here Q^{AF} and Q^{BF} values less than 1 indicate a direct loss of information; still for proper evaluation of fusion loss, one must be able to discriminate it from fusion artefacts, that result in Q^{AF} and $Q^{BF} < 1$. The $Q^{AB/F}$ approach distinguishes utilizing gradient strength in the inputs and fused images. Thus, each location is classified as follows: F contains artefacts if the gradient strengths in F is greater than in inputs; contrarily, weak gradient in F implies a loss of input details. The overall fusion details is then calculated as the perceptually weighted local fusion loss, specified as $1 - Q^{AF}$ and $1 - Q^{BF}$ for both input A and B, Eq. (12), merged over locations in which the signal gradient in the inputs is stronger as compared to fused image, i.e. in which the $r_{n,m}$ flag is 1, Eq. (13),

$$L^{AB/F} = \frac{\sum_{\forall n,m} r_{n,m} [(1 - Q_{n,m}^{AF}) w_{n,m}^A + (1 - Q_{n,m}^{BF}) w_{n,m}^B]}{\sum_{\forall n,m} w_{n,m}^A + w_{n,m}^B} \quad (12)$$

$$r_{n,m} = \begin{cases} 1, & if g_{n,m}^F < g_{n,m}^A \, or \, g_{n,m}^F < g_{n,m}^B \\ 0, & otherwise \end{cases} \quad (13)$$

4.3 Fusion artifacts

Fusion artefacts $N^{AB/F}$ are visual details injected into the fused image through the fusion procedure which does not correspond to any of the input. Fused artefacts are inherently erroneous data which reduces the subsequent utility of the fusion image and can have major effects in some fusion operations. Fusion artefacts can be analysed using the adopted framework as gradient details which exist in F but not in any of the source image. Local estimate of fusion artefacts, often known as fusion noise, $N_{n,m}$, are currently computed as fusion loss at regions in which fusion gradients are strong than input, Eq. (14). Overall fusion artefacts for the fusion processes $A, B \to F$ is calculated as the perceptually weighted combination of fusion noise estimates across the entire fused image, Eq. (15).

$$N_{n,m} = \begin{cases} 2 - Q_{n,m}^{AF} - Q_{n,m}^{BF}, & if \quad g_{n,m}^F > (g_{n,m}^A \& g_{n,m}^B) \\ 0, & otherwise \end{cases} \quad (14)$$

$$N^{AB/F} = \frac{\sum_{\forall n,m} N_{n,m} (w_{n,m}^A + w_{n,m}^B)}{\sum_{\forall n,m} w_{n,m}^A + w_{n,m}^B} \quad (15)$$

In addition to the above three evaluation metrics, the fused images were compared and assessed by means of mean absolute deviation (MAD) and standard deviation (SD) between original images and fused images. MAD is the sum of absolute differences between the pixel values of source image and the fused image divided by number of observations. It is utilised to measure the standard error of the fused image. Smaller the value of MAD better is the quality. It is given by

$$MAD = \frac{\sum_{i=1}^{n} |A_i - F_i|}{n} \quad (16)$$

where A_i and F_i are the pixel values of the source and fused image, respectively.

Similarly, SD is utilized to depict the contrast of the fused image. A low SD represents that the data points tend to be close to the mean of the set, while a high SD represents that the data points are spread out over a wider range of values. Smaller the value of SD better is the image quality. It is given by.

$$\text{Standard deviation}(\sigma) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \qquad (17)$$

where x_i is the data vector and x are the mean value.

5 Experimental results and analysis

5.1 Experimental setup

Experimental environment consists of hardware components and software platforms. The PC configuration is Intel(R) Core (TM) i7-6700CPU@3.40 GHz and an NVIDIA GeForce GTX 1050-Ti GPU. The experimental configuration equipped with 64-bit windows 10, Caffe, Matlab for deep learning algorithms, the entire network in our setup incorporates the training network and the testing network.

5.2 Image database

The experiments have been carried out using several pairs of medical datasets to validate an effectiveness of our method. However, we have considered and presented only three pairs of CT-MRI medical datasets in this paper as shown in Fig. 2. The datasets are indicated as Dataset1, Dataset2 and Dataset3, respectively. The datasets are obtained from -https://drive.google.com/drive/mobile/folders/0BzXT0LnoyRqlY2d0UTJnb2ZoMk0

5.3 Comparative image fusion methods

Our image fusion algorithm is compared with various prevailing image fusion algorithms such as Anisotropic diffusion (Bavirisetti and Dhuli 2016a), *m*-DCT (Naidu 2010), DCHWT (Kumar 2013), DWT and PCA (Vijayarajan and Muttan 2015), WBD (Liu and Wang Nov. 2013), ATF (Rodriguez-Sánchez et al. 2011), FPDE (Bavirisetti and Xiao 2017), Guided filter (Li et al. 2013), GFS (Bavirisetti et al. 2017), CBF (Shreyamsha Kumar 2015) WLS (Ma et al. 2017), *m*-SVD (Naidu 2011), structure aware (Li et al. 2018a), Saliency detection (Bavirisetti and Dhuli 2016b), *m*-PCNN (Wang and Ma 2008), DSIFT (Liu et al. 2015), GTF (Ma et al. 2016), DLF (Li et al. 2018b), Gradient domain (Paul et al. 2016), DCNN(Liu et al. 2017), Laplacian (Wang and Chang 2011),FFIF (Zhan et al. 2017) and Quadtree (Bai et al. 2015). All of these algorithms have used with default parameters setting given by the concerned researchers.

5.4 Analysis of fusion performance

A key purpose behind image fusion is to preserve uniform details in the images corresponding to the original images
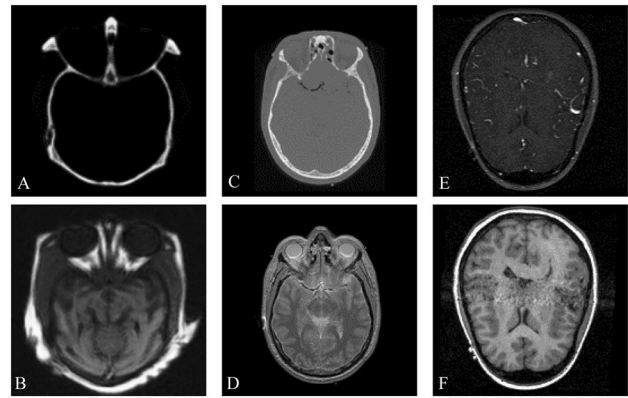

Fig. 2 Source images: Dataset1 (**A, B**), Dataset2 (**C, D**), Dataset3 (**E, F**) (**A, C, E** represents CT images; **B, D, E** represents MRI images)

and the prevalence of artefacts should be minimum. Certain evaluation metrics are required to presents the efficiency of our fusion method. This paper has utilized three objective evaluation metrics (as discussed in Sect. 4) to presents the performance and validity of our such as fusion information score($Q^{AB/F}$), fusion loss, ($L^{AB/F}$) and fusion artefacts($N^{AB/F}$). These three-evaluation metrics are significant and commonly used metrics which provides an in-depth analysis of fusion efficiency. $Q^{AB/F}$ presents total information transfers from input image to fused image. $L^{AB/F}$ presents cumulative loss of information during the fusion process and $N^{AB/F}$ presents noise or artefacts added in the fused image due because of fusion operation. The performance of fused image is appropriate when $Q^{AB/F}$ have high value and $L^{AB/F}$ and $N^{AB/F}$ should possess less values. Here A and B indicate two input image and F indicate resultant fused image. The results of evaluation metrics are presented in Table 1. As it is known that the main idea of image fusion is to incorporates maximum relevant information from both input images into final fused image. However, fusion result cannot be determined purely by seeing the fused image or by estimating objective measures only. It must be evaluated both visually and quantitatively utilising fusion performance measures. This following section demonstrates both visual analyses and objective analysis of different fusion techniques.

5.5 Qualitative analysis

The brain datasets presented in Fig. 2 that are obtained using CT and MRI modalities. As stated earlier, CT images capture bone structures and hard tissues whereas MRI imaging may capture soft tissues in brain. Therefore, by applying the fusion process, it is necessary to combine complete relevant information from different imaging modalities into single image for adequate diagnosis and treatment of diseases

Table 1 Objective image fusion evaluation for (a) Dataset1 (b) Dataset2 and (c) Dataset3

| Method	Dataset1			Dataset 2			Dataset 3		
	$Q^{AB/F}$	$L^{AB/F}$	$N^{AB/F}$	$Q^{AB/F}$	$L^{AB/F}$	$N^{AB/F}$	$Q^{AB/F}$	$L^{AB/F}$	$N^{AB/F}$
Proposed	0.9020	0.0826	0.0146	0.7525	0.2381	0.0094	0.8251	0.1441	0.0308
anisotropic diff (Bavirisetti and Dhuli 2016a)	0.6594	0.3399	0.0006	0.6867	0.3096	0.0037	0.8172	0.1827	**0.0002**
m-DCT (2010)	0.8098	0.1496	0.0406	0.7455	0.2017	0.0528	0.8433	**0.1086**	0.0481
DCHWT (Kumar 2013)	0.8145	0.1820	0.0035	0.7185	0.2782	0.0033	0.8556	0.1384	0.0060
DWT + PCA (Vijayarajan and Muttan 2015)	0.6238	0.3762	*0.0000*	0.7796	0.2204	*0.0000*	0.8348	0.1652	*0.0000*
DLF (Li et al. 2018b)	0.8726	0.2551	0.0126	0.7165	**0.1672**	0.0141	0.8315	0.1662	0.0033
WBD (Liu and Wang 2013)	0.6300	0.3600	0.0100	0.7500	0.2300	0.0200	0.8300	0.1500	0.0200
ATF (Rodriguez-Sánchez et al. 2011)	0.8600	0.1100	0.0300	0.7600	0.2000	0.0400	0.8400	*0.0800*	0.0800
FPDE (Bavirisetti and Xiao 2017)	0.6509	0.3486	**0.0005**	0.6641	0.3352	**0.0007**	0.8125	0.1875	0.0000
Guided filtering (Li et al. 2013)	0.8616	0.1368	0.0016	0.7630	0.2361	0.0009	0.8594	0.1346	0.0060
GFS (Bavirisetti et al. 2017)	*0.9152*	*0.0736*	0.0112	0.7844	0.2100	0.0056	0.8566	0.1258	0.0176
CBF (Shreyamsha Kumar 2015)	0.8826	0.1059	0.0115	0.7531	0.2422	0.0046	0.8479	0.1267	0.0253
DCNN (Liu et al. 2017)	0.8561	0.1761	0.0162	0.7261	0.2561	0.0161	0.8144	0.6162	0.0131
WLS (Ma et al. 2017)	0.8623	0.1235	0.0142	0.7550	0.2346	0.0104	0.8588	0.1273	0.0139
m-SVD (Naidu 2011)	0.7406	0.2583	0.0010	0.7569	0.2369	0.0062	0.7555	0.2424	0.0022
Structure aware (Li et al. 2018a)	**0.9112**	0.0805	0.0083	*0.8360*	*0.1488*	0.0152	0.8652	0.1285	0.0062
Saliency detect (Bavirisetti and Dhuli 2016b)	0.8611	0.1221	0.0168	0.7846	0.1975	0.0178	**0.8686**	0.1172	0.0142
m-PCNN (Wang and Ma 2008)	0.1480	0.8500	0.0020	0.5150	0.4780	0.0070	0.2710	0.7270	0.0020
DSIFT (Liu et al. 2015)	0.8800	0.1100	0.0100	0.8000	0.1900	0.0100	0.8300	0.1700	0.0000
GTF (Ma et al. 2016)	0.2449	0.6810	0.0741	0.4367	0.4509	0.1124	0.5492	0.3837	0.0672
Gradient domain (Paul et al. 2016)	0.8650	0.1330	0.0020	0.8060	0.1900	0.0040	0.8600	0.1300	0.0100
Laplacian (Wang and Chang 2011)	0.9030	**0.0860**	0.0110	**0.8200**	0.1600	0.0200	*0.8800*	0.1100	0.0100
FFIF (Zhan et al. 2017)	0.3169	0.5964	0.0868	0.3199	0.5619	0.1182	0.5386	0.2718	0.1896
Quadtree (Bai et al. 2015)	0.9062	0.0819	0.0119	0.8046	0.1927	0.0027	0.8429	0.1505	0.0065

Italics indicates top best value; bold colour indicates second best value

properly. Figure 3 shows a visual result of various fusion of Dataset1. The representative MRI and CT images are demonstrated in Fig. 2. Figure 3a–w, shows the fused images of other algorithms and Fig. 3x demonstrates a fused image of our algorithm. The result demonstrates that the visual performance and contrast of fused images of Anisotropic diffusion, *m*-DCT, DWT and PCA, FPDE, WBD, DCNN and GTF fusion methods are not up to par resulting in certain visual distortions in the fused images. Through visual analysis, the fusion images of Laplacian, DSIFT, Structure Aware, DLF, ATF and Quadtree method looks good. However as compared to other methods, the proposed method produces a visually enhanced and undistorted images. The contrast between object and the background are negligible and the objects appears blurred. Figure 4 shows the visual comparison of the quality of different techniques for image Dataset2. Figure 4a–w, presents the fused images of other techniques used for comparison. Figure 4x shows a fused image of the proposed technique.

Anisotropic diffusion, WLS, DCHWT, FPDE, WBD, *m*-SVD, and FFIF techniques are unable to integrate overall complimentary information of the paired input images

property as illustrated in Fig. 4 where there is some detail loss and also fusion performance is poor. The DLF, Quadtree, DSIFT, ATF, Structure Aware methods are capable enough to integrate some essential information and generated the visually appealing results. As can be observed in DTNN and DLF, a relatively absolute object region is obtained and the contrast of object background is improved but the background resolution is insignificant. However, when compares to other techniques, the fused image of proposed technique provides visually more details. Figure 5 shows the visual results of different fusion techniques for image Dataset3. Figure 5a–u shows the fusion results of different algorithms utilized for comparison. Figure 5x presents a fusion result of the proposed method. Guided filter, anisotropic diffusion, *m*-DCT, WLS, DCNN, DCHWT, FPDE, WBD, and *m*-SVD algorithms are unable to properly combine all of the complimentary information of paired source images as shown in Fig. 5. When compared to the input images, there is some details loss and the performance of the fusion result is low. Structure Aware, Quadtree, Laplacian, DSIFT, DLF, ATF, FFIF algorithms are capable to retails essential details and able to generate an aesthetically

Fig. 3 Visual quality analysis of numerous fusion techniques on Dataset1: **a** anisotropic diffusion **b** *m*-DCT **c** DCHWT **d** DWT and PCA **e** WBD **f** ATF **g** FPDE **h** guided filter **i** GFS **j** CBF **k** WLS **l** *m*-SVD **m** structure aware **n** saliency detection **o** *m*-PCNN **p** DSIFT **q** GTF **r** gradient domain **s** Laplacian **t** FFIF **u** Quadtree, **v** DCNN **w** DLF and **x** proposed method
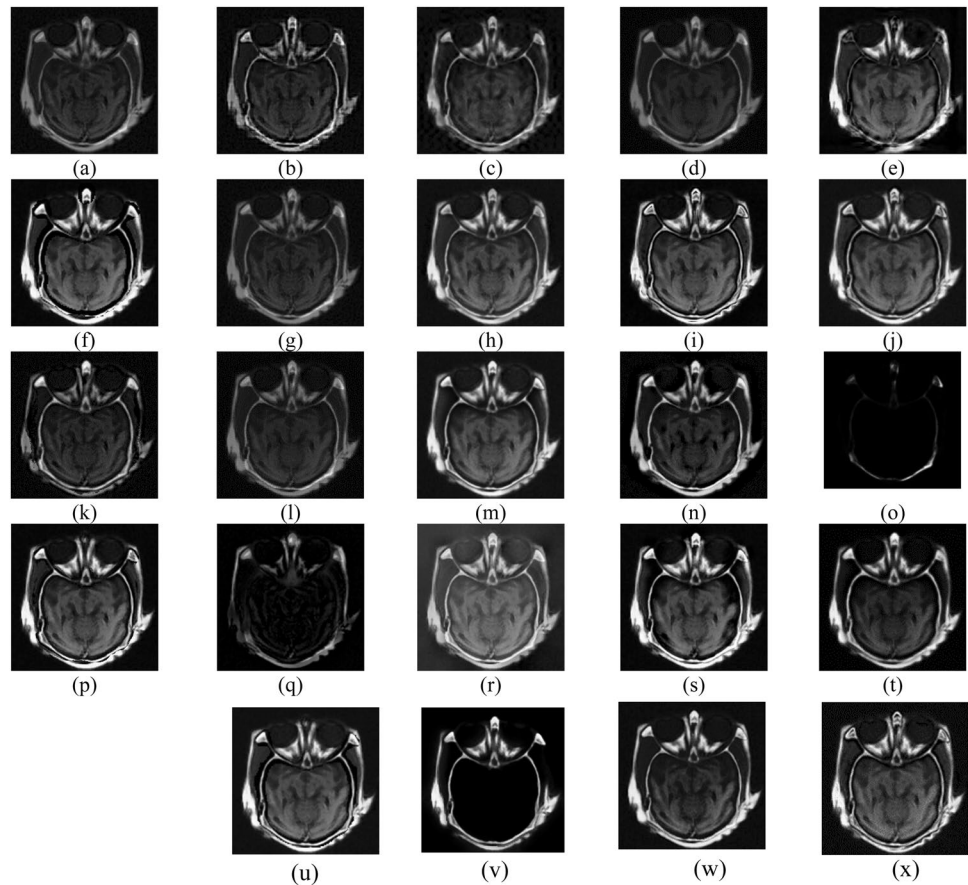

pleasing result. However, when comparing to other algorithms, fusion result of our method provides visually pleasing and more information with excellent contrast. The fused image of our algorithm in terms of all the datasets exhibits superior visual performance and comparable quantitative value as compare to other methods. Also, the experimental result on the different sets of images shows that our method running time is faster than the various benchmark traditional fusion algorithms. Figures 3x, 4 and 5x are the fused image of our algorithm which incorporates better contrast objects and the scene of high-resolution. In terms of visual performance, our method fusion result has more information and extracts high salient information from input images which indicate that our algorithm outperforms prevailing benchmark image fusion methods.

5.6 Quantitative analysis

Our method is quantitatively analysed in contrast to different fusion methods using the fusion metrics $Q^{AB/F}$, $L^{AB/F}$ and $N^{AB/F}$. During the fusion process, $Q^{AB/F}$ indicates total information transfers from input images into fused image, $L^{AB/F}$ indicates total information loss and $N^{AB/F}$ indicates noise or artefacts added during fusion process. Any method

should have higher value for $Q^{AB/F}$ and minimal values in case of $L^{AB/F}$ and $N^{AB/F}$ for better performance. Table 1a demonstrates the quantitative performance of different image fusion techniques as well as the proposed algorithm for image Dataset1. The GTF, *m*-PCNN, and FFIF methods have the lowest value for the statistic $Q^{AB/F}$. The highest $L^{AB/F}$ value is found in GTF, *m*-PCNN and, DCNN, FFIF. Similarly, FFIF, GTF, *m*-DCT and ATF possess large $N^{AB/F}$ metric value. The GFS, Laplacian, Quadtree and the proposed method perform consistently across all measures. However, the proposed technique outperforms all other fusion techniques. Table 1b demonstrates the result of performance metrics of image Dataset2 for various approaches. For the fusion metrics $Q^{AB/F}$, the technique such as FFIF, GTF, and *m*-PCNN have the lowest performance. For DLF, DCNN, FFIF, GTF, *m*-PCNN and Anisotropic diffusion, $L^{AB/F}$ is high. The $N^{AB/F}$ value of FFIF and GTF is the highest. However, for overall fusion performance demonstrates in Table 1a–c, the proposed technique shows consistency, stability and significant performance in all quantitative matrices as comparison to other methods.

Table 1c demonstrates the results of performance evaluation metrics of image Dataset3 for various approaches. For the fusion metrics $Q^{AB/F}$, it can be noted that DCNN,

Fig. 4 Visual performance analysis of different fusion techniques on Dataset2: **a** anisotropic diffusion **b** *m*-DCT **c** DCHWT **d** DWT and PCA **e** WBD **f** ATF **g** FPDE **h** guided filter **i** GFS **j** CBF **k** WLS **l** *m*-SVD **m** structure aware **n** saliency detection **o** *m*-PCNN **p** DSIFT **q** GTF **r** gradient domain **s** Laplacian **t** FFIF **u** Quadtree, **v** DCNN **w** DLF and **x** proposed method
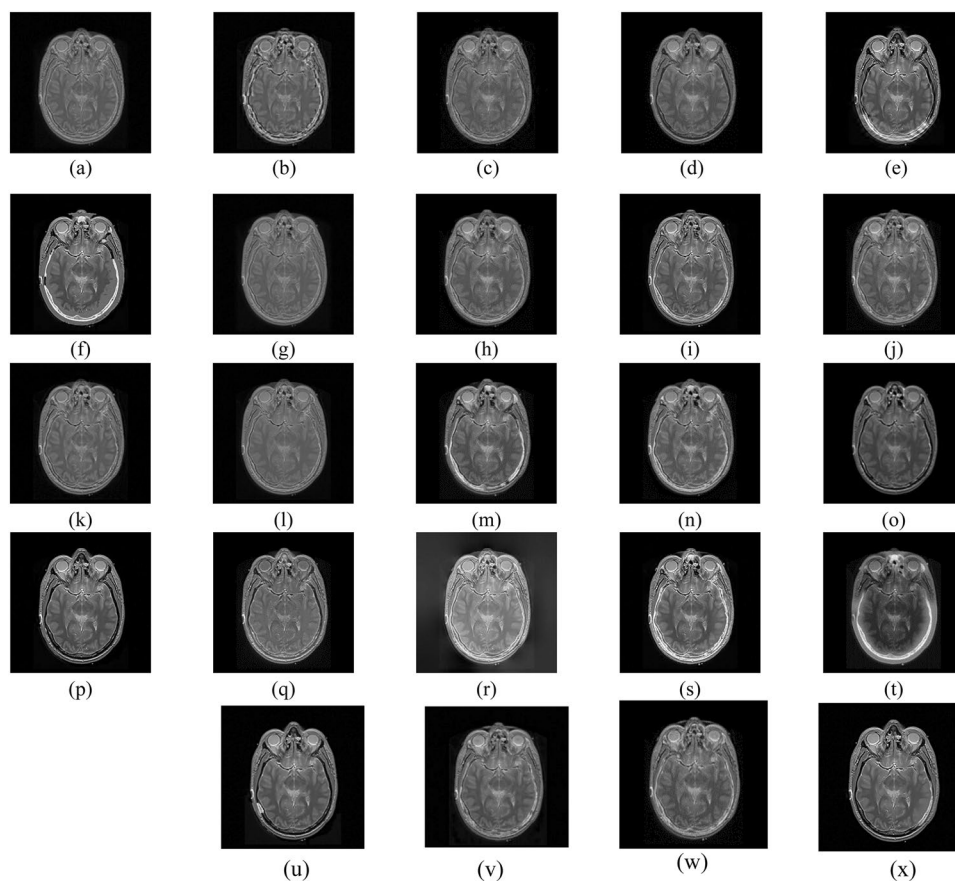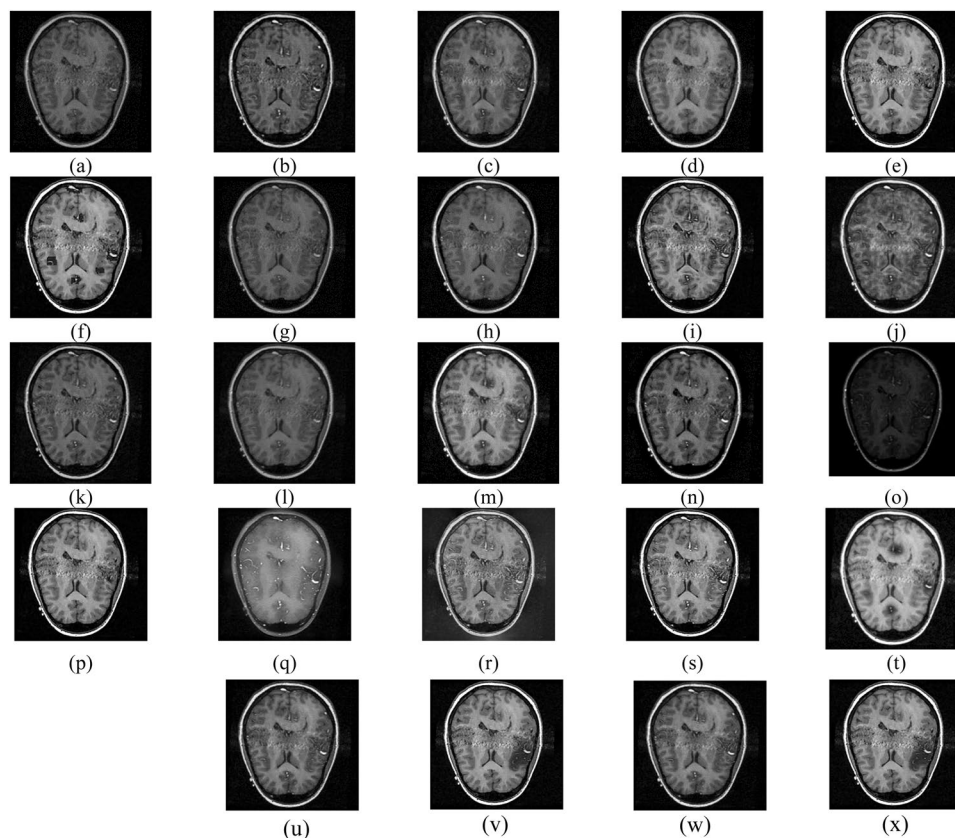

Fig. 5 Visual quality analysis of different fusion techniques on Dataset3: **a** anisotropic diffusion **b** *m*-DCT **c** DCHWT **d** DWT and PCA **e** WBD **f** ATF **g** FPDE **h** guided filter **i** GFS **j** CBF **k** WLS **l** *m*-SVD **m** structure aware **n** saliency detection **o** *m*-PCNN **p** DSIFT **q** GTF **r** gradient domain **s** Laplacian **t** FFIF **u** Quadtree, **v** DCNN **w** DLF and **x** proposed method

Fig. 6 Statistical analysis of fusion performance in case of mean absolute deviation (MAD) and standard deviation (SD) of Dataset1
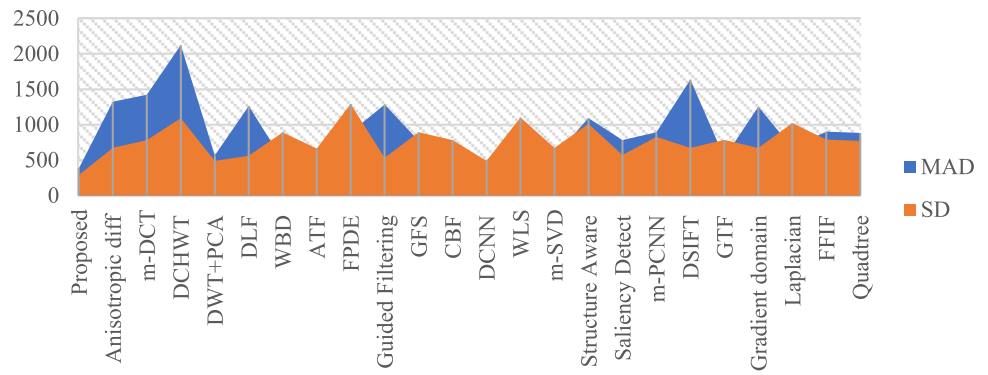

Fig. 7 Statistical analysis of fusion performance in case of mean absolute deviation (MAD) and standard deviation (SD) of Dataset2
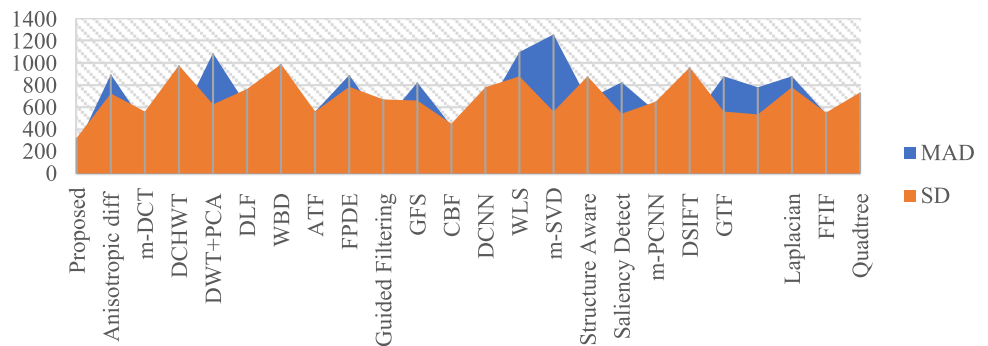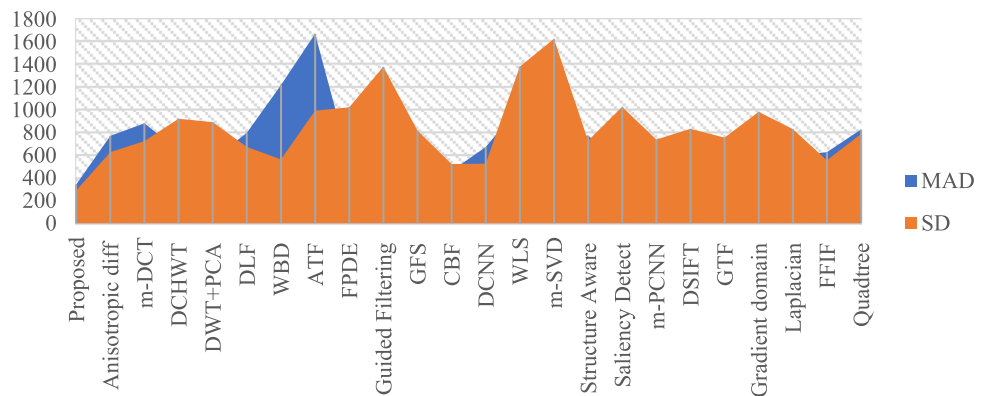

Fig. 8 Statistical analysis of fusion performance in case of mean absolute deviation (MAD) and standard deviation (SD) of Dataset3

Chi-square test of $Q^{AB/F}$ for proposed method				
	Dataset1	Dataset2	Dataset3	Chi Square Test Formula:
Observed (O)	0.9020	0.7525	0.8251	
Expected (E)	1.0000	1.0000	1.0000	$\chi^2 = \sum \frac{(O-E)^2}{E}$
Difference (O-E)	0.0980	0.2475	0.1749	
$(O-E)^2$	0.0096	0.0612	0.0306	
$(O-E)^2/E$	0.0096	0.0612	0.0306	$\sum (O-E)^2/E = 0.1014$
$\chi^2 = 0.1014$,　Degree of freedom (DF) = 2,　$\alpha = 0.05$,　Critical value= 5.99				
$\chi^2_{(2)} = 0.1014$, p < 0.05				
$\chi^2 <$ Critical value				

Fig. 9 Statistical analysis using Chi square test in terms of fusion information score ($Q^{AB/F}$) for all three datasets pairs

Gradient domain, Laplacian, Saliency detect, guided filter, GFS, WLS optimum has a high performance. For FFIF, GTF, *m*-PCNN and FFIF, the value of $L^{AB/F}$ is high and similarly ATF has high $N^{AB/F}$ value. Figures 6, 7 and 8 presents the mean absolute deviation (MAD) and standard deviation (SD) of all the algorithms for three different pairs of images to further analyse the effectiveness of our algorithm.

and to further prove the excellence of our algorithm. In Figs. 6, 7 and 8, we observed that the mean absolute deviation and standard deviation of proposed method is smaller than other state-of-the-art methods. The smaller values of both the parameters indicates that there is an increase in

information and the fusion performance is improved. Hence, the proposed method showed the best fusion performance in terms of image quality.

In order to further substantiate the versatility of the proposed algorithm, it is ideal of the method of be consistent in performance on varied types of datasets. The chi square test is excellent statistical analysis to numerate the difference between the observed and the expected value. Chi-square test explores the statistical relationship between the categorical variables. We have computed chi-square test in terms of fusion information score ($Q^{AB/F}$) for three dataset pairs as shown in Fig. 9. As per the analysis employing chi-square test the value computed is less than the critical value. Therefore, the hypothesis is correct i.e. there is no significant difference between the parametric numeration of the proposed method on categorical datasets used. The analysis further supports the robustness of the proposed algorithm for different variety of datasets. In the context of overall visual performance and fusion metrics, the proposed method shows consistency and comparable performance for all the datasets as compared to other state-of-the-art algorithms as shown by the experimental results and statistical analysis.

6 Conclusions

In this paper, we have presented a novel multimodal medical image fusion method based on combination of guided filter and image statistics in shearlet transform domain in order to fuse MRI and CT image datasets. The first step is to employ shearlet transform which yields base layers and detail layers, then use guided filter with high epsilon value to generate weights of input paired images. These weights are then added to base layers. These base and detail layers are then fuse separately utilizing GFS fusion rule and choose max fusion rule. In the end, inverse Shearlet transform is used to obtained final fused image. Numerous experiments are conducted on the different pairs of multimodal medical images to demonstrate that our proposed image fusion method can retain tissue and structural information satisfactorily than most of the existing benchmark algorithms. Subjectively, our proposed method improves the contrast and brightness of the fused image which is suitable for human visual perception. Three objective evaluation metrics are utilized to validate the performance of the fusion results with other benchmark methods. Experimental findings indicate that our method can adequately retain the edge information within a given range, better display an information of source images and ensure that no additional artefacts or information are added during fusion operation. In addition, our method is able to maintain both edge information of MRI images and structural information of CT images. Our proposed algorithm proposed in this paper can be used in medical diagnosis to

enhance accuracy and efficiency. However, it can be noted that the proposed method has a great scope to be improved further in the future.

Data availability statement The datasets for experiments are obtained from https://drive.google.com/drive/mobile/folders/0BzXT0Lnoy RqlY2d0UTJnb2ZoMk0

References

Albahli S, Rauf HT, Arif M, Nafis MT, Algosaibi A (2021) Identification of thoracic diseases by exploiting deep neural networks. Neural Netw 5:6

Amin J, Sharif A, Gul N, Anjum MA, Nisar MW, Azam F, Bukhari SAC (2020a) Integrated design of deep features fusion for localization and classification of skin cancer. Pattern Recognit Lett 131:63–70

Amin J, Sharif M, Anjum MA, Raza M, Bukhari SAC (2020b) Convolutional neural network with batch normalization for glioma and stroke lesion detection using MRI. Cogn Syst Res 59:304–311

Bai X, Zhang Y, Zhou F, Xue B (2015) Quadtree-based multi-focus image fusion using a weighted focus-measure. Inf Fus 22:105–118

Bavirisetti DP, Dhuli R (2016a) Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen–Loeve transform. IEEE Sens J 16(1):203–209

Bavirisetti DP, Dhuli R (2016b) Two-scale image fusion of visible and infrared images using saliency detection. Infrared Phys Technol 76:52–64

Bavirisetti D, Xiao G (2017) Multi-sensor image fusion based on fourth order partial differential equations. In: 2017 20th international conference on information fusion (fusion). IEEE, pp 1–9

Bavirisetti DP, Kollu V, Gang X, Dhuli R (2017) Fusion of MRI and CT images using guided image filter and image statistics. Int J Imaging Syst Technol 27(3):227–237

Dogra A, Kadry S, Goyal B, Agrawal S (2020) An efficient image integration algorithm for night mode vision applications. Multim Tools Appl 79(15):10995–11012

Faragallah O, El-Hoseny H, El-Shafai W, El-Rahman WA, El-Sayed HS, El-Rabaie E-S, El-Samie FA, Geweid GGN (2021) A comprehensive survey analysis for present solutions of medical image fusion and future directions. IEEE Access 9:11358–11371

Goyal B, Lepcha DC, Bhateja V, Lay-Ekuakille A (2021) Measurement and analysis of multi-modal image fusion metrics based on structure awareness using domain transform filtering. Measurement 182:109663

Ji XX (2016) An improved image fusion method of infrared image and SAR Image via shearlet and sparse representation. In: 2016 8th international conference on intelligent human-machine systems and cybernetics (IHMSC), vol 1. IEEE, pp 74–77

Jose J, Gautam N, Tiwari M, Tiwari T, Suresh A, Sundararaj V, Rejeesh MR (2021) An image quality enhancement scheme employing adolescent identity search algorithm in the NSST domain for multimodal medical image fusion. Biomed Signal Process Control 66:102480

Kaur M, Singh D (2021) Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks. J Ambient Intell Humaniz Comput 12(2):2483–2493

Khare A, Khare M, Srivastava R (2021) Shearlet transform based technique for image fusion using median fusion rule. Multim Tools Appl 80(8):11491–11522

Kumar BS (2013) Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform. SIViP 7(6):1125–1143

Kurban T (2021) Fusion of remotely sensed infrared and visible images using Shearlet transform and backtracking search algorithm. Int J Remote Sens 42(13):5087–5104

L Yu, Chen X, Peng H, Wang Z (2017) Multi-focus image fusion with a deep convolutional neural network. Inf Fus 36:191–207

Lepcha DC, Goyal B, Dogra A (2020) Image fusion based on cross bilateral and rolling guidance filter through weight normalization. Open Neuroimag J 13(1):51–61

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

Li W, Xie Y, Zhou H, Han Y, Zhan K (2018a) Structure-aware image fusion. Optik 172:1–11

Li H, Wu X-J, Kittler J (2018b) Infrared and visible image fusion using a deep learning framework. In: 2018b 24th international conference on pattern recognition (ICPR). IEEE, pp 2705–2710

Li Y, Zhao J, Lv Z, Li J (2021) Medical image fusion method by deep learning. Int J Cogn Comput Eng 2:21–29

Liu Y, Wang Z (2013) Multi-focus image fusion based on wavelet transform and adaptive block. J Image Graphics 18(11):1435–1444

Liu Yu, Liu S, Wang Z (2015) Multi-focus image fusion with dense SIFT. Inf Fus 23(1):139–155

Liu Yu, Wang L, Cheng J, Li C, Chen X (2020) Multi-focus image fusion: a survey of the state of the art. Inf Fus 64:71–91

Ma J, Chen C, Li C, Huang J (2016) Infrared and visible image fusion via gradient transfer and total variation minimization. Inf Fus 31:100–109

Ma J, Zhou Z, Wang B, Zong H (2017) Infrared and visible image fusion based on visual saliency map and weighted least square optimization. Infrared Phys Technol 82:8–17

Naidu VPS (2010) Discrete cosine transform-based image fusion. Def Sci J 60(1):48–54

Naidu VPS (2011) Image fusion technique using multi-resolution singular value decomposition. Def Sci J 61(5):479–484

Nair RR, Singh T (2021) An optimal registration on shearlet domain with novel weighted energy fusion for multi-modal medical images. Optik 225:165742

Panguluri SK, Mohan L (2020) Efficient DWT based fusion algorithm for improving contrast and edge preservation. Int J Adv Comput Sci Appl 11(10):123–131

Parvathy S, Velmurugan SP, Sampson J (2020) A novel approach in multimodality medical image fusion using optimal shearlet and deep learning. Int J Imaging Syst Technol 30(4):847–859

Paul S, Sevcenco IS, Agathoklis P (2016) Multi-exposure and multi-focus image fusion in gradient domain. J Circuits Syst Comput 25(10):1650123

Petrovic V, Xydeas C (2005) Objective image fusion performance characterisation. In: Tenth IEEE international conference on computer vision (ICCV'05) volume 1, vol 2. IEEE, pp 1866–1871

Połap D (2019) Analysis of skin marks through the use of intelligent things. IEEE Access 7:149355–149363

Połap D, Srivastava G (2021) Neural image reconstruction using a heuristic validation mechanism. Neural Comput Appl 33(17):10787–10797

Rajinikanth V, Kadry S, Taniar D, Damaševičius R, Rauf HT (2021) Breast-cancer detection using thermal images with marine-predators-algorithm selected features. In: 2021 seventh international conference on bio signals, images, and instrumentation (ICBSII), pp 1–6. IEEE

Rodriguez-Sánchez R, García JA, Fdez-Valdivia J (2011) From computational attention to image fusion. Pattern Recognit Lett 32(14):1778–1795

Shreyamsha Kumar BK (2015) Image fusion based on pixel significance using cross bilateral filter. Signal Image Video Process 9(5):1193–1204

Tawfik N, Elnemr HA, Fakhr M, Dessouky MI, El-Samie FEA (2021) Survey study of multimodality medical image fusion methods. Multim Tools Appl 80(4):6369–6396

Toet A, Hogervorst MA (2016) Multiscale image fusion through guided filtering. In: Target and background signatures II, vol. 9997, p. 99970J. International Society for Optics and Photonics

Ullah H, Ullah B, Wu L, Abdalla FYO, Ren G, Zhao Y (2020) 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

Vijayarajan R, Muttan S (2015) Discrete wavelet transform based principal component averaging fusion for medical images. AEU Int J Electron Commun 69(6):896–902

Wang W, Chang F (2011) A multi-focus image fusion method based on Laplacian pyramid. JCP 6(12):2559–2566

Wang Z, Ma Y (2008) Medical image fusion using m-PCNN. Inf Fus 9(2):176–185

Wang L, Dou J, Qin P, Lin S, Gao Y, Wang R, Zhang J (2021) Multimodal medical image fusion based on nonsubsampled shearlet transform and convolutional sparse representation. Multim Tools Appl 80(30):36401–36421

Zhan K, Xie Y, Wang H, Min Y (2017) Fast filtering image fusion. J Electron Imaging 26(06):1