

Article

CSID: A Novel Multimodal Image Fusion Algorithm for Enhanced Clinical Diagnosis

Shah Rukh Muzammil ¹, Sarmad Maqsood ², Shahab Haider ¹
and Robertas Damaševičius ^{2,*}

¹ Department of Computer Science, City University of Science and Information Technology, Peshawar 25000, Pakistan; kashafshah904@gmail.com (S.R.M.); dr.shahab39@gmail.com (S.H.)

² Department of Software Engineering, Kaunas University of Technology, Kaunas 51368, Lithuania; sarmad.maqsood@ktu.edu

* Correspondence: robertas.damasevicius@ktu.lt

Received: 11 September 2020; Accepted: 29 October 2020; Published: 5 November 2020



Abstract: Technology-assisted clinical diagnosis has gained tremendous importance in modern day healthcare systems. To this end, multimodal medical image fusion has gained great attention from the research community. There are several fusion algorithms that merge Computed Tomography (CT) and Magnetic Resonance Images (MRI) to extract detailed information, which is used to enhance clinical diagnosis. However, these algorithms exhibit several limitations, such as blurred edges during decomposition, excessive information loss that gives rise to false structural artifacts, and high spatial distortion due to inadequate contrast. To resolve these issues, this paper proposes a novel algorithm, namely Convolutional Sparse Image Decomposition (CSID), that fuses CT and MR images. CSID uses contrast stretching and the spatial gradient method to identify edges in source images and employs cartoon-texture decomposition, which creates an overcomplete dictionary. Moreover, this work proposes a modified convolutional sparse coding method and employs improved decision maps and the fusion rule to obtain the final fused image. Simulation results using six datasets of multimodal images demonstrate that CSID achieves superior performance, in terms of visual quality and enriched information extraction, in comparison with eminent image fusion algorithms.

Keywords: medical image processing; image fusion; multimodal medical image; image decomposition; sparse representation

1. Introduction

Image processing manipulates input source images to extract the maximum possible information. The information obtained is exploited for several applications, including remote sensing, malware analysis, clinical diagnosis, etc. [1–5]. However, the latter requires greater attention as enhanced clinical diagnosis remains the top priority around the world [6]. In this regard, clinical imaging plays a vital role in modern day health care systems, where Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) are among the most extensively used imaging modalities [7–9]. This allows radiologists to analyze the human body and generate different patterns, which are used in clinical analysis [10]. These images provide anatomical statistics [7]; however, the extraction of purposeful functional details from an individual image remains a critical issue. This demands multimodal image fusion, which integrates the complementary information of images from different modalities to produce an enhanced fused image through simulation, thereby providing enriched anatomical and functional information [6,7,11–13].

There are several multimodal image fusion algorithms [7,14,15], which are divided into two major classes, namely the spatial and transform domains [2]. The spatial domain produces fused

images based on the coefficient deduced from pixels, sections, or blocks without transformation [14,16], which produces false structural artifacts in the resulting fused images. Conversely, the transform domain merges the respective transform coefficients and employs the inverse transformation to yield better fused images. To this end, MultiScale Transformation (MST) includes contourlet transform [17], discrete wavelet transform [18], non-subsampled contourlet transform [19], and curvelet transform [20], which are used for multimodal image fusion.

There are a number of the MST-based algorithms. For example, the authors in [18] proposed a novel Discrete Wavelet Transform-based (DWT) algorithm. The proposed algorithm selects the low frequency domain coefficients by employing the maximum sharpness focus measure method, whereas the high frequency sub-band coefficients are chosen on the basis of maximum neighboring energy-based fusion. Similarly, Non-Subsampled Contourlet Transform (NSCT) [19] decomposes the input source images into a series of high frequency sub-bands and one low frequency sub-band. The work in [19] also proposed activity measures for low-pass and high-pass sub-bands to enhance fusion. The sub-bands obtained from the aforementioned process are merged based on the corresponding activity measure. Finally, the inverse NSCT is applied on the merged sub-bands to build a fused image. Furthermore, the Laplacian Pyramid (LP) [21] decomposes the source images into different low-pass filtered images and produces a pyramid structure, where the visual quality of the resulting image remains proportional to the pyramid levels. Moreover, Guided Filtering-based Fusion (GFF) [22] and Convolutional Sparse Representation (CSR) [23] decompose the images into base and detail layers. GFF aims to enhance the spatial consistency of fused images, whereas CSR preserves the detailed information from the input source images. Additionally, Convolutional Sparsity-based Morphological Component Analysis (CSMCA) [24] integrates MCA and CSR into a novel optimization framework to enhance the fusion process. Another algorithm, namely Convolutional Neural Network (CNN) [25] adopts the Siamese convolutional network to create weight maps. These weight maps integrate pixel activity information from the input source images and employ the local similarity-based strategy to adapt fusion modes for the decomposed coefficients. However, the aforementioned algorithms exhibit several limitations, such as the blurring effect near strong edges during image decomposition and increased information loss, which impact the originality of the fused images, thereby producing high spatial distortion.

To resolve the aforementioned issues, we propose a novel algorithm for multimodal image fusion, namely Convolutional Sparse Image Decomposition (CSID), having the following contributions.

1. We employ contrast stretching and the spatial gradient method to extract edges from the input source images.
2. We propose the use of the cartoon-texture decomposition that creates an over-complete dictionary.
3. We propose a modified Convolutional Sparse Coding (CSC) method.
4. Finally, our proposed algorithm uses enhanced decision maps and a fusion rule to obtain the fused image.
5. Additionally, this simulation study reveals that the CSID algorithm achieves superior performance, in terms of visual quality and enriched information extraction, in comparison with other image fusion algorithms, as it will be discussed in Section 5.

The rest of the paper is organized as follows. Section 2 critically reviews the eminent related work on multimodal image fusion. Section 3 details the proposed algorithm. Section 4 presents the objective evaluation metrics. Section 5 evaluates the performance of the suggested algorithm in comparison with the state-of-the-art eminent algorithms. Finally, Section 6 concludes the paper with discussion on future research aims.

2. Related Work

Modern healthcare systems actively use multimodal image fusion for diagnosis [10]. This section critically reviews the eminent work on multimodal clinical image fusion.

Recently, the MST and Sparse Representation (SR) techniques have gained significant popularity in the transform domain and have produced positive results in medical image analysis [26]. However, these methods have shortcomings, such as (i) the “max-1” rule induces spatial inconsistency in a fused image when different modalities are captured from the source images [27], (ii) the MST-based filters used for the SR-based image fusion [28] are time-dependent due to the training of dictionary and its optimization, and (iii) these algorithms are also unable to decompose several types of images [12]. Another challenge is the complicated oriented shape of source images that cannot be precisely categorized through an already trained dictionary [28]. To address these issues, the authors in [29] propose a training model that employs the well-known K-means algorithm [30]. Research in the domain of multimodal image fusion have produced promising outcomes; however, there are several drawbacks. The authors in [15,25] propose neural network-based fusion algorithms that efficiently adjust and fit the training parameters, but these algorithms are not capable of representing information from multiple sources [31].

Moreover, learning-based algorithms are also found to be useful in multimodal image fusion [32]. To this end, SR in combination with learning-based multimodal medical image fusion strategies are gaining interest of the research community [32]. The works in [23,27,33] employ the SR-based algorithms for image fusion. Similarly, the authors in [34] propose enhanced sparse representation orthogonal matching pursuit (SR-OMP) algorithms. Furthermore, the work in [24] presents another SR-based morphological component analysis model for pixel level image fusion. However, the blurring effect during decomposition restricts the performance of the proposed model. The authors in [35] present a multimodal image fusion algorithm that employs the SR-based cartoon-texture decomposition. However, it also faces the blurring issue during decomposition that results in considerable information loss. Additionally, the pyramid transformation [36] algorithm exhibits limited performance due to inaccuracy in information capturing and path details. Arif et al. [37] present a multimodal image fusion algorithm based on the curvelet transform and genetic algorithm (GA). Here, GA solves the suspicions and evaluates an optimized fused image. Kaur et al. [38] propose another image fusion algorithm that employs deep belief networks. Maqsood et al. [2] present a two scale image decomposition technique, where the spatial gradient-based edge detection method is used to acquire the detail layer and the SR rule is used to construct a fused image. This method produces improved image fusion results, however, it still experiences false structured artifacts in the fused image. Shahdoosti et al. [39] propose a sparse representation in the tetrolet domain for medical image fusion; however, this approach falls a victim of overlapping artifacts in the fused images.

From the literature survey, it is found that the SR-based image fusion algorithms have the advantage of better information extraction in comparison with other fusion algorithms. However, there are several issues that require urgent attention, such as (i) blurring effect near strong edges during image decomposition, (ii) appearance of false structured artifacts in fused image, and (iii) reduced contrast that results in high spatial distortion. To resolve these issues, we propose a novel image fusion algorithm that is detailed in the following section.

3. The Proposed Convolutional Sparse Image Decomposition (CSID) Algorithm

This section presents our proposed novel algorithm for multimodal image fusion, namely, Convolutional Sparse Image Decomposition (CSID). CSID comprises six phases that include contrast enhancement, edge detection, cartoon and texture decomposition, enhanced CSC-based sparse coding, sparse coefficient maps fusion, and fused image reconstruction, as depicted in Figure 1. These phases are detailed in the following subsections.

We take a source image (I_i) having $P \times Q$ dimensions, where $P = 1, 2, 3, \dots, p$, $Q = 1, 2, 3, \dots, q$, and $i \in [1, 2]$ that represents the CT and MRI images, respectively. CSID starts with contrast enhancement, which is detailed in the following subsection.

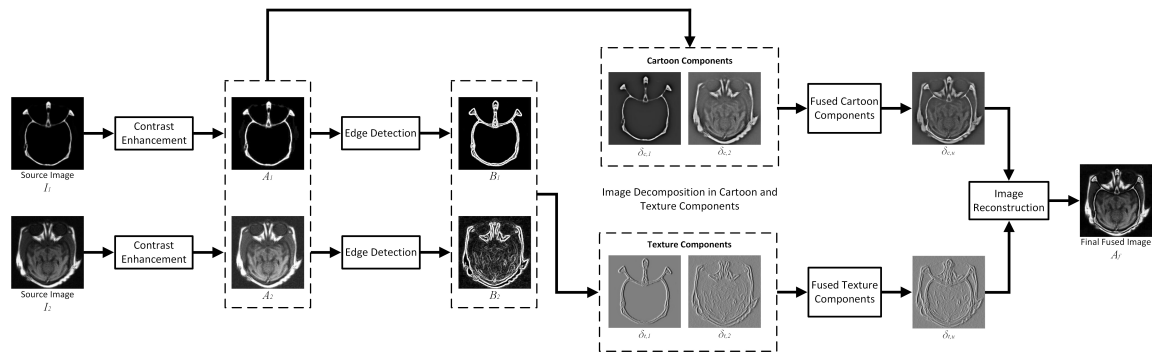


Figure 1. Procedural flowchart of the proposed CSID algorithm.

3.1. Contrast Enhancement

Contrast enhancement is a major preprocessing step in image fusion for diagnostics processes [40]. No-Reference Image Quality Assessment (NR-IQA) and Non-Parametric Histogram Equalization (NPHE) are commonly used for contrast enhancement [41,42]. NR-IQA employs histogram equalization and uses structural-similarity index metric to generate images with enhanced contrast, whereas NPHE employs modified spatial transformation-based adaptive contrast enhancement. However, these techniques require manual parameter tuning that limits their performance in accurately reflecting the image contrast with respect to an input image. To resolve the aforementioned issues, CSID employs the Bio Inspired Multi Exposure Fusion (BIMEF) framework [40] that improves contrast and preserves mean brightness of the source images. BIMEF uses illumination estimation for the construction of a weighted matrix. Thus, we start with the detection of optimal exposures using a camera response model, thereby, producing synthetic images that are better exposed in the regions in comparison with source images. Furthermore, we apply the weight matrix, with an appropriate exposure, upon the synthetic image, which is then merged with the source image for contrast enhancement. Here, to conserve the contrast of an image, a weighted matrix is associated with the scene brightness that is computed as follows [40],

$$\hat{I} = A^\phi, \tag{1}$$

where A denotes the scene brightness map and ϕ is a parameter managing the degree of enhancement. Moreover, since the highest regional maxima is better identified using a max function in comparison with a min function [43,44], CSID computes the dark regions (R), based upon initial estimation of brightness for each pixel (x), as [40],

$$R(p, q) = \max \hat{I}. \tag{2}$$

Since absolute brightness has local consistency for the boundaries with same structures, A eliminates the textural edges and builds a significant structure of an image. CSID optimizes A as [40],

$$X(p, q) = \min_A ||A - R||_2^2 + \varphi ||W \circ \nabla A||_1, \tag{3}$$

where $||\star||_1$ and $||\star||_2$ represent the l_1 and l_2 norms, respectively, ∇ is the first order derivative filter consisting $\nabla_h A$ and $\nabla_v A$ as horizontal and vertical components, respectively. φ denotes the coefficient and W refers to the weighted matrix. The weighted matrix is further refined to obtain significant edges in an image as [40],

$$Z_d(p, q) = \frac{1}{|\sum_{y \in d(p, q)} \nabla_f X(p, q)| + \kappa}, j \in (h, v), \tag{4}$$

where $|\star|$ denotes the positive value operator, $d(p, q)$ is the neighborhood window pointed at pixels p and q , and κ represents the constant to evade zero denominator. Moreover, we use (5) to evaluate A_i that minimizes complexity. A_i is then applied upon the source image to generate the final outcome of this phase, i.e., an image with enhanced contrast.

$$A_i = \min_A \sum_{p,q} \left((A(p,q) - d(p,q))^2 + \varphi \sum_{j \in \{h,v\}} \frac{Z_d(p,q) (\nabla_f(A(p,q)))^2}{|\nabla_f X(p,q)| + \epsilon} \right). \quad (5)$$

The results of contract enhancement are illustrated in Figure 2. After contrast enhancement, CSID enters the second phase, which is detailed in the following subsection.

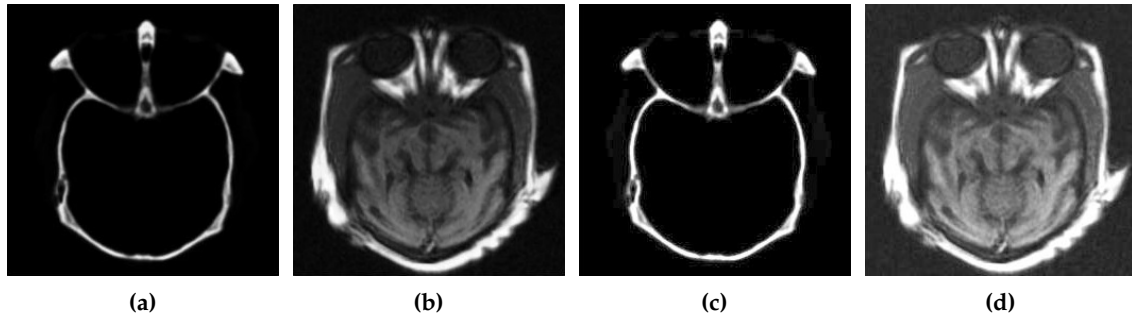


Figure 2. Contrast enhancement results, where (a,b) represent the source images and (c,d) show images with enhanced contrast using BIMEF.

3.2. Edge Detection

Edge detection finds the boundaries of objects in an image through identification of brightness discontinuities [45]. To this end, CSID performs edge detection in the image, obtained after contrast enhancement, by employing the Sobel operator [45]. Edge detection yields better performance when applied upon the images with enhanced contrast in comparison with original source images, as it is demonstrated in Figure 3.

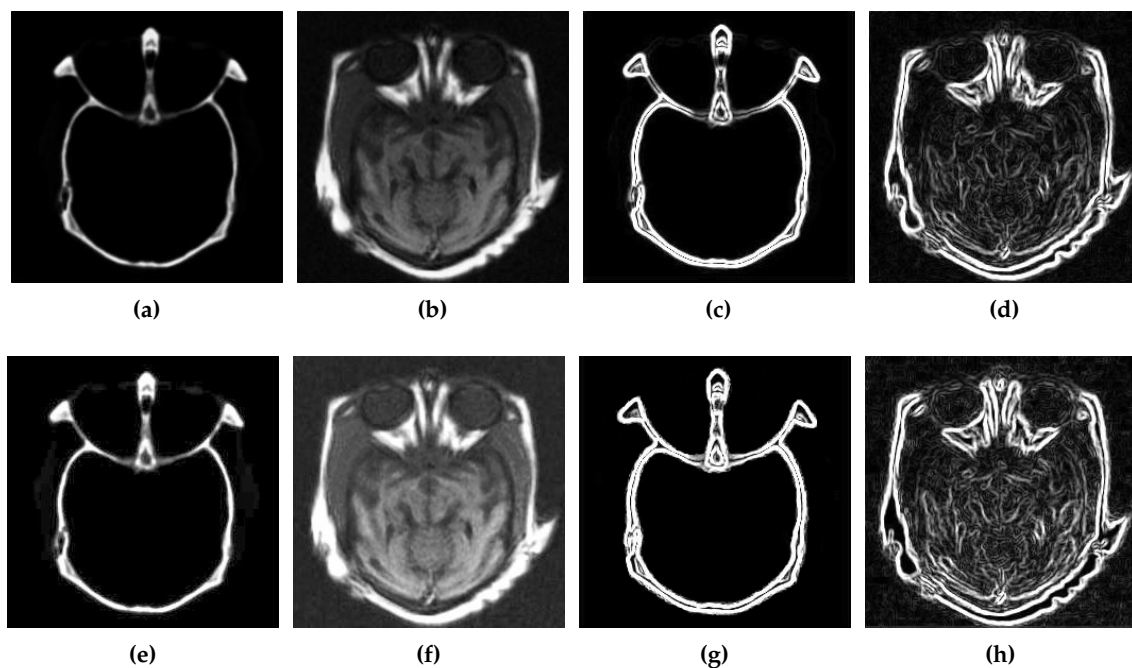


Figure 3. Edge detection results, where (a,b) are the original source images, (c,d) represent the gradients of (a,b) using the Sobel method, (e,f) shows the images with enhanced contrast using BIMEF, and (g,h) represent the gradients of (e,f) using the Sobel method.

For edge detection, CSID includes the image gradient approximation, where each location is either the corresponding gradient vector or the norm of this vector. The image is convoluted with the first kernel from left to right and the gradient for the X coordinate is obtained as,

$$B_p = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}. \tag{6}$$

Similarly, the gradient for the Y coordinate is obtained by convoluting the first kernel from top to bottom as,

$$B_q = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \tag{7}$$

Furthermore, the image gradient vectors, obtained using (6) and (7), are used to find edges as,

$$B_i = \sqrt{(B_p)^2 + (B_q)^2}. \tag{8}$$

Figure 3 depicts a comparison of edge information in the source and enhanced images obtained after completion of the first two phases of our proposed CSID algorithm. Figure 3a,b present the source CT and MRI images, respectively, while their edge maps are demonstrated in Figure 3c,d, respectively. Furthermore, Figure 3e,f include the improved CT and MRI images obtained after contrast enhancement (as detailed in Section 3.1) and their respective gradient maps are shown in Figure 3g,h. Here, improved edge detection is observed in the image with an enhanced contrast in comparison with edge detection in the original source image. On completion of this phase, CSID proceeds to the third phase, which is discussed in the following subsection.

3.3. Cartoon and Texture Decomposition

Cartoon-texture decomposition divides an image into the geometric cartoon and texture components, which removes the background interference. To this end, we propose a modification to the legacy Convolutional Sparse Coding (CSC) [46]. In our proposed modified CSC model, a similarity threshold is maintained to compute residual correlation similarity between the selected sparsest coefficients and the other coefficients in the sparse coding phase, which is discussed in Section 3.4. This expands the coefficients set, thereby, obtaining more suitable coefficients for SR. Furthermore, it also accelerates the process of sparse coding, and the residual correlation similarity easily minimizes the target error for each image patch signal. Moreover, multiple similar patches are used to represent a single image patch that further enhances the fused image by avoiding the blurring effect. An optimized solution of the CSC problem using (9) is generated as,

$$\begin{aligned} \delta_{c,t} = \delta_{c,u}, \delta_{t,u} = \min_{\varphi_{c,u}, \varphi_{t,u}} \frac{1}{2} & \left\| B_i - \sum_{u=1}^{U_c} h_{c,u} * \varphi_{c,u} - \sum_{u=1}^{U_t} h_{t,u} * \varphi_{t,u} \right\|_2^2 \\ & + \nu_c \sum_{u=1}^{U_c} \|\varphi_{c,u}\|_1 + \nu_t \sum_{u=1}^{U_t} \|\varphi_{t,u}\|_1, \end{aligned} \tag{9}$$

where $q_c = \{h_{c,u}\}_{u=1}^{U_c}$ and $q_t = \{h_{t,u}\}_{u=1}^{U_t}$ represent the sets of dictionary filters for SR of the cartoon and texture components. $\varphi_{c,u}$ and $\varphi_{t,u}$ are the sparse coefficients that estimate q_c and q_t when convolved with filters $\{h_{c,u}\}$ and $\{h_{t,u}\}$, respectively, and ν_c and ν_t are the positive regularization parameters. The optimization problem is solved iteratively over $\varphi_{c,u}$ and $\varphi_{t,u}$. As, $\varphi_{t,u}$, $h_{t,u}$, and $h_{c,u}$ are fixed, the accompanying issue is settled for the updated $\varphi_{c,u}$ as,

$$\delta_{c,u} = \min_{\varphi_{c,u}} \frac{1}{2} \left\| A_i - \sum_{u=1}^{U_t} h_{t,u} * \varphi_{t,u} - \sum_{u=1}^{U_c} h_{c,u} * \varphi_{c,u} \right\|_2^2 + \nu_c \sum_{u=1}^{U_c} \|\varphi_{c,u}\|_1. \tag{10}$$

Similarly, for the updated $\varphi_{t,u}$, keeping $\varphi_{c,u}$ fixed, the accompanying issue is settled as,

$$\delta_{t,u} = \min_{\varphi_{t,u}} \frac{1}{2} \left\| B_i - \sum_{u=1}^{U_c} h_{c,u} * \varphi_{c,u} - \sum_{u=1}^{U_t} h_{t,u} * \varphi_{t,u} \right\|_2^2 + \nu_t \sum_{u=1}^{U_t} \|\varphi_{c,t}\|_1. \tag{11}$$

Alternating Direction Method of Multipliers (ADMM)-based CSC [47] is used to address the aforementioned two issues in (10) and (11). This completes the cartoon and texture decomposition phase and allows CSID to proceed to the next phase, which is detailed in the following subsection.

3.4. Enhanced CSC-Based Sparse Coding

CSID employs our modified CSC model for cartoon and texture layer decomposition using $\{\varphi_{c,u}\}_{u=1}^{U_c}$ and $\{\varphi_{t,u}\}_{u=1}^{U_t}$, which represent the sparse coefficient vectors of the cartoon and texture components, respectively. Moreover, the same coefficient vectors are used to evaluate sparse coefficient maps in the next phase of our proposed CSID algorithm, as detailed in the following subsection.

3.5. Sparse Coefficient Maps Fusion

CSID applies the l_1 -norm of the sparse coefficient maps as the activity level measurement of the enhanced images, which remains a common approach adopted in several SR-based image fusion techniques [2,48,49]. Sparse coefficient maps fusion uses an attribute j ($j \in \{c, t\}$) that refers to the cartoon and the texture components and $\varphi_{j,1:U_j}^n(p, q)$ that uses the U_j dimensional vector consisting coefficients of $\varphi_{j,u}^n$ at points (p, q) . Hence, the initial activity level map $\zeta_j^n(p, q)$ map is obtained as,

$$\zeta_j^n(p, q) = \|\varphi_{j,1:U_j}^n(p, q)\|_1, j \in \{c, t\}. \tag{12}$$

A window-based averaging scheme is then applied for noise removal and enhancement of robustness to misregistration. Thus, the activity level map $\tilde{\zeta}_j^n(p, q)$ is computed as,

$$\tilde{\zeta}_j^n(p, q) = \frac{\sum_{x=-m_j}^{m_j} \sum_{y=-m_j}^{m_j} \zeta_j^n(p+x, q+y)}{(2m_j+1)^2}, j \in \{c, t\}, \tag{13}$$

where m_c and m_t refer to the window size for cartoon and texture components, respectively. Finally, CSID employs the “choose-max” rule to obtain the fused coefficient maps $\{\varphi_{j,u}^f\}_{u=1}^{U_j}$ with $j \in \{c, t\}$ as,

$$\varphi_{j,1:U_j}^f(p, q) = \varphi_{j,1:U_j}^{n^\circ}(p, q), n^\circ = \arg \max_n (\tilde{\zeta}_j^n(p, q)). \tag{14}$$

This completes the sparse coding phase that leads CSID to the final phase, which is detailed in the following subsection.

3.6. Fused Image Reconstruction

This phase is responsible to fuse enhanced CT and MRI images obtained from the aforementioned phases of CSID. The final fused image A_f is obtained through the linear combination of $\{\varphi_{c,u}^f\}_{u=1}^{U_c}$ and $\{\varphi_{t,u}^f\}_{u=1}^{U_t}$ as,

$$A_f = \sum_{u=1}^{U_c} h_{c,u} * \varphi_{c,u}^f + \sum_{u=1}^{U_t} h_{t,u} * \varphi_{t,u}^f. \tag{15}$$

This phase completes the multimodal fusion process through our proposed CSID algorithm. The cartoon component includes edges, round and anisotropic structure parts, whereas the texture component contains detailed texture information, periodic behaviors and several levels of noise data. This enables the proposed CSID algorithm to surpass the limitations of the existing fusion techniques

(as detailed in Section 2) by allowing the reconstruction of lost information in the CT and MRI images. The reconstruction process results in sharper images with enriched information. The fusion of such enhanced images improves accuracy during clinical diagnosis. The next section discusses the objective evaluation metrics used for the performance evaluation of our proposed CSID algorithm.

4. Objective Evaluation Metrics

The objective performance evaluation metrics include mutual information, entropy, feature mutual information, spatial structural similarity, and visual information fidelity, which are used by state-of-the-art works [27,50–53]. These metrics are defined in the following subsections.

4.1. Mutual Information (MI)

MI [50] computes the common information among two discrete variables as follows:

$$MI = \sum_{l=1}^n \sum_{m=1}^n H_{ij}(l, m) \log_2 \frac{H_{ij}(l, m)}{H_i(l)H_j(m)}, \quad (16)$$

where $H_{ij}(l, m)$ denotes the combined probability density distribution of the grayscale image in i and j . $H_i(l)$ and $H_j(m)$ refer to the probability density distribution of the grayscale image in i and j , respectively. MI expresses the sum of mutual information between each source image and the fused image. A larger value of MI refers to increased information extracted from the input source images.

4.2. Entropy (EN)

EN [27] refers to the measure of information randomness in a sample, which is expressed as,

$$EN(x) = - \sum_{l=0}^{N-1} H_i(l) \log_2 H_i(l), \quad (17)$$

where N is the number of gray levels, which is taken as 256 in this work, and $H_i(l)$ is the normalized histogram of the fused image i .

4.3. Feature Mutual Information (FMI)

FMI [51] computes the feature mutual information. A non-reference performance metric for fusion methods is determined as,

$$FMI_m^{i,j} = \frac{1}{N} \sum_{l=1}^N \frac{I_l(m, i)}{S_l(m)S_l(i)} + \frac{I_l(m, j)}{S_l(m) + S_l(j)}, \quad (18)$$

where N represents the number of sliding windows, $S_l(m)$ is the entropy of the n th window in an image m , $I_l(m, i)$ refers to the regional common information between the n th window of images m and i . Similarly, $I_l(m, j)$ is the regional MI between the n th window of images m and j . $FMI_m^{i,j}$ indicates the amount of edge information transmitted into the fused image from the source images. Here, $FMI_m^{i,j}$ remains proportional to the image quality, i.e., a greater value of $FMI_m^{i,j}$ yields better quality fused image.

4.4. Spatial Structural Similarity (SSS) $Q^{AB/F}$

$Q^{AB/F}$ SSS [52] is an edge-based fusion quality evaluation metric, which determines the quantity of transmitted edge information into the fused image from input images. $Q^{AB/F}$ for a set of source images is computed as,

$$Q^{AB/F} = \frac{\sum_{l=1}^m \sum_{j=1}^n (Q^{AB}(i, j)W^A(i, j) + Q^{BF}(i, j)W^B(i, j))}{\sum_{l=1}^m \sum_{j=1}^n (W^A(i, j) + W^B(i, j))}, \quad (19)$$

where $Q^{AB/F}(i, j)$ denotes the information transferred from a source image into the fused image for the pixel location (i, j) and $W^B(i, j)$ is the weight for a pixel location (i, j) . Here, pixels with a higher gradient value influence $Q^{AB/F}$ more in comparison with pixels having a lower gradient value. Thus, $W^A(i, j) = [Grad(x, y)]^T$: T remains constant.

4.5. Visual Information Fidelity (VIF)

VIF [53], being a perceptual distortion metric, stands as an important index for image quality assessment. In the context of image fusion, VIF evaluates the performance by calculating common data between a source image and its corresponding fused image. Since VIF provides accurate distortions identification, this work takes average VIF value for the performance evaluation of the given set of algorithms, as shall be discussed in Section 5.

5. Performance Evaluation

We evaluate and compare the proposed CSID algorithm to Discrete Wavelet Transform (DWT) [18], Dual Tree Complex Wavelet Transform (DTCWT) [54], Laplacian Pyramid (LP) [21], Guided Filtering based Fusion (GFF) [22], Non-Subsampled Contourlet Transform (NSCT) [19], Non-Subsampled Shearlet Transform domain-Parameter Adaptive Pulse Coupled Neural Network (NSST-PAPCNN) [55], Convolutional Sparse Representation (CSR) [23], Convolutional Sparsity based Morphological Component Analysis (CSMCA) [24], and Convolutional Neural Network (CNN) [25]. The following subsection details the simulation parameters used in this paper.

5.1. Simulation Setup

Simulation results are derived using MATLAB R2020b (MathWorks Inc., MA, USA), which is used by state-of-the-art methods for multimodal image fusion due to its extensive built-in libraries support [2,23–25]. The hardware platform includes Intel Core *i7 – 9750H* 2.59 Giga Hertz processor with 16 GB memory running Microsoft Windows 10 (Microsoft, WA, USA). The multimodal brain image datasets (Data-1 through Data-6) are obtained from [56], which are composed of the CT and MR images. For performance evaluation, selected 500 grayscale images are taken from each of the aforementioned datasets. Input images dimensions are standardized as 256×256 pixels. Both qualitative and quantitative analysis are performed for the performance evaluation that are detailed in the following subsections.

5.2. Results and Discussion

Six different datasets of multimodal images, referred as Data-1 through Data-6, are used in the simulations. Figure 4 depicts sample images from the aforementioned datasets. The fusion results, generated by our proposed CSID algorithm and the aforementioned eminent fusion algorithms, are shown in Figures 5–10. Each result presented is averaged over 20 replicated simulation runs by keeping all the parameters fixed and changing the random seed values. The following subsections demonstrate and discuss the obtained results.

5.2.1. Qualitative Analysis of the Given Set of Algorithms for Multimodal Fusion

This subsection presents the results based on visual observations of the images generated through our proposed CSID algorithm in comparison with the aforementioned algorithms using different datasets, i.e., Data-1 through Data-6. Visual quality comparison of the Data-1 dataset using different fusion methods, i.e., DWT, DTCWT, LP, GFF, NSCT, NSST-PAPCNN, CSR, CSMCA, CNN, and the proposed algorithm are shown in Figure 5a through Figure 5j, respectively. A CT image gives information about hard tissues and their structures, whereas an MRI image indicates information regarding soft tissues. For better diagnosis, it is essential to merge critical information of the aforementioned images into one fused image [12]. In this regard, the aforementioned set of algorithms perform multimodal image fusion.

The qualitative results shown in Figure 5 depict inferior performance, in terms of contrast and visual effect, for DWT (Figure 5a), DTCWT (Figure 5b), NSCT (Figure 5e), and CSR (Figure 5g). Note that these algorithms are not capable of preserving information in the fused image, which relates to the objective evaluation metric MI that remains proportional to the level of information extraction. Additionally, Section 5.2.2 further validates this claim through quantitative analysis, where DWT, DTCWT, NSCT, and CSR exhibit lower MI score in comparison with other algorithms. Moreover, GFF (Figure 5d) and NSST-PAPCNN (Figure 5f), yield better results, when compared with DWT, DTCWT, NSCT, and CSR algorithms, by avoiding information loss. However, the lack of noise removal results in over enhancement of the structural features in these algorithms. CSMCA (Figure 5h) and CNN (Figure 5i) further improve the visual quality, where enhanced visualization remains an outcome of lesser information loss. Finally, our proposed CSID algorithm (Figure 5j) yields clear, high contrast and superior visual quality and preserves the salient features, which include considerably enhanced bone structure and soft tissues information in comparison with other given algorithms.

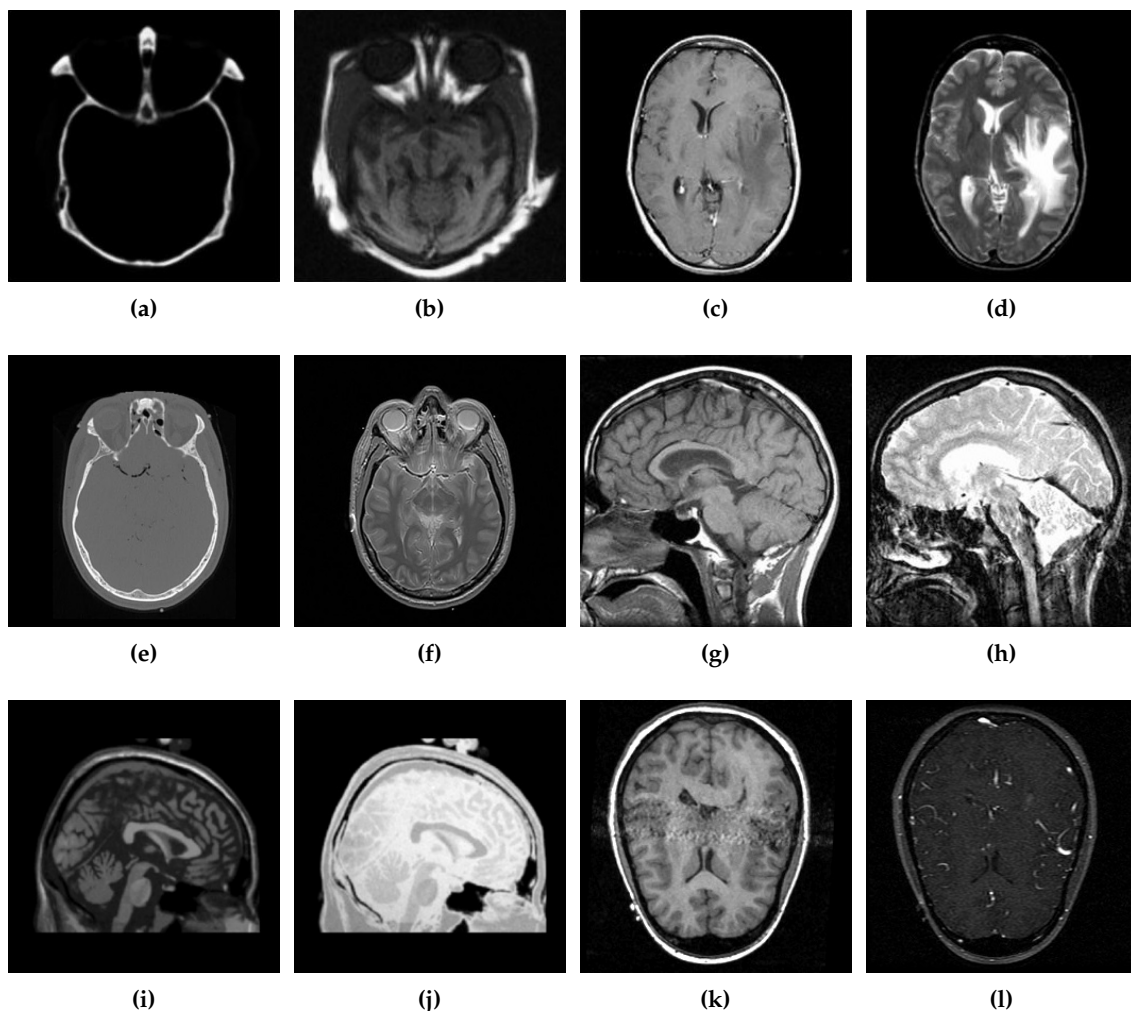


Figure 4. Sample source images from the given datasets (Data-1 through Data-6), where (a,b) \in Data-1, (c,d) \in Data-2, (e,f) \in Data-3, (g,h) \in Data-4, (i,j) \in Data-5, and (k,l) \in Data-6.

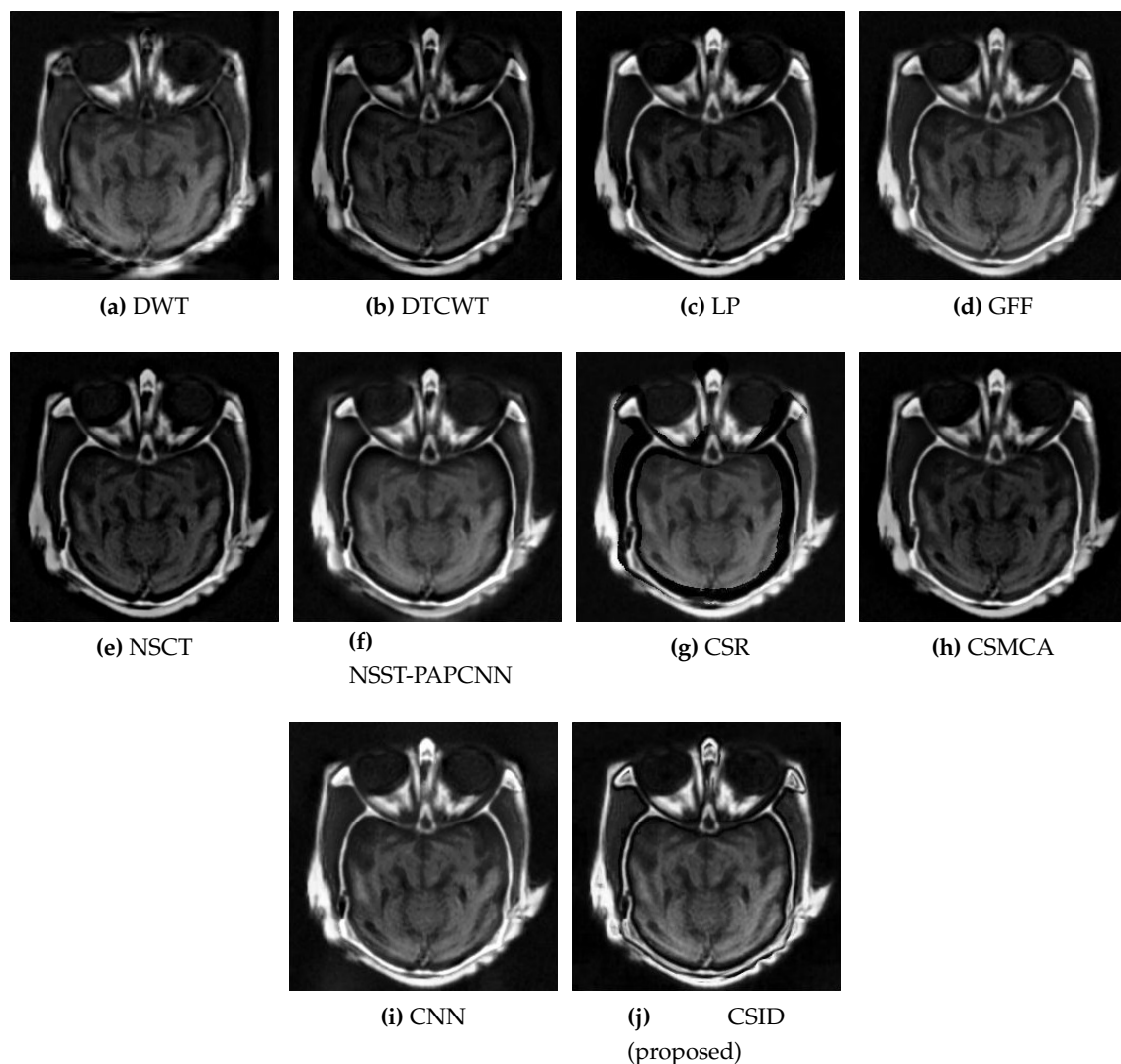


Figure 5. Comparative analysis, based upon visual observation, of the resultant fused images generated by the given set of algorithms using Data-1 dataset.

Similarly, Figure 6 shows the qualitative visual analysis for the Data-2 dataset. As a result of the smaller MI score, DWT (Figure 6a), NSCT (Figure 6e), and CSR (Figure 6g) show the highest information loss, among the selected algorithms, for the overlapping areas in MRI and CT images that result in visual deformations of the fused images. Similarly, CNN (Figure 6i) also does not remain effective in transferring information from the source images. The following section (Section 5.2.2) provides quantitative analysis with respect to the given objective evaluation metrics (as discussed in Section 4) that affirms the aforementioned statements. Moreover, in addition to MI, $FMI_m^{x,y}$ and $Q^{AB/F}$ evaluation metrics also remain critical that relates to accuracy in the resultant fused images. Although, GFF, NSST-PAPCNN, and CSMCA provide better results in comparison with DWT, NSCT and CSR by conveying complementary information into the fused image, but these algorithms lack accuracy (as shall be discussed in Section 5.2.2 through quantitative analysis). In the end, note that our proposed algorithm (Figure 6j) provides better visual effects in comparison with the other aforementioned algorithms, due to its improved information extraction and edge detection abilities.

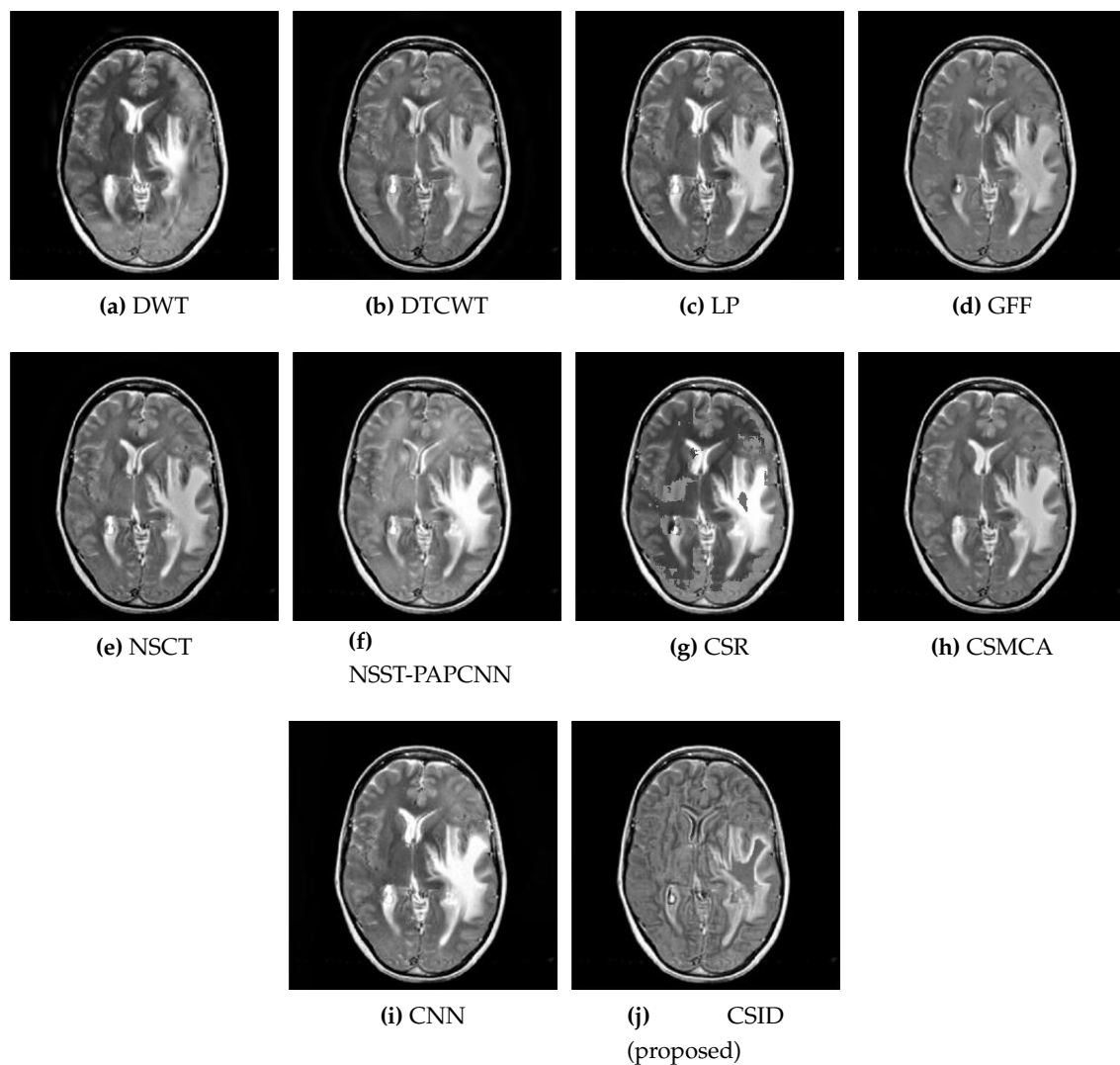


Figure 6. Comparative analysis, based upon visual observation, of the resultant fused images generated by the given set of algorithms using Data-2 dataset.

Moreover, the qualitative results for the Data-3 dataset are demonstrated in Figure 7. Visual results of DWT (Figure 7a), DTCWT (Figure 7b), LP (Figure 7c), CSR (Figure 7g), and CSMCA (Figure 7h), face issues in edge regions, as these algorithms do not effectively preserve information from the source images. The visual quality of GFF (Figure 7d), NSCT (Figure 7e), NSST-PAPCNN (Figure 7f), and CNN (Figure 7i) remain slightly better, however, these algorithms still remain ineffective to reduce the information loss considerably. Furthermore, the results of our proposed CSID algorithm (Figure 7j) are found the best among all the aforementioned algorithms, as it efficiently preserves edges and texture details.

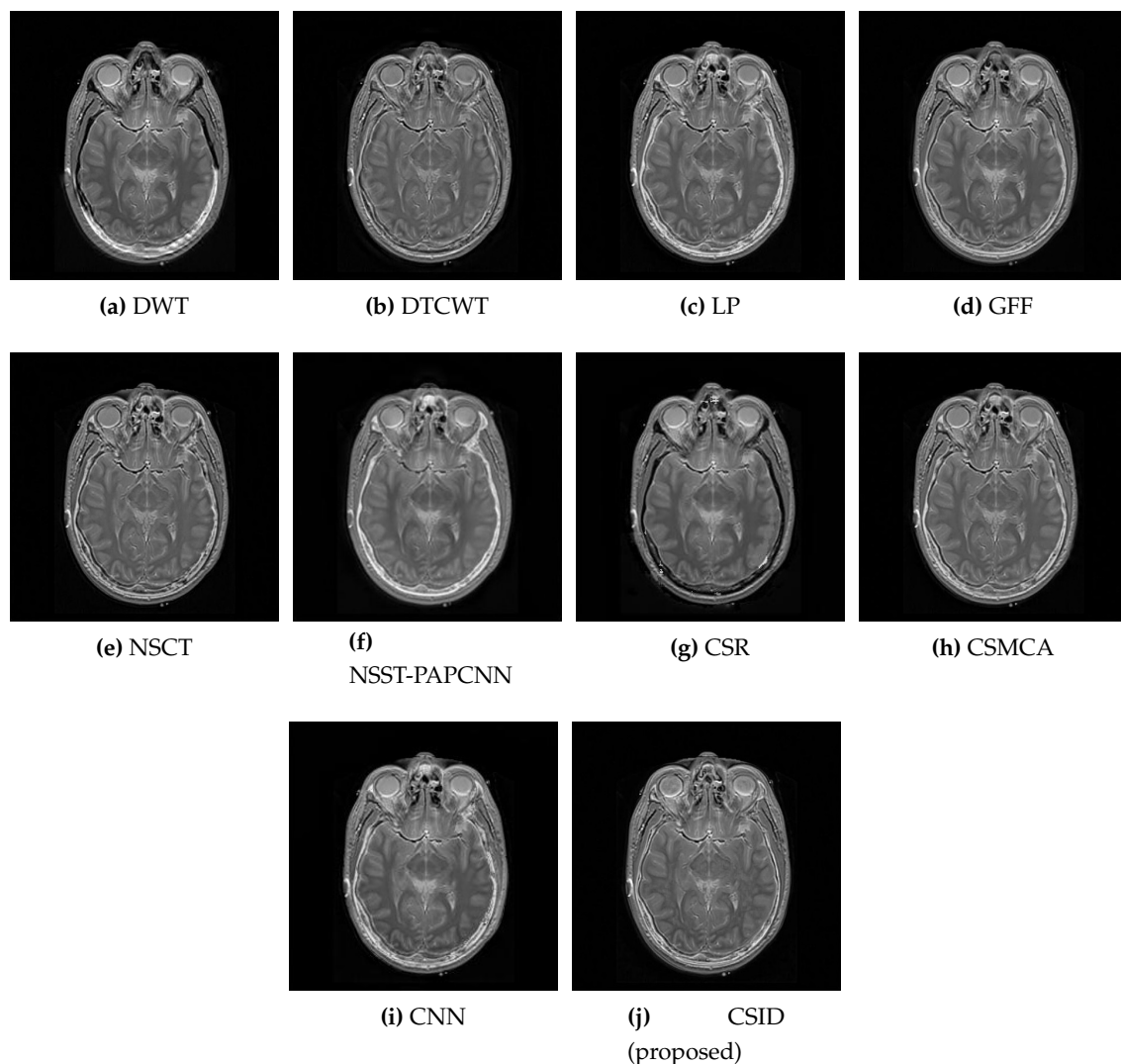


Figure 7. Comparative analysis, based upon visual observation, of the resultant fused images generated by the given set of algorithms using Data-3 dataset.

Image fusion performance is further evaluated using the Data-4 dataset for all the aforementioned algorithms and the corresponding qualitative results are shown in Figure 8. Some image fusion algorithms, such as DWT (Figure 8a), DTCWT (Figure 8b), GFF (Figure 8d), NSST-PAPCNN (Figure 8f), and CSMCA (Figure 8h), show smaller $Q^{AB/F}$ scores that impact sharpness of the resultant fused images. Moreover, these algorithms also experience distorted regions due to lower VIF scores in comparison with other algorithms, as shall be discussed in the quantitative analysis performed in Section 5.2.2. Additionally, LP (Figure 8c) and CSR (Figure 8g) are also found incapable of retaining originality due to increased information loss. NSCT (Figure 8e) and CNN (Figure 8i) provide comparatively improved results, as these algorithms provide effective information integration. Here again, CSID accomplishes the best performance in comparison with the aforementioned algorithms, as it remains capable of transferring more details and provides better contrast.

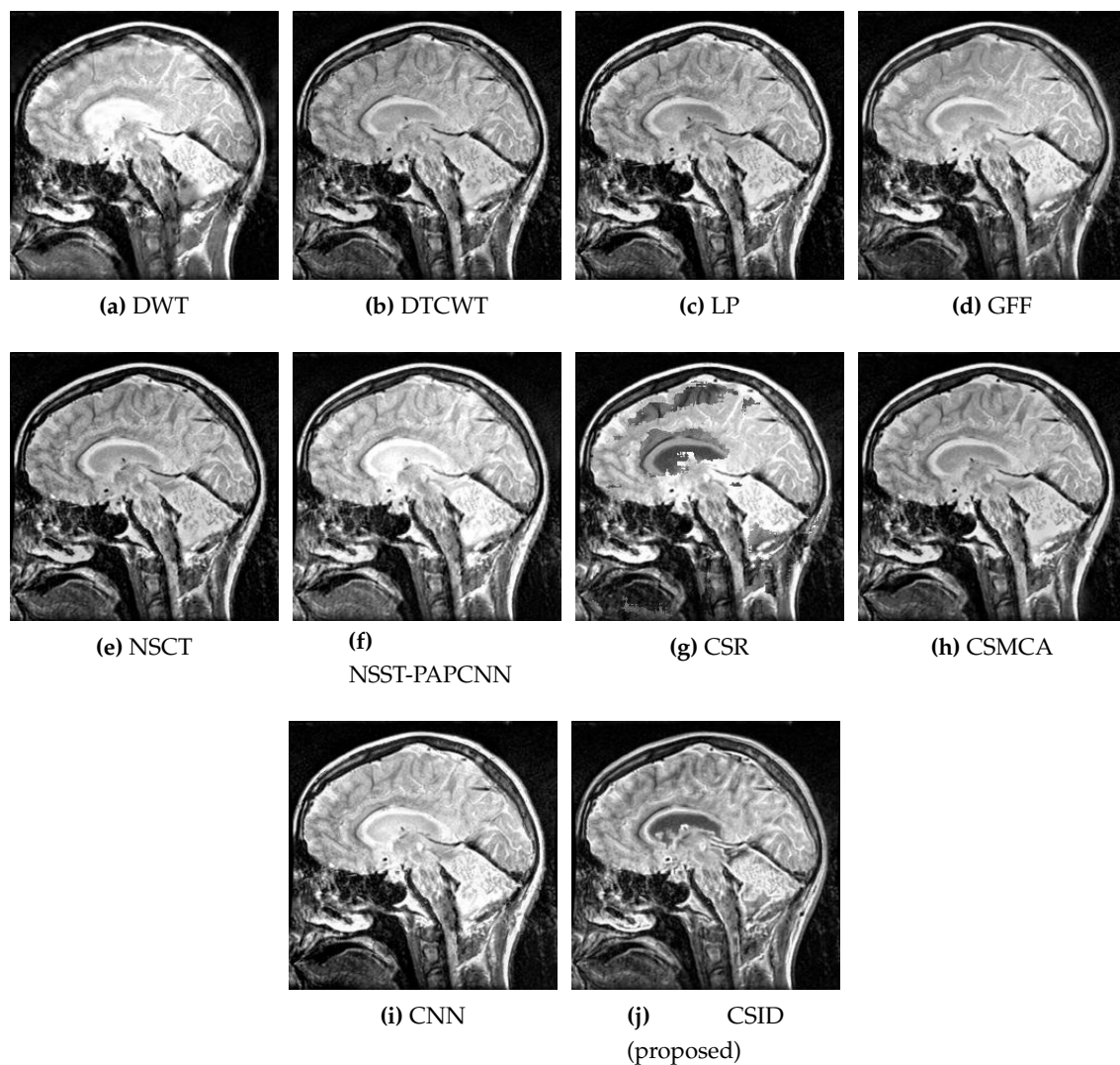


Figure 8. Comparative analysis, based upon visual observation, of the resultant fused images generated by the given set of algorithms using Data-4 dataset.

Furthermore, the aforementioned set of algorithms is evaluated using the Data-5 dataset and the corresponding qualitative results are shown in Figure 9.

Here, DWT (Figure 9a), DTCWT (Figure 9b), NSCT (Figure 9e), NSST-PAPCNN (Figure 9f), and CSR (Figure 9g) provide limited structural information. Moreover, LP (Figure 9c), GFF (Figure 9d), CSMCA (Figure 9h), and CNN (Figure 9i) experience faded edges. However, our proposed CSID algorithm (Figure 9j) shows the superior performance, in comparison with the other aforementioned algorithms, due to its enhanced energy information preservation.

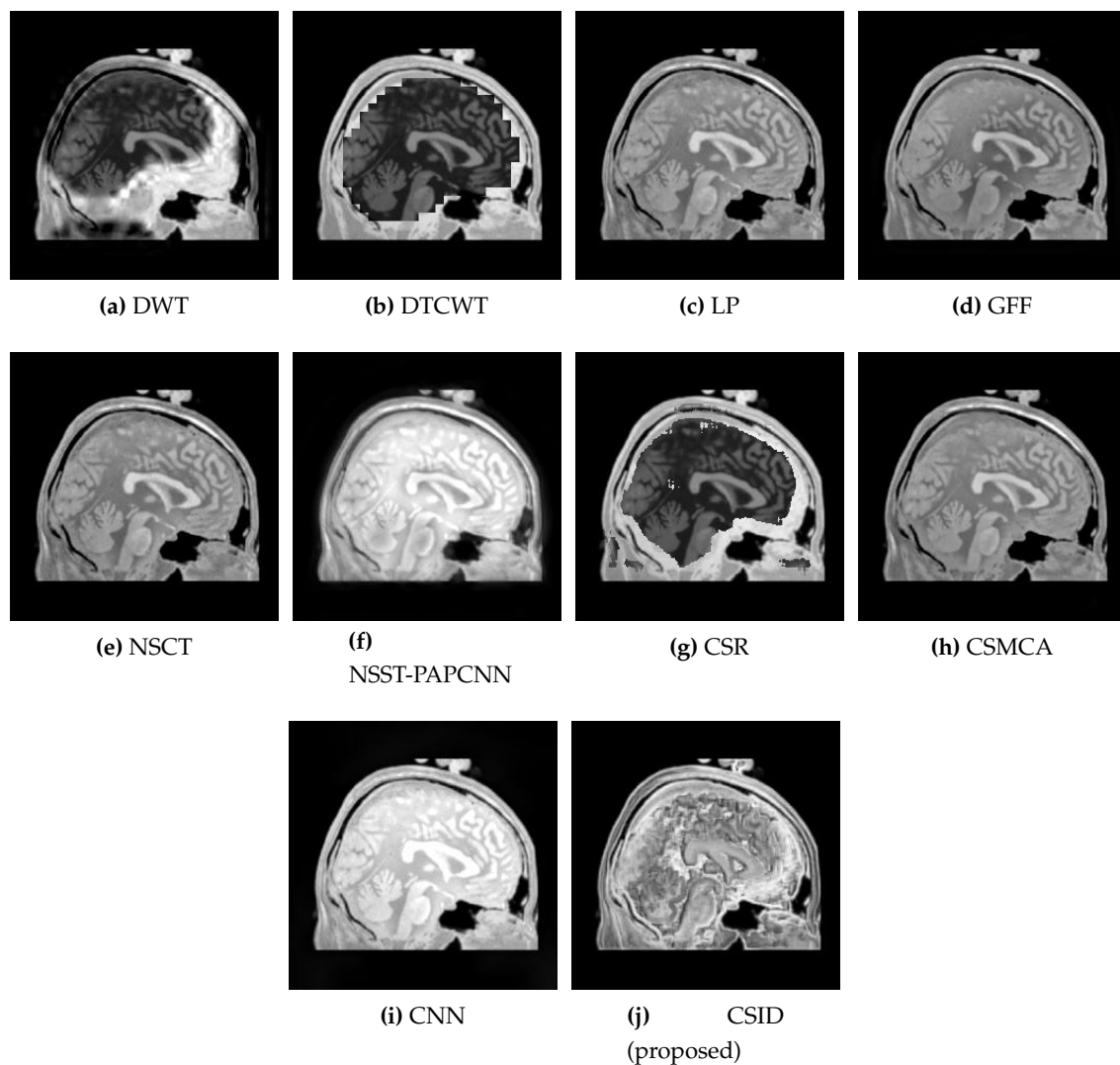


Figure 9. Comparative analysis, based upon visual observation, of the resultant fused images generated by the given set of algorithms using Data-5 dataset.

Finally, CSID and the aforementioned set of algorithms are evaluated using the Data-6 dataset, where Figure 10 demonstrates the corresponding qualitative results. All the algorithms, other than CSID, are unable of extracting detailed information that results in blurred fused images. To this end, our proposed CSID algorithm shows improved edge detection and provides enhanced contrast, in comparison with all the aforementioned algorithms that yield better visualization.

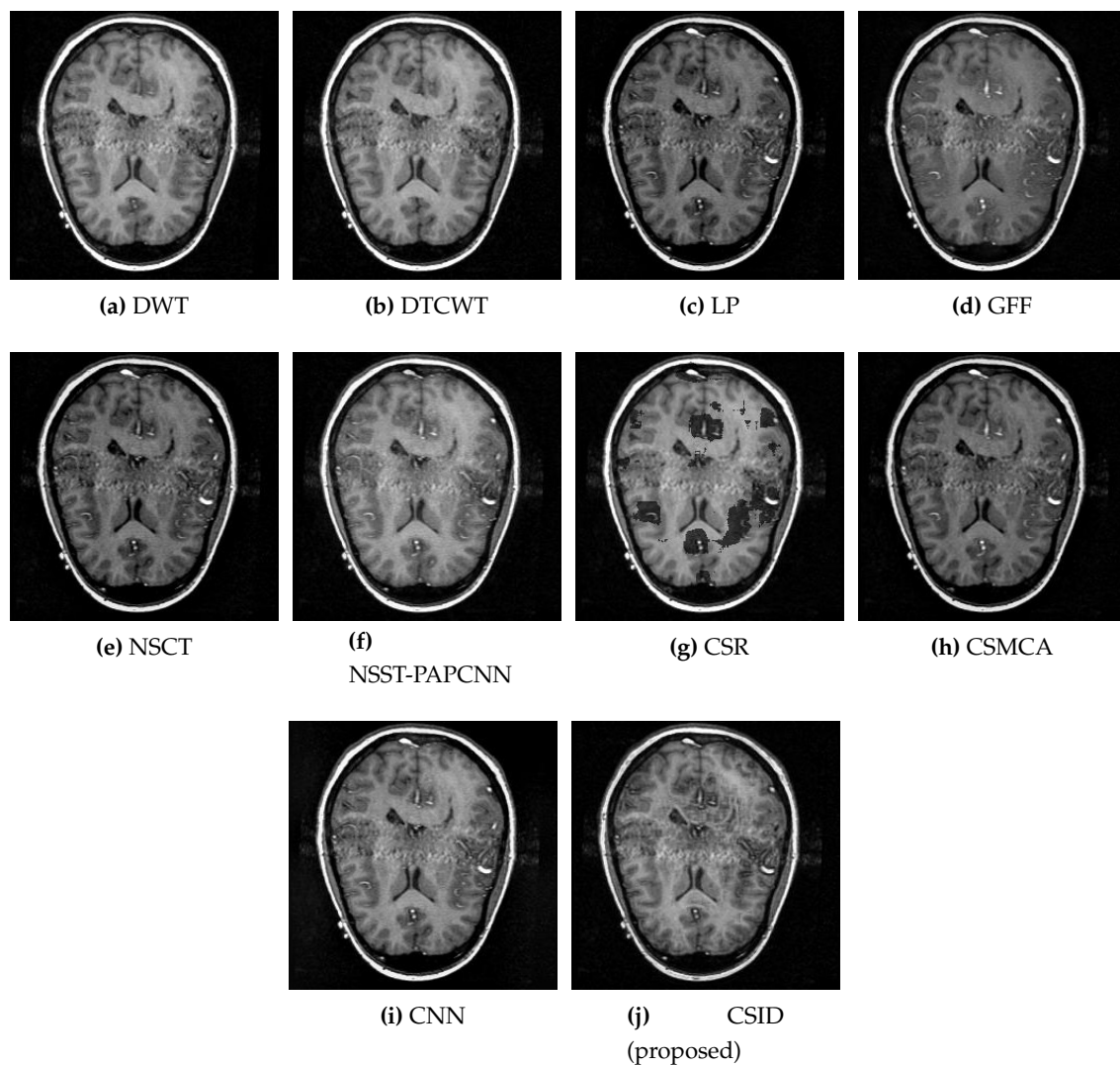


Figure 10. Comparative analysis, based upon visual observation, of the resultant fused images generated by the given set of algorithms using Data-6 dataset.

5.2.2. Quantitative Analysis of the Given Set of Algorithms for Multimodal Fusion

Tables 1 and 2 show the results obtained from DWT, DTCWT, LP, GFF, NSCT, NSST-PAPCNN, CSR, CSMCA, CNN, and CSID against the objective metrics, such as MI, EN, FMI, $Q^{AB/F}$, and VIF (as detailed in Section 4). Scores obtained for these metrics remain proportional to the quality of the resultant fused image. Therefore, a smaller score indicates missing information and false structured artifacts during fusion process, whereas a higher score results in enhanced fused images. To this end, all the highest scores obtained are highlighted in bold in Tables 1 and 2. These results demonstrate that our proposed CSID achieves higher MI, EN, FMI, $Q^{AB/F}$, and VIF scores in comparison with all the other image fusion algorithms using different datasets, i.e., Data-1 through Data-6. This indicates improved performance for CSID, as it remain capable of extracting enriched information from the input source images, thereby, preserving enhanced edges details and yields enhanced visual quality.

Table 1. The quantitative comparison of fusion methods. Best values are shown in bold.

Images	Fusion Methods	MI [50]	EN [27]	$FMI_m^{x,y}$ [51]	$Q^{AB/F}$ [52]	VIF [53]	Time (s)
Data-1	DWT [18]	2.1141	6.1512	0.7654	0.6656	0.4065	3.644
	DTCWT [54]	2.1044	6.2074	0.8341	0.6454	0.3976	6.645
	LP [21]	2.5508	6.2724	0.7412	0.6321	0.4141	1.699
	GFF [22]	3.4313	6.7971	0.9032	0.7849	0.4864	3.004
	NSCT [19]	2.2087	6.1488	0.7612	0.6872	0.3864	2.100
	NSST-PAPCNN [55]	2.4665	6.9551	0.4559	0.6968	0.9015	5.083
	CSR [23]	2.087	6.4871	0.3712	0.6327	0.8041	24.037
	CSMCA [24]	2.5863	6.3274	0.4751	0.7373	0.9088	76.700
	CNN [25]	3.5248	6.7541	0.7712	0.7992	0.8991	10.696
Proposed CSID	3.9649	6.9971	0.9781	0.8021	0.9897	4.065	
Data-2	DWT [18]	3.5472	5.5481	0.8493	0.6922	0.5593	3.649
	DTCWT [54]	3.5201	6.2074	0.8341	0.6756	0.5521	6.544
	LP [21]	3.5908	5.6692	0.8568	0.6571	0.4352	1.783
	GFF [22]	3.8595	5.8459	0.8596	0.5919	0.4295	3.024
	NSCT [19]	3.5110	5.5703	0.8498	0.6837	0.5435	2.112
	NSST-PAPCNN [55]	3.5462	7.7278	0.5597	0.5136	0.8393	5.144
	CSR [23]	3.8744	6.0867	0.5614	0.6667	0.4715	23.441
	CSMCA [24]	3.5008	7.6182	0.5728	0.5772	0.8615	74.994
	CNN [25]	4.2014	7.8421	0.7458	0.6969	0.8015	10.447
Proposed CSID	4.8821	8.0142	0.8850	0.7199	0.8715	4.051	
Data-3	DWT [18]	3.0523	7.1581	0.9438	0.7542	0.5369	3.702
	DTCWT [54]	3.0871	7.1287	0.9361	0.7414	0.5348	6.414
	LP [21]	3.1847	7.0536	0.8914	0.7499	0.4832	1.955
	GFF [22]	4.0609	5.2463	0.9013	0.6788	0.4486	3.287
	NSCT [19]	3.7394	7.1873	0.9197	0.7101	0.5132	2.089
	NSST-PAPCNN [55]	3.7147	5.3329	0.5536	0.5956	0.8825	5.090
	CSR [23]	3.9478	5.0398	0.8657	0.6342	0.7226	23.339
	CSMCA [24]	3.3098	5.0064	0.4679	0.6397	0.9048	76.018
	CNN [25]	4.0183	6.9420	0.9224	0.7301	0.9755	9.581
Proposed CSID	4.4388	7.5970	0.9744	0.7842	0.9891	4.049	

Table 2. The quantitative comparison of fusion methods. Best values are shown in bold.

Images	Fusion Methods	MI [50]	EN [27]	$FMI_m^{x,y}$ [51]	$Q^{AB/F}$ [52]	VIF [53]	Time (s)
Data-4	DWT [18]	3.5962	4.7393	0.3823	0.5835	0.9027	3.645
	DTCWT [54]	3.6632	4.8551	0.8339	0.6921	0.6679	6.643
	LP [21]	3.4733	4.6547	0.7690	0.6391	0.9255	1.774
	GFF [22]	3.4514	4.4081	0.9047	0.6470	0.4961	3.132
	NSCT [19]	3.8544	4.5360	0.8395	0.7093	0.7769	2.143
	NSST-PAPCNN [55]	3.3372	5.0598	0.5401	0.6076	0.8960	5.232
	CSR [23]	3.6584	4.7695	0.8471	0.6655	0.8467	22.998
	CSMCA [24]	3.4007	4.3896	0.4939	0.6601	0.9027	75.802
	CNN [25]	4.2540	5.1748	0.8421	0.7441	0.9408	10.113
Proposed CSID	4.6987	5.9459	0.9814	0.8023	0.9947	4.122	
Data-5	DWT [18]	4.0214	4.6386	0.4777	0.5782	0.7592	3.650
	DTCWT [54]	4.2985	4.7687	0.4885	0.6257	0.5573	6.625
	LP [21]	4.4128	4.8825	0.5241	0.6825	0.5826	1.874
	GFF [22]	4.7093	5.2982	0.7849	0.7259	0.7928	3.332
	NSCT [19]	3.9309	4.9304	0.6908	0.6827	0.7469	2.139
	NSST-PAPCNN [55]	4.1937	4.9809	0.7360	0.6887	0.6993	5.403
	CSR [23]	4.5094	5.0297	0.6997	0.6259	0.5067	23.422
	CSMCA [24]	5.0924	5.9330	0.7485	0.7759	0.8257	76.112
	CNN [25]	5.1118	5.9989	0.8697	0.8267	0.8881	10.691
Proposed CSID	5.2471	6.2874	0.8847	0.8728	0.8971	4.041	
Data-6	DWT [18]	3.6877	4.8474	0.5570	0.4938	0.5551	3.647
	DTCWT [54]	3.6439	4.8839	0.5683	0.5097	0.6086	6.245
	LP [21]	3.9482	4.9029	0.6019	0.6287	0.6239	1.963
	GFF [22]	4.1675	5.0098	0.7829	0.6876	0.7452	3.504
	NSCT [19]	3.8888	4.8729	0.7067	0.6431	0.7884	2.146
	NSST-PAPCNN [55]	4.0671	4.9038	0.7149	0.6835	0.7763	5.113
	CSR [23]	3.7432	4.4597	0.6839	0.5334	0.6720	23.483
	CSMCA [24]	4.5810	4.9997	0.8097	0.7482	0.8027	76.772
	CNN [25]	4.6744	5.2779	0.8527	0.7983	0.8341	10.834
Proposed CSID	4.8887	5.8209	0.8817	0.8497	0.8748	4.047	

In the past few decades, non-invasive applications (like multimodal fusion) have gained tremendous popularity among the healthcare professionals that adds ease and accuracy to the diagnostic process [57,58]. CSID aims to enhance clinical diagnostics by improving the multimodal fusion. We acquired the expert opinion of two healthcare professionals (one radiologist and one physician, whose help we kindly acknowledge) based upon the visual observation of the resultant fused images generated through the given set of algorithms. These experts appreciated the enhanced results generated by CSID in comparison with other state-of-the-art algorithms. Furthermore, it was added by the experts that CSID enables detailed information extraction along with clearer edge detection to yield enhanced fused images, which remain promising for better clinical diagnosis.

5.2.3. Statistical Analysis of the Results

We used the non-parametric Friedman's test and the post-hoc Nemenyi test to analyze how the analyzed methods differ from each other. The Nemenyi test calculates a Critical Difference (CD) using the Tukey's distribution, and any difference in the ranks between method ranks that is greater than the CD is considered as significantly different [59]. In Figure 11, we used the values from Tables 1 and 2 to calculate average data fusion method ranks. The results of the Nemenyi test show that the

proposed CSID method achieved better values than all other methods when evaluated in terms of MI , EN , $FMI_m^{x,y}$, $Q^{AB/F}$, and VIF scores for six datasets (from Data-1 to Data-6), however the advantage over the next best method CNN [25] was not statistically significant (difference between mean ranks < 2.4734 , Friedman's $p < 0.001$).

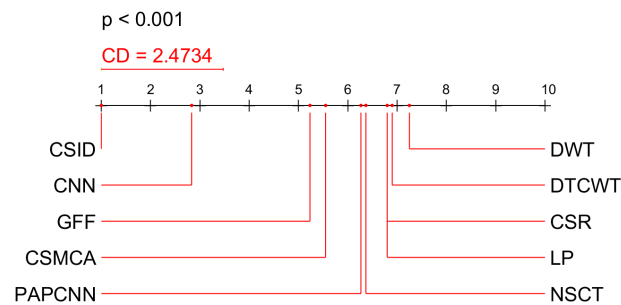


Figure 11. Results of Nemenyi test on different scores evaluating data fusion methods while using six datasets (Data-1 through Data-6). CSID is the proposed method.

5.2.4. Computational Efficiency

This subsection evaluates the computational efficiency of our proposed CSID algorithm in comparison with DWT, DTCWT, LP, GFF, NSCT, NSST-PAPCNN, CSR, CSMCA, and CNN. Tables 1 and 2 show the execution time (in seconds) for each of the aforementioned algorithms when applied on the given datasets, i.e., Data-1 through Data-6. The results shown in Tables 1 and 2 demonstrate that LP exhibits the smallest execution time among the aforementioned algorithms, whereas CSMCA bears the highest execution time. Considering our proposed CSID algorithm, it has smaller execution time in comparison with DTCWT, NSST-PAPCNN, CSR, and CSMCA, and higher execution time than DWT, LP, GFF and NSCT. This is because CSID employs the cartoon-texture component gradient-based feature extraction that enhances image visualization, as shown in Section 5.2.1. Since the main aim of this work is to enhance visualization, a tradeoff in terms of slight increase in execution time remains affordable. Moreover, execution time minimization will be taken as future extension of this work.

6. Conclusions

Multimodal medical image fusion has gained a firm stature in modern day healthcare systems. There are several fusion algorithms, which merge multiple input source images to extract detailed information that is exploited to enhance the clinical diagnosis. However, these algorithms have several limitations, such as blurring edges during decomposition, excessive information loss that results in false structured artifacts, and high spatial distortion due to inadequate contrast. This work aims to resolve the aforementioned issues and proposes a novel CSID algorithm that performs contrast stretching and identifies edges by using spatial gradient. CSID proposes the use of cartoon-texture decomposition that creates an overcomplete dictionary. Moreover, this work proposes a modification to the legacy convolutional sparse coding method, and employs enhanced decision maps and fusion rule to obtain the final fused image. Simulation results demonstrate that CSID attains improved performance, in terms of visual quality and enriched information extraction, as compared with other known fusion algorithms. Future work will aim on reducing the execution time of CSID to enable rapid image fusion. Furthermore, the extension of CSID to provide applications, such as visible-infrared image, multi-exposure image, and multi-focus image fusions, can also be taken as a future research direction.

Author Contributions: Conceptualization: S.H., S.R.M., and S.M.; methodology: S.H., S.R.M., and S.M.; software: S.R.M., and S.M.; validation: S.H., S.R.M., and S.M.; formal analysis: S.H., S.R.M., and R.D.; investigation: S.H., S.M., and R.D.; data curation: S.H., and R.D.; writing—original draft preparation: S.H., S.R.M., and S.M.; supervision: S.H. and R.D.; project administration: S.H., and R.D.; funding acquisition: S.M., and R.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We are thankful to Ayesha Amin (MBBS, FCPS (Radiology), FRCR) working as Radiologist at Maroof International Hospital, Islamabad, Pakistan, and Huma Riaz (MBBS, FCPS (Haematology)) working as Assistant Professor at Hayatabad Medical Complex, Peshawar, Pakistan, for spending their precious time in evaluating the visual results presented in this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Liu, X.; Liu, Q.; Wang, Y. Remote sensing image fusion based on two-stream fusion network. *Inf. Fusion* **2020**, *55*, 1–15. [[CrossRef](#)]
2. Maqsood, S.; Javed, U. Biomedical Signal Processing and Control Multi-modal Medical Image Fusion based on Two-scale Image Decomposition and Sparse Representation. *Biomed. Signal Process. Control* **2020**, *57*, 101810. [[CrossRef](#)]
3. Chen, G.; Li, C.; Wei, W.; Jing, W.; Woźniak, M.; Blažauskas, T.; Damaševičius, R. Fully convolutional neural network with augmented atrous spatial pyramid pool and fully connected fusion path for high resolution remote sensing image segmentation. *Appl. Sci.* **2019**, *9*, 1816. [[CrossRef](#)]
4. Nisa, M.; Shah, J.H.; Kanwal, S.; Raza, M.; Khan, M.A.; Damaševičius, R.; Blažauskas, T. Hybrid malware classification method using segmentation-based fractal texture analysis and deep convolution neural network features. *Appl. Sci.* **2020**, *10*, 4966. [[CrossRef](#)]
5. Bernardo, L.S.; Quezada, A.; Muñoz, R.; Maia, F.M.; Pereira, C.R.; Wu, W.; de Albuquerque, V.H.C. Handwritten pattern recognition for early Parkinson's disease diagnosis. *Pattern Recognit. Lett.* **2019**, *125*, 78–84. [[CrossRef](#)]
6. Gambhir, D.; Manchanda, M. Waveatom transform-based multimodal medical image fusion. *Signal Image Video Process.* **2019**, *13*, 321–329. [[CrossRef](#)]
7. Manchanda, M.; Sharma, R. An improved multimodal medical image fusion algorithm based on fuzzy transform. *J. Vis. Commun. Image Represent.* **2018**, *51*, 76–94. [[CrossRef](#)]
8. Ke, Q.; Zhang, J.; Wei, W.; Damaševičius, R.; Woźniak, M. Adaptive independent subspace analysis of brain magnetic resonance imaging data. *IEEE Access* **2019**, *7*, 12252–12261. [[CrossRef](#)]
9. Wei, W.; Zhou, B.; Polap, D.; Wozniak, M. A regional adaptive variational PDE model for computed tomography image reconstruction. *Pattern Recognit.* **2019**, *92*, 64–81. [[CrossRef](#)]
10. Guo, Z.; Li, X.; Huang, H.; Guo, N.; Li, Q. Deep Learning-Based Image Segmentation on Multimodal Medical Imaging. *IEEE Trans. Radiat. Plasma Med. Sci.* **2019**, *3*, 162–169. [[CrossRef](#)]
11. Khan, M.A.; Ashraf, I.; Alhaisoni, M.; Damaševičius, R.; Scherer, R.; Rehman, A.; Bukhari, S.A.C. Multimodal Brain Tumor Classification Using Deep Learning and Robust Feature Selection: A Machine Learning Application for Radiologists. *Diagnostics* **2020**, *10*, 565. [[CrossRef](#)] [[PubMed](#)]
12. Maqsood, S.; Javed, U.; Riaz, M.M.; Muzammil, M.; Muhammad, F.; Kim, S. Multiscale Image Matting Based Multi-Focus Image Fusion Technique. *Electronics* **2020**, *9*, 472. [[CrossRef](#)]
13. James, A.P.; Dasarathy, B.V. Medical image fusion: A survey of the state of the art. *Inf. Fusion* **2014**, *19*, 4–19. [[CrossRef](#)]
14. Hermessi, H.; Mourali, O.; Zagrouba, E. Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain. *Neural Comput. Appl.* **2018**, *30*, 2029–2045. [[CrossRef](#)]
15. Wang, L.; Li, B.; Tian, L.F. Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients. *Inf. Fusion* **2014**, *19*, 20–28. [[CrossRef](#)]
16. Li, H.; Qiu, H.; Yu, Z.; Li, B. Multifocus image fusion via fixed window technique of multiscale images and non-local means filtering. *Signal Process.* **2017**, *138*, 71–85. [[CrossRef](#)]
17. Yang, S.; Wang, M.; Jiao, L.; Wu, R.; Wang, Z. Image fusion based on a new contourlet packet. *Inf. Fusion* **2010**, *11*, 78–84. [[CrossRef](#)]
18. Yang, Y. A novel DWT based multi-focus image fusion method. *Procedia Eng.* **2011**, *24*, 177–181. [[CrossRef](#)]

19. Li, H.; Qiu, H.; Yu, Z.; Zhang, Y. Infrared and visible image fusion scheme based on NSCT and low-level visual features. *Infrared Phys. Technol.* **2016**, *76*, 174–184. [[CrossRef](#)]
20. Nencini, F.; Garzelli, A.; Baronti, S.; Alparone, L. Remote sensing image fusion using the curvelet transform. *Inf. Fusion* **2007**, *8*, 143–156. [[CrossRef](#)]
21. Du, J.; Li, W.; Xiao, B. Union laplacian pyramid with multiple features for medical image fusion. *Neurocomputing* **2016**, *194*, 326–339. [[CrossRef](#)]
22. Li, S.; Kang, X.; Hu, J. Image fusion with guided filtering. *IEEE Trans. Image Process.* **2013**, *22*, 2864–2875. [[PubMed](#)]
23. Liu, Y.; Chen, X.; Ward, R.K.; Wang, Z.J. Image Fusion With Convolutional Sparse Representation. *IEEE Signal Process. Lett.* **2016**, *23*, 1882–1886. [[CrossRef](#)]
24. Liu, Y.; Chen, X.; Ward, R.K.; Wang, Z.J. Medical Image Fusion via Convolutional Sparsity Based Morphological Component Analysis. *IEEE Signal Process. Lett.* **2019**, *26*, 485–489. [[CrossRef](#)]
25. Liu, Y.; Chen, X.; Cheng, J.; Peng, H. A medical image fusion method based on convolutional neural networks. In Proceedings of the 2017 20th International Conference on Information Fusion (Fusion), Xi'an, China, 10–13 July 2017.
26. Yang, B.; Li, S. Visual attention guided image fusion with sparse representation. *Optik (Stuttg.)* **2014**, *125*, 4881–4888. [[CrossRef](#)]
27. Liu, Y.; Liu, S.; Wang, Z. A general framework for image fusion based on multi-scale transform and sparse representation. *Inf. Fusion* **2015**, *24*, 147–164. [[CrossRef](#)]
28. Li, S.; Yin, H.; Fang, L. Group-sparse representation with dictionary learning for medical image denoising and fusion. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 3450–3459. [[CrossRef](#)]
29. Kim, M.; Han, D.K.; Ko, H. Joint patch clustering-based dictionary learning for multimodal image fusion. *Inf. Fusion* **2016**, *27*, 198–214. [[CrossRef](#)]
30. Basar, S.; Adnan, A.; Khan, N.H.; Haider, S. Color image segmentation using K-mean classification on RGB histogram. In Proceedings of the Recent Advances In Telecommunications, Informatics And Educational Technologies, Istanbul, Turkey, 15–17 December 2014; pp. 257–262.
31. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.; Ciompi, F.; Ghafoorian, M. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)]
32. Wang, K.; Zheng, M.; Wei, H.; Qi, G.; Li, Y. Multi-modality medical image fusion using convolutional neural network and contrast pyramid. *Sensors* **2020**, *20*, 2169. [[CrossRef](#)]
33. Zhang, Q.; Levine, M.D. Robust multi-focus image fusion using multi-task sparse representation and spatial context. *IEEE Trans. Image Process.* **2016**, *25*, 2045–2058. [[CrossRef](#)] [[PubMed](#)]
34. Xing, C.; Wang, M.; Dong, C.; Duan, C.; Wang, Z. Using Taylor Expansion and Convolutional Sparse Representation for Image Fusion. *Neurocomputing* **2020**, *402*, 437–455. [[CrossRef](#)]
35. Li, Y.; Sun, Y.; Huang, X.; Qi, G.; Zheng, M.; Zhu, Z. An image fusion method based on sparse representation and Sum Modified-Laplacian in NSCT Domain. *Entropy* **2018**, *20*, 522. [[CrossRef](#)]
36. Li, W.; Jia, L.; Du, J. Multi-Modal Sensor Medical Image Fusion Based on Multiple Salient Features with Guided Image Filter. *IEEE Access* **2019**, *7*, 173019–173033. [[CrossRef](#)]
37. Arif, M.; Wang, G. Fast curvelet transform through genetic algorithm for multimodal medical image fusion. *Soft Comput.* **2020**, *24*, 1815–1836. [[CrossRef](#)]
38. Kaur, M.; Singh, D. Fusion of medical images using deep belief network. *Cluster Comput.* **2020**, *23*, 1439–1453. [[CrossRef](#)]
39. Shahdoosti, H.R.; Mehrabi, A. Multimodal image fusion using sparse representation classification in tetrolet domain. *Digital Signal Process.* **2018**, *79*, 9–22. [[CrossRef](#)]
40. Ying, Z.; Li, G.; Gao, W. A Bio-Inspired Multi-Exposure Fusion Framework for Low-light Image Enhancement. *arXiv* **2017**, arXiv:1711.00591.
41. Yan, J.; Li, J.; Fu, X. No-Reference Quality Assessment of Contrast-Distorted Images using Contrast Enhancement. *arXiv* **2019**, arXiv:1904.08879.
42. Poddar, S.; Tewary, S.; Sharma, D.; Karar, V.; Ghosh, A.; Pal, S.K. Non-parametric modified histogram equalisation for contrast enhancement. *IET Image Process.* **2013**, *7*, 641–652. [[CrossRef](#)]
43. Schindelin, J.; Rueden, C.T.; Hiner, M.C.; Eliceiri, K.W. The ImageJ ecosystem: An open platform for biomedical image analysis. *Mol. Reprod. Dev.* **2015**, *82*, 518–529. [[CrossRef](#)] [[PubMed](#)]

44. Pinheiro, P.O.; Collobert, R. From image-level to pixel-level labeling with convolutional networks. In Proceedings of the 28th IEEE conference on computer vision and pattern recognition, CVPR 2015, Boston, MA, USA, 7–12 June 2015; pp. 1713–1721.
45. Gao, W.; Zhang, X.; Yang, L.; Liu, H. An improved Sobel edge detection. In Proceedings of the 3rd International Conference on Computer Science and Information Technology, Chengdu, China, 9–11 July 2010; pp. 67–71.
46. Zhang, H.; Patel, V.M. Convolutional sparse and low-rank coding-based image decomposition. *IEEE Trans. Image Process.* **2017**, *27*, 2121–2133. [[CrossRef](#)]
47. Wohlberg, B. Efficient algorithms for convolutional sparse representation. *IEEE Trans. Image Process.* **2016**, *25*, 301–315. [[CrossRef](#)] [[PubMed](#)]
48. Yang, B.; Li, S. Pixel-level image fusion with simultaneous orthogonal matching pursuit. *Inf. Fusion* **2012**, *13*, 10–19. [[CrossRef](#)]
49. Jiang, Y.; Wang, M. Image fusion with morphological component analysis. *Inf. Fusion* **2014**, *18*, 107–118. [[CrossRef](#)]
50. Hossny, M.; Nahavandi, S.; Vreighton, D. Comments on information measure for performance of image fusion. *Electron. Lett.* **2008**, *44*, 1066–1067. [[CrossRef](#)]
51. Haghghat, M.B.A.; Aghagolzadeh, A.; Seyedarabi, H. A non-reference image fusion metric based on mutual information of image features. *Comput. Electr. Eng.* **2011**, *37*, 744–756. [[CrossRef](#)]
52. Petrovi, V.S.; Xydeas, C.S. Sensor noise effects on signal-level image fusion performance. *Inf. Fusion* **2003**, *4*, 167–183. [[CrossRef](#)]
53. Han, Y.; Cai, Y.; Cao, Y.; Xu, X. A new image fusion performance metric based on visual information fidelity. *Inf. Fusion* **2013**, *14*, 127–135. [[CrossRef](#)]
54. Yu, B.; Jia, B.; Ding, L.; Cai, Z.; Wu, Q. Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion. *Neurocomputing* **2016**, *182*, 1–9. [[CrossRef](#)]
55. Yin, M.; Liu, X.; Liu, Y.; Chen, X. Medical Image Fusion With Parameter-Adaptive Pulse Coupled-Neural Network in Nonsampled Shearlet Transform Domain. *IEEE Trans. Instrum. Measur.* **2019**, *68*, 49–64. [[CrossRef](#)]
56. Zhu, Z.; Chai, Y.; Yin, H.; Li, Y.; Liu, Z. A novel dictionary learning approach for multi-modality medical image fusion. *Neurocomputing* **2016**, *214*, 471–482. [[CrossRef](#)]
57. Madanagopal, R. Medical fusion imaging: Paving the way for better diagnosis of tumours. *Health Manag.* **2014**, 1–3. Available online: <https://healthmanagement.org/c/healthmanagement/issuearticle/medical-fusion-imaging-paving-the-way-for-better-diagnosis-of-tumours> (accessed on 1 September 2020).
58. Amini, N.; Fatemizadeh, E.; Behnam, H. MRI-PET image fusion based on NSCT transform using local energy and local variance fusion rules. *J. Med. Eng. Technol.* **2014**, *38*, 211–219. [[CrossRef](#)]
59. Demsar, J. Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* **2006**, *7*, 1–30.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).