



# DTCWTASODCNN: DTCWT based Weighted Fusion Model for Multimodal Medical Image Quality Improvement with ASO Technique & DCNN

Vineeta Singh\* & Vandana Dixit Kaushik

<sup>1</sup>Department of Computer Science and Engineering, Harcourt Butler Technical University, East Campus, Nawabganj, Kanpur 208 002, Uttar Pradesh, India

*Received 15 October 2021; revised 16 July 2022; accepted 17 July 2022*

Medical image fusion approaches are sub-categorized as single-mode as well as multimodal fusion strategies. The limitations of single-mode fusion approaches can be resolved by introducing a multimodal fusion approach. Multimodal medical image fusion approach is formed by integrating two or more medical images of similar or dissimilar modalities aims to enhance the image quality and to preserve the image information. Hence, this paper introduced a new way to meld multimodal medical images via utilizing developed weighted fusion model relied on Dual Tree Complex Wavelet Transform (DTCWT) for fusing the multimodal medical image. Here, the two medical images are considered for image fusion process and we have implied DTCWT to the medical images for generating four sub-bands partition of the source medical images. The Renyientropy-based weighted fusion model is used to combine the weighted coefficient of DTCWT of images. The final fusion process is carried out using Atom Search Sine Cosine Algorithm (ASSCA)-based Deep Convolutional Neural Network (DCNN). Moreover, the simulation work output demonstrated for developed fusion model gained the superior outcomes relied on key indicators named as Mutual Information i.e. MI, Peak Signal to Noise Ratio abbreviated as PSNR as well as Root Mean Square Error, in short RMSE with the values of 1.554, 40.45 dB as well as 5.554, correspondingly.

**Keywords:** Atom search optimization, DCNN, Image quality improvement, Medical imaging, Weighted average model

## Introduction

The image fusion process has been performed at various forms of data, some of them are feature, signal, pixel and representative forms.<sup>1</sup> The image fusion process comes under the domain of image processing. Some of the image processing techniques involve deblurring, denoising, image enhancement and so on.<sup>2</sup> The intention of image processing technique is to improve the modalities of images.<sup>3</sup> For medical applications, the image quality enhancement technique is utilized for improving the contrast of processed images. In traditional fusion scheme, the fusing of medical images with complimentary information produces various issues in the image modalities, which can be resolved with Renyi-entropy fusion schemes. Image fusion is a process of joining various medical features into a unique image so that quality of images may be enhanced without losing the medical image information. In medical imaging research field, image fusion is one of the effective and promising diagnostic schemes employed. The melded

pictures procured from different modalities or devices are significant in various applications, involving human perception, robotics, intelligent systems<sup>4</sup>, visual recognition<sup>5</sup>, object detection, concealed weapon detection, video surveillance, under sea detection, object tracking, and clinical diagnosis etc.<sup>6</sup> Primary intention for fusing medical images is to achieve a high/better resolution image that facilitates large amount of information further helping in accurate and faster diagnosis. For that, two pictures of same organ are combined, and afterward accumulate the colossal amount of data for the diagnosis of organ, after that it is compared with non-melded pictures.<sup>7</sup>

Mainly two varieties of image fusion strategies exist, named as spatial domain image fusion strategies as well as transform domain image fusion strategies. Image fusion strategies involving spatial domain, considers pixel values for gaining desired result. Some of the spatial domain approaches involves schemes dependent on high pass filtering as well as hue intensity saturation technique, Principal Component Analysis (PCA) as well as Brovey scheme.<sup>8</sup> Moreover, the spatial domain approach operates based on the concept of spectral distortion as

\*Author for Correspondence  
E-mail: vineeta.singh.cs@gmail.com

well as spatial distortion in melded picture. Meanwhile, in transform domain-dependent approaches, at first, medical image is transformed into different (another) domain and further fusion techniques are carried out based on transformed image. Hence, the inverse transform is applied for gaining resultant image.<sup>9</sup> The fusion process is done using various levels of information including feature, signal, and pixel levels. The multimodal medical image fusion approaches are categorized into three levels, such as decision level, pixel and feature level. The pixel transformation is limited using pixel-level fusion, which preserves the data from fused images. Moreover, the decision level fusion processes are carried out based on feature selection schemes using various systems, like forecast, fuzzy logic as well as voting. In addition, the feature-level fusion processes are done based on texture and fine detail and textures and edges of images. Some of the widely transforms-based image fusion schemes are Discrete Wavelet Transform (DWT), and pyramid transform techniques.<sup>10,11</sup> Other image fusion technique involves curvelet, complex wavelet, contourlet and so on.<sup>12</sup>

The main motivation of this research is to propose a novel fusion technique by designing and developing DTCWT-based weighted fusion model to fuse multimodal medical images. Initially, the two input images are processed separately using DTCWT in order to partition the images into sub-bands. After that, the Renyi-entropy-based weighted fusion model is used to fuse the weighted coefficient of DTCWT. Then, the ASSCA-based Deep CNN is employed to perform the final fusion for acquiring the multimodal medical images.

The main contribution of this research is:

*Develop DTCWT-based Weighted Fusion Model:*

The multimodal fusion process is done by partitioning the images into smaller sub-bands using DTCWT, where the weighted coefficient of DTCWT is fused using Renyi-entropy.

### Motivation and Related Work

The melding of the images have been accompanied through fusion process for combining the data from numerous images of same scenario into an individual image, which normally comprises the imperative features of same images. The resultant combined picture is more suitable for human perception along with for machine vision and for remaining image processing strategies. This is the motivation of image fusion technique.

Wendazhao *et al.*<sup>13</sup> developed the Multi-level CNN (MLCNN) for performing the multi-focus picture fusion process. Moreover, this method extracted the effective multilevel features, which was provided an optimal outcome. Although, the MLCNN fusion process attained an optimal fusion outcome, the computational complexity of this method was high. Boyuan Ma *et al.*<sup>14</sup> devised the deep learning-dependent fusion scheme to accompany the multi-focus picture fusion process. Here, the fusion process was done based on the outcome of decision map with pixel-based weighted average rule. Moreover, the errors exists in the fused images were removed using consistency verification approaches. However, this method was failed to process with real world applications. Hao Zhang *et al.*<sup>15</sup> modelled Generative Adversarial Network (GAN) with two kind of constraints named as adaptive as well as gradient joint constraints for achieving the fusion of multi-focus images. Here, the Laplacian operator and gradient map was applied to the input images in order to perform the fusion. The processing time of multi-image fusion process was low. However, this method attained maximum loss functions, which produced the information loss. Chionmaya *et al.*<sup>16</sup> designed the adaptive Dual Channel Pulse coupled neural network abbreviated as PA-DCPCNN scheme for performing multi-focus picture fusion. Further images were converted into sub-groups utilizing Non-sub Sampled Contourlet Transform (NSCT). The sub-groups were fused using PA-DCPCNN for acquiring the fused outcome. However, this method was failed to process with multi-exposure and infrared-visible image fusion.

### Challenges

The challenges faced during the assessment of medical image fusion approaches are organized as follows.

The developed MLCNN model by Zhao *et al.*<sup>13</sup> achieved good performance for the images with complex features. However, the input images with great misregistered scenario may produce some artifacts, which reduced the performance. Hence, the challenge lies on introducing more robust fusion scheme based on CNN for overcoming this issue.

The image fusion model devised via Ma *et al.*<sup>14</sup> was formed by combining unsupervised learning and conventional image processing approaches, which provides an effective fused outcome. However, this method was not suitable for new image fusion

processes, like infrared visible and multi-exposure fusion.

The effectiveness of image fusion model can be enhanced by designing broader multi-modal deep learning process in order to merge multi-focus images.<sup>15</sup> Moreover, a unified model can be deliberated to perform multiple multi-modal image fusion process, involving medical image, infrared as well as visible image fusion.

The challenges of PA-DCPCNN model lies on subjecting some computationally effective multi-scale multi-direction transform for decreasing cost of computation.<sup>16</sup>

### Experimental Work

#### Proposed Fusion Model

Multimodal clinical picture (medical image) combination is the mix of at least two (may involve more than two images) pictures at same or different modalities from same scene. Through this research, we are aiming towards to augment the quality of image without affecting its local information. This paper devised the developed DTCWT-based weighted fusion model for fusing the multimodal medical image. Here, two images are considered as the input, and is fed to the DTCWT<sup>17</sup> in order to convert the images into sub-groups. Once the images are converted into sub-groups, further fusion procedure is implemented based on weighted model relied on Renyi entropy. The output attained from the fusion is then given to the final fusion phase, which is done based on Deep CNN. The training process of Deep CNN<sup>18,19</sup> is done using atom search optimization algorithm<sup>20</sup> in combination with sine cosine algorithm<sup>21</sup> named as ASSCA. We have demonstrated the details of proposed fusion technique in Fig. 1.

Let us assume  $Q_1$  and  $Q_2$  be the two multimodal images acquired from database, which contains  $n$  number of images. Initially, the input images are partitioned using DTCWT such that the four sub bands are obtained. These bands are fused using Renyi entropy-based weighted model, thereby initial fused output is obtained.

#### Dual Tree Complex Wavelet Transform

We have utilized DTCWT for computation of the complex transform of image with two dissimilar two-dimensional DWT (2D-DWT) scheme. The decomposed image contains a real and imaginary part. The block diagram of DTCWT is described in Fig. 2, which contains the top and bottom filter bank. From the figure, the top filter bank with low pass filter, in short named as LPF, as well as high pass filter, in short, named as HPF, is represented as  $x_0(u)$  and  $x_1(u)$ , respectively, which is used to calculate the wavelet and scale coefficient of real part. Similarly, the bottom filter bank with LPF as well as HPF is represented as  $y_0(u)$  and  $y_1(u)$ , which is used to measure the wavelet and scale coefficient of imaginary part. In the initial phase, the image performs line transformation, such as  $[y_0(u), y_1(u)]$ , and then it performs rank transformation, like  $[x_0(u), x_1(u)]$ . In the second phase, the image performs line transformation, like  $[x_0(u), x_1(u)]$ , and then it performs rank transformation, like  $[y_0(u), y_1(u)]$ . For every level of DTCWT, the image is decomposed into four sub-regions, namely Low-High (LH), High-High (HH), High-Low (HL) as well as Low-Low (LL).

Let us assume the noisy image, which is mathematically represented as,

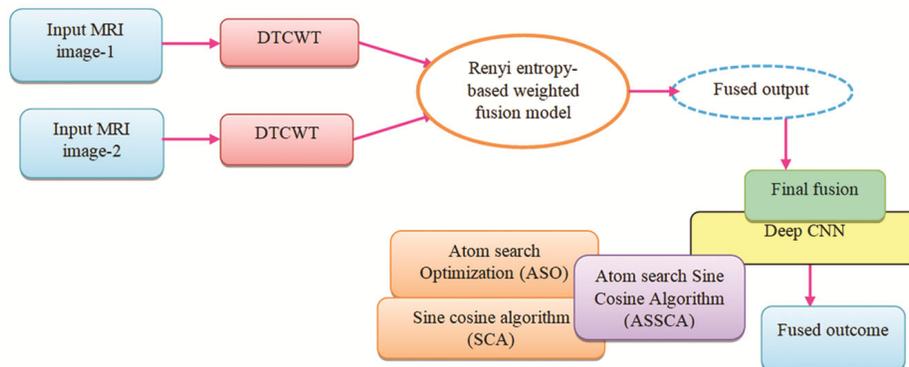


Fig. 1 — Block diagram of developed DTCWT-based weighted fusion model for Multimodal medical image fusion

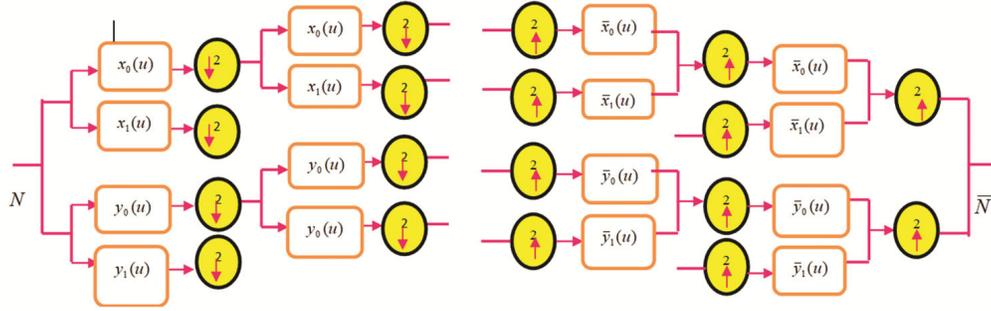


Fig. 2 — Block diagram of DTCWT

$$Q_{noisy} = Q_{original} \otimes V \quad \dots(1)$$

where,  $Q_{noisy}$  indicates the noisy image,  $\otimes$  represents the operator of noise model,  $Q_{original}$  depicts the original image.

$$J = +V \quad \dots(2)$$

where,  $J$  represents the wavelet coefficient of captured image,  $\lambda$  indicates the wavelet coefficient of original image, which is obtained from noise  $V$ , and  $V$  represents the wavelet coefficient of noise. Here,  $\lambda$  is measured using maximum posterior probability, which is represented as,

$$\hat{\lambda}(J) = \arg \left( \max_{\lambda} K_{\lambda|J}(\lambda|J) \right) \quad \dots(3)$$

where,  $K$  indicates the posterior probability function. Equation (3) can be modified as,

$$\hat{\lambda}(J) = \arg \left( \max_{\lambda} (K_V(V) \times K_{\lambda}(\lambda)) \right) \quad \dots(4)$$

$$K_V(V) = \left( \frac{1}{2\pi\delta^2} \right) \times \exp \left( -\frac{V_1^2 + V_2^2}{2\delta^2} \right) \quad \dots(5)$$

where,  $K_{\lambda}(\lambda)$  indicates the bivariate probability distributed function, which is indicated as,

$$K_{\lambda}(\lambda) = \left( \frac{3}{2\pi\delta_{\lambda}^2} \right) \times \exp \left( -\frac{\sqrt{3}}{\delta} \sqrt{\lambda_1^2 + \lambda_2^2} \right) \quad \dots(6)$$

where,  $\delta_{\lambda}^2$  indicates the signal variance,  $\lambda_1$  indicates the wavelet coefficient at first level,  $\lambda_2$  represents the wavelet coefficient at next level. The posterior probability with upper value  $\lambda_1$  is represented as,

$$\hat{\lambda}_1 = \frac{\left( \sqrt{J_1^2 + J_2^2} - \frac{\sqrt{3}\lambda^2}{\delta_{\lambda}} \right)_{+}}{\sqrt{J_1^2 + J_2^2}} \times J_1 \quad \dots(7)$$

Moreover, the noise variance  $\delta^2$  is estimated using median filter, which is given by,

$$\delta^2 = \frac{\text{Median}(|J_i|)}{0.675}; \text{ where, } J_i \in \text{subband}_{HH} \quad \dots(8)$$

In this model, there is a need to calculate

$$\delta^2 = \hat{\delta}^2 + \delta \quad \dots(9)$$

$$\hat{\delta}_J^2 = \frac{1}{S} \sum_{J_i \in D(y)} J_i^2 \quad \dots(10)$$

where,  $S$  indicates the wavelet coefficient size.

From Eq. (9), the value of  $\hat{\delta}^2$  can be rewritten as,

$$\hat{\delta}^2 = \sqrt{(\hat{\delta}_J^2 + \delta)} \quad \dots(11)$$

Moreover, the obtained outcome of DTCWT is represented as  $G_1^{DTCWT}$  and  $G_2^{DTCWT}$ .

### Renyi-Entropy-based Weighted Fusion

This section explains the Renyi-entropy-based weighted fusion model for combining the multimodal medical image for attaining the fused clinical picture (medical image). The Renyi entropy is a factor, which is used to compute the important information exist in the image. The advantage of Renyi entropy is that it reduces the information loss while performing image fusion. The Renyi-entropy-based weighted fusion is expressed by combining weighted coefficient of DTCWT of images with Renyi entropy values. Here, the Renyi-entropy-based weighted fusion model

combines the DTCWT of input image  $G_1$  and  $G_2$  with Renyi entropy of image  $G_1$  and  $G_2$ , which is represented as,

$$f = \mu G_1^{DTCWT} + (1 - \mu) G_2^{DTCWT} \quad \dots(12)$$

where,  $\mu$  indicates the Renyi-entropy of  $(G_1, G_2)$ , and  $G_1^{DTCWT}$  and  $G_2^{DTCWT}$  indicates the DTCWT of image  $I_1$  and  $I_2$ . The fused output of Renyi-entropy-based weighted fusion model is represented as  $f$ .

**ASSCA-based Deep CNN for Final Fusion**

This section explains the final fusion process using Deep CNN with ASSCA. The final fusion process is performed to augment the contrast of fused image.

The input of Deep CNN is illustrated as  $f$ . Here, the fusion procedure is performed using Deep CNN in which the optimal fusion score is selected using ASSCA. The ASSCA is obtained by integrating ASO and SCA and is acquired by adapting the benefits of both optimization techniques.

**Structural Design of Deep CNN**

The Deep CNN architecture comprised of three layers, like Fully Connected (FC), convolutional (conv), and pooling (POOL) layers, where each layer performs separate operations. Here, FC layer is responsible for performing fusion process.

-*Conv Layers*: Here, conv layer extracts important information exists in the image, which is modelled as,

$$E = \{E_1, E_2, \dots, E_b, \dots, E_p\} \quad \dots (13)$$

where,  $p$  signifies total conv layers, and  $E_b$  denotes  $b^{th}$  conv layer. The units situated in  $(l, m)$  attains output, which is expressed as,

$$(E_\sigma^b)_{l,m} = (\varepsilon_\sigma^b)_{l,m} + \sum_{y=1}^{Y_1^{\sigma-1}} \sum_{r=-h_1^\sigma}^{h_1^\sigma} \sum_{v=-h_2^\sigma}^{h_2^\sigma} (Z_{b,y}^\sigma)_{r,v} * (E_\sigma^{b-1})_{l+z,m+v} \quad \dots(14)$$

where,  $*$  denotes conv operator,  $(E_\sigma^b)_{l,m}$  denotes static feature maps from convolution layer  $\sigma$ , and  $Y_1^{l-1}$  denotes overall feature maps, and  $(Z_{b,y}^\sigma)_{r,v}$  signifies

weights, which is trained using ASSCA algorithm,  $\varepsilon_\sigma^b$  signifies the bias of  $\sigma^{th}$  convolution layer,  $(E_\sigma^{b-1})_{l+z,m+v}$  signifies the feature map of previous convolution layer and  $Z_{b,y}^\sigma$  denotes the kernel function.

-*ReLU*: ReLU portrays an activation module, which accelerates a maximum efficiency. The outcome generated with ReLU layer is given as,

$$E_s^q = fun(E_y^{b-1}) \quad \dots(15)$$

where,  $fun()$  signifies activation function in  $y^{th}$  layer.

-*FC Layers*: The patterns produced using conv as well as pooling layers are exposed as an input to FC layers to perform image fusion. The output produced using FC is modelled as,

$$V_\sigma^b = W(a_\sigma^b) \text{ with } a_\sigma^b = \sum_{y=1}^{Y_1^{\sigma-1}} \sum_{r=-h_1^\sigma}^{h_1^\sigma} \sum_{v=-h_2^\sigma}^{h_2^\sigma} (Z_{b,y}^\sigma)_{r,v} * (E_\sigma^{b-1})_{l+z,m+v} \quad \dots(16)$$

where,  $(Z_{b,y}^\sigma)_{r,v}$  signifies weight associating  $(l, m)$  in  $r^{th}$  feature map of layer  $\sigma - 1$  and  $s^{th}$  unit in layer  $\sigma$ . The output of Deep CNN is illustrated as  $F_a^*$ , which is trained using ASSCA.

**ASSCA for the Training Process of Deep CNN**

This section explains the ASSCA for training the weights of Deep CNN. Here, ASSCA model is adapted to select optimal value in order to perform an effective fusion. The ASSCA is modelled by amalgamation of ASO<sup>20</sup> as well as SCA<sup>21</sup>, that's shaped with the aid of using the advantages of each algorithm. The design of SCA is based on the sine and cosine properties, which provides better searching and global optimal outcome. Moreover, the advantage of SCA is that, it has local optima avoidance and poses high exploration. ASO is designed based on the principle of molecular dynamics, which handles the complex optimization issues. Hence, ASO algorithm is merged with SCA in order to select an optimal fusion score, which ensures the combined ASSCA model provides a better balance among exploration as well as exploitation. The final updated equation of ASSCA algorithm is portrayed as follows.

$$C_o(z+1) = \frac{c_o(z)j_1j_2\text{Sin}(j_2)}{c_o(z)j_1j_2\text{Sin}(j_2) - \eta e^{-T}} \left[ \frac{C_o(z) + \text{Rand}_{j_1}(t) - \lambda \left(1 - \frac{z-1}{I}\right) e^{-\frac{20z}{I}}}{\sum_{k \in \text{Best}} \frac{\text{Rand}_{j_1} [2 \times (d_{ob}(z))^3] - (d_{ob})^7}{c_o(z)}} \right] \dots(17)$$

$$\frac{C_i(z) - C_o(z)}{\|C_o(z), C_i(z)\|_2} \frac{\eta e^{-T} C_o(z) [1 - j_1 \text{Sin}(j_2) + j_1 j_3 \text{Sin}(j_2)]}{c_o(z) j_1 j_3 \text{Sin}(j_2)}$$

where,  $j_1$  signify movement direction,  $j_2$  portrays the movement towards outer target,  $j_3$  signifies target random weight, and  $j_4$  switches among sine and cosine components,  $C_o(z)$  illustrates the current solution,  $\lambda$  shows depth weight,  $I$  designate highest iterations,  $T_{best}$  signifies optimal fitness value,  $c_o(z)$  signifies mass at  $z^{th}$  iteration of  $o^{th}$  atom, multiplier weight has been illustrated through function  $\eta$ ,  $i_o(z+1)$  indicates at  $(z+1)$  iteration the  $o^{th}$  atom velocity,  $C_o(z)$  denotes at  $z^{th}$  iteration the  $o^{th}$  atom location,  $\text{Rand}_o$  designate arbitrary number in  $[0,1]$ ,  $T_o(z)$  denotes acceleration at  $z^{th}$  iteration of  $o^{th}$  atom, and further  $T_o(z)$  indicates velocity of  $o^{th}$  atom at  $z^{th}$  iteration.

**Fitness Measure:** Here, the fitness function with minimum value is considered as an optimal outcome, which is evaluated as,

$$MS_{err} = \frac{1}{e} \sum_{a=1}^e [F_a - F_a^*]^2 \dots(18)$$

where,  $F_a$  denotes an expected outcome, and  $F_a^*$  signifies the predicted outcome of Deep CNN,  $e$  characterizes data sample count, where  $1 < a \leq e$ .

**Simulation Details**

The developed multimodal clinical picture (medical image) fusion algorithm is executed over MATLAB tool, Windows 10 OS, Intel I3 core processor as well as 2GB RAM using BRATS 2018 dataset. The BRATS 2018 dataset<sup>22</sup> comprised of multimodal 3D brain MRIs as well as brain tumour segmentation of ground truth samples involving per case 4 MRI modalities, namely FLAIR, T2, T1c as well as T1. The segmentation images contain 3 tumor sub regions, such as augmenting tumor, necrotic as well as peritumoral edema and non-enhancing tumor core. Moreover, the medical images are lending out from 19 institutions involving distinct MRI scanners.

**Evaluation Measure**

The performance of developed fusion model is assessed using various performance measures with MI, RMSE and PSNR.

**MI:** MI is utilized to ascertain the measure of shared reliance between the two pictures, and is computed as,

$$U = \frac{1}{2} [R[C(w, z), L(w, z)]MI[D(w, z), L(w, z)]] \dots(19)$$

where,  $C(w, z)$  and  $D(w, z)$  represents the two input, while the fused image is depicted through  $L(w, z)$ .

**RMSE:** RMSE is utilized to gauge the error produced via fusion model, and is computed as,

$$M = \frac{1}{2} [M[C(w, z), L(w, z)] + M[D(w, z), L(w, z)]] \dots(20)$$

where,  $C(w, z)$  and  $D(w, z)$  represents the two input, while the fused image is denoted as  $L(w, z)$ .

**PSNR:** Definition for PSNR may be described as the ratio of maximal possible power of signal to affected noise and its unit is denoted as decibel (dB), which is described as,

$$PSNR = 10 \log_{10} \left( \frac{R_{max}^2}{MSE} \right) \dots(21)$$

where,  $R_{max}$  denotes the peak value of pixel, while MSE resembles mean square error.

**Results and Discussion**

Here authors explicate the experimental outcomes of devised DTCWT-based weighted fusion model for multimodal medical image fusion. We have shown input image samples through Fig. 3a and b, while final fused output image is shown through Fig. 3c).

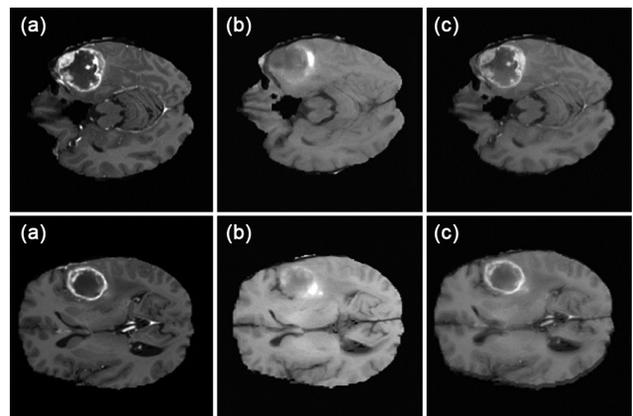


Fig. 3 — Experimental results:(a) Input MRI image-1, (b) Input MRI image-2, (c) Final fused image (Row I and II)

**Comparative Methods & Analysis**

The developed fusion model is compared with existing techniques, like DWT+Holoentropy Whale Fusion (HW Fusion) +SP – Whale<sup>12</sup>, DWT+undecimated DWT (UD WT)+ Genetic Algorithm (GA)<sup>23</sup>, DWT+ASSCA+Renyi entropy<sup>24</sup> and ASSCA-based Deep CNN.<sup>25</sup> The effectiveness of developed fusion model is evaluated with performance measure by changing the db2 and sym2 of DWT with various numbers of images.

**Image Fusion Assessment based on Sym2**

This section describes the assessment of image fusion model dependent on sym2 by changing the image numbers regarding the assessment parameters, like MI, PSNR, and RMSE. We have demonstrated assessment of MI through Fig. 4 a. When the number of image is 1, then the MI value attained by the DWT+HW Fusion+SP – Whale method is 0.9924, DWT+UDWT+GA measured the MI values of 1.1724, DWT+ASSCA+Renyi Entropy model attained the MI values of 1.2063, ASSCA-based Deep CNN method attained the MI values of 1.402, and the developed DTCWT-based weighted fusion model measured the MI values of 1.4462, correspondingly. When the number of image is 5, then the Maximal MI value of 1.554 has been attained through proposed DTCWT-based weighted fusion scheme while others recent State-of-the-Art fusion models yielded MI values of 0.972, 1.175, 1.234, 1.5 for DWT+HW Fusion+SP-Whale, DWT+UDWT+GA, DWT+ASSCA+Renyi Entropy and ASSCA based Deep CNN fusion model respectively. The assessment of PSNR is shown in Fig. 4 b. When the number of image is 3, then the PSNR value measured by the developed fusion model is 29.83 dB, whereas the existing approaches attained the PSNR values of 19.36 dB, 20.75 dB, 21.63 dB and 26.66 dB. We have enlisted the assessment of RMSE through Fig. 4c. The RMSE value attained by the developed DTCWT-based weighted fusion model is 5.554, and the existing approaches attained the RMSE value of 24.90, 20.53, 16.15, and 7.058 for the number of images is 5.

**Image Fusion Assessment based on db2**

This section describes the assessment of image fusion model based on db2 by varying the number of images with respect to the performance metrics. The assessment of mutual information is shown in Fig. 5a. The MI value attained by the developed DTCWT-based weighted fusion model is 1.456, and the

existing approaches attained the MI value of 1.004, 1.175, 1.216, and 1.403 for the number of images is 2. The assessment of PSNR is shown in Fig. 5b. When the number of image is 4, then the PSNR value attained by the DWT+HWFusion+SP – Whale method is 18.28 dB, DWT+UDWT+GA measured the PSNR value of 22.01 dB, DWT+ASSCA+ Renyi Entropy model attained the PSNR value of 23.66 dB, ASSCA\_basedDeepCNN method attained the PSNR value of 26.95 dB and the developed DTCWT-based weighted fusion model measured the PSNR value of

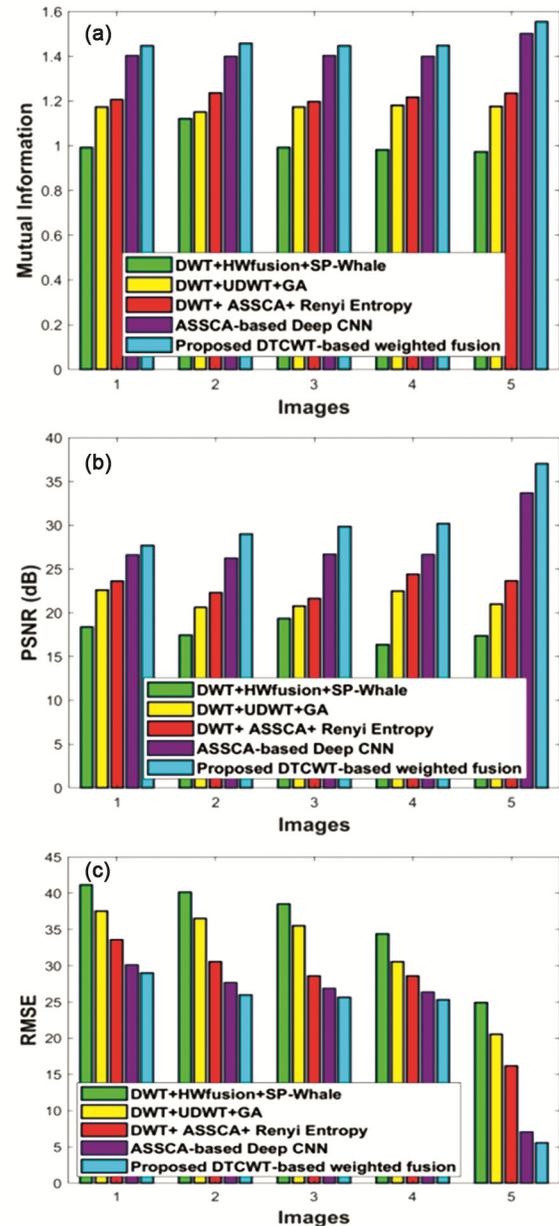


Fig. 4 — Assessment of fused model using sym2 relied on a) MI, b) PSNR and c) RMSE

Table 1 — Comparative Analysis of the Fusion Algorithms

Variations	Evaluation metrics	DWT+HWFusion+SP-Whale	DWT+UDWT+T+GA	DWT+ASSCA+Renyi Entropy	ASSCA_based Deep CNN	Proposed Algorithm
Sym2	MI	0.972	1.175	1.234	1.5	<b>1.554</b>
	PSNR (dB)	17.36	20.98	23.64	33.65	37.03
	RMSE	24.90	20.53	16.15	7.058	<b>5.554</b>
db2	MI	0.992	1.174	1.251	1.398	1.432
	PSNR (dB)	18.28	22.66	23.86	30.60	<b>40.45</b>
	RMSE	23.59	21.16	16.60	9.740	6.598

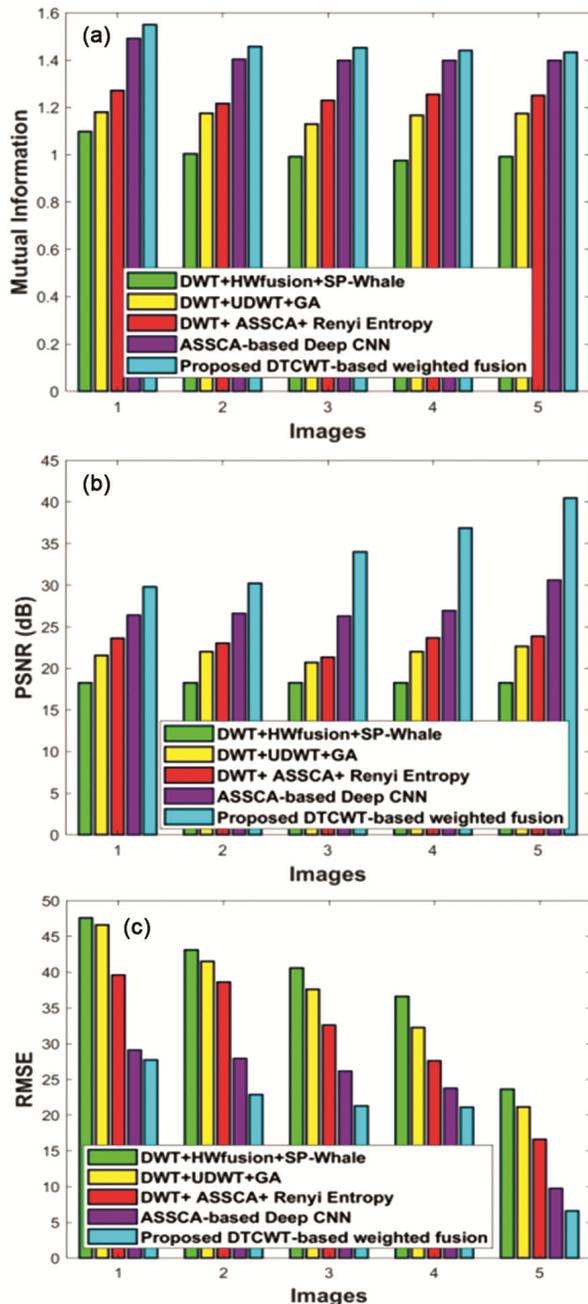


Fig. 5 — Assessment of fused model using db2 relied on (a) MI, (b) PSNR, (c) RMSE

36.87 dB, correspondingly. When the number of image is 5, Proposed DTCWT based weighted fusion scheme yielded maximum PSNR value of 40.45 dB based on db2 while other existing techniques yielded PSNR values of 18.28 dB, 22.66 dB, 23.86 dB, 30.60 dB respectively for DWT+HWFusion+SP-Whale, DWT+UDWT+GA, DWT+ASSCA+Renyi Entropy, ASSCA\_based Deep CNN correspondingly. The assessment of RMSE is shown in Fig. 5c. If the number of image is 3, then the RMSE value measured by the developed fusion model is 21.28, whereas the existing approaches attained the RMSE values of 40.59, 37.57, 32.60 and 26.15.

**Comparative Discussion**

We have demonstrated the comparative discussion of developed DTCWT-based weighted fusion model through Table 1 in a summarized form. From the table, the developed DTCWT-based weighted fusion model attained the Maximal MI value of 1.554 based on sym2 and existing techniques yielded MI value of 0.972, 1.175, 1.234, 1.5 respectively.

Our proposed model yielded minimal RMSE value of 5.554 based on sym2, while other existing models yielded RMSE value of 24.90, 20.53, 16.15, 7.058.

Our proposed model yielded maximum PSNR value of 40.45dB based on db2, while others existing fusion models yielded 18.28 dB, 22.66 dB, 23.86 dB, 30.60 dB.

**Conclusions**

This paper devises a new approach named DTCWT based weighted fusion model for Multimodal medical image fusion. Here we have proposed a newly designed framework for enhancing quality of multimodal medical images. We have taken into consideration the two medical images for image fusion procedure. Here focused regions were identified and fused, moreover fine details were also extracted in the fused image which could help in medical diagnosis for medical practitioners. The experimental results demonstrated that the developed

fusion strategy accomplished the better result dependent on MI using sym2 with a maximum value of 1.554, RMSE using sym2 with a least value of 5.554 and PSNR using db2 with a maximum value of 40.45 dB. Currently we carried out our experimental work on brain MRI images taken from BRATS dataset. In future we may also test other typical medical modalities as well as incorporation of different wavelet transform techniques and optimization techniques that will also be an interesting part of research and development.

## References

- 1 Bhateja V, Patel H, Krishn A, Sahu A & Lay-Ekuakille A, Multimodal medical image sensor fusion framework using cascade of wavelet and contourlet transform domains, *IEEE Sens J*, **15(12)** (2015) 6783–6790.
- 2 Li X, He M & Roux M, Multifocus image fusion based on redundant wavelet transform, *IET Image Process*, **4(4)** (2010) 283–293.
- 3 Rodrigues D, Virani H A & Kutty S, Multimodal image fusion techniques for medical images using wavelets, *Image*, **2(3)** (2014) 310–313.
- 4 Srivastava V, Tripathi B K & Pathak V K, Hybrid computation model for intelligent system design by synergism of modified EFC with neural network, *Int J Inf Technol Decis Mak*, **14(1)** (2015) 17–41.
- 5 B Liu Y, Guo Y & Georgiou T & Lew M S, Fusion that matters, convolutional fusion networks for visual recognition, *Multimed Tools Appl*, **77** (2018) 29407–29434, <https://doi.org/10.1007/s11042-018-5691-4>
- 6 Singh V & Kaushik V D, HoEnTOA: Holoentropy and Taylor assisted optimization based a novel image quality enhancement algorithm for multi-focus image fusion, *J Sci Ind Res*, **80(10)** (2021) 875–886, Doi: <http://nopr.niscair.res.in/handle/123456789/58230>.
- 7 Nikolov S, Hill P, Bull D & Canagarajah N, Wavelets for image fusion, in *Wavelets in Signal and Image Analysis, Computational Imaging and Vision*, vol. 19 edited by A A Petrosian & F G Meyer (Springer, Dordrecht) 2001, [https://doi.org/10.1007/978-94-015-9715-9\\_8](https://doi.org/10.1007/978-94-015-9715-9_8)
- 8 Cheng J, Liu H, Liu T, Wang F & Li H, Remote sensing image fusion via wavelet transform and sparse representation, *ISPRS J Photogramm Remote Sens*, **104** (2015) 158–173.
- 9 Rahman S M, Ahmad M O & Swamy M N S, Contrast-based fusion of noisy images using discrete wavelet transform, *IET Image Process*, **4(5)** (2010) 374–384.
- 10 Deng A, Wu J & Yang S, An image fusion algorithm based on discrete wavelet transform and canny operator, in *Advanced Research on Computer Education, Simulation and Modeling*, vol 175 edited by S Lin & Huang X (Springer, Berlin, Heidelberg) 2011, [https://doi.org/10.1007/978-3-642-21783-8\\_6](https://doi.org/10.1007/978-3-642-21783-8_6).
- 11 Vijayarajan R & Muttan S, Discrete wavelet transform based principal component averaging fusion for medical images, *Int J Electron Commun*, **69(6)** (2015) 896–902.
- 12 Venkatrao P H & Damodar S S, HwFusion: Holoentropy and SP-Whale optimisation-based fusion model for magnetic resonance imaging multimodal image fusion, *IET Image Process*, **12(4)** (2018) 572–581.
- 13 Zhao W, Wang D & Lu H, Multi-focus image fusion with a natural enhancement via joint multi-level deeply supervised convolutional neural network, *IEEE Trans Circuits Syst Video Technol*, **29(4)** (2019) 1102–1115.
- 14 Ma B, Zhu Y, Yin X, Ban X, Huang H & Mukeshimana M, Sef-fuse: An unsupervised deep model for multi-focus image fusion, *Neural Comput Appl*, (2020) 1–12.
- 15 Zhang H, Le Z, Shao Z, Xu H & Ma J, MFF-GAN: An unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion, *Inf Fusion*, **66** (2020) 40–53.
- 16 Panigrahy C, Seal A & Mahato N K, Fractal dimension based parameter adaptive dual channel PCNN for multi-focus image fusion, *Opt Lasers Eng*, **133** (2020), <https://doi.org/10.1016/j.optlaseng.2020.106141>.
- 17 Kumar B, Ranjan R K & Husain A, A multi-objective enhanced fruit fly optimization (MO-EFOA) framework for despeckling SAR images using DTCWT based local adaptive thresholding, *Int J Remote Sens*, **42(14)** (2021) 5497–5518.
- 18 TuF, Yin S, Member, Ouyang P, Tang S, Liu L & Wei S, Deep convolutional neural network architecture with reconfigurable computation patterns, *IEEE Trans Very Large Scale Integr VLSI Syst*, **25(8)** (2017) 2220–2233.
- 19 Kavitha S & Thyagarajan K K, Efficient DWT-based fusion techniques using genetic algorithm for optimal parameter estimation, *Soft Comput*, **21(12)** (2017) 3307–3316.
- 20 Zhao W, Wang L & Zhang Z, Atom search optimization and its application to solve a hydrogeologic parameter estimation problem, *Knowl Based Syst*, **163** (2019), 283–304.
- 21 Seyedali M, SCA: A sine cosine algorithm for solving optimization problems, *Knowl Based Syst*, **96** (2016) 120–133.
- 22 Multimodal brain tumor segmentation challenge 2018 (BraTS) dataset taken from <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=37224922> &quot;(23 July 2021).
- 23 Kavitha S & Thyagarajan K K, Efficient DWT-based fusion techniques using genetic algorithm for optimal parameter estimation, *Soft Comput*, **21(12)** (2017) 3307–3316.
- 24 Singh V & Kaushik V D, Renyi entropy and atom search sine cosine algorithm for multi focus image fusion, *Signal Image Video Process*, **15** (2021) 903–12.
- 25 Singh V & Kaushik V D, WeAbDeepCNN: Weighted Average Model and ASSCA based Two Level Fusion Scheme For Multi-Focus Images, *J Sci Ind Res*, **80(10)** (2021) 905–914, Doi: <http://nopr.niscair.res.in/handle/123456789/58227>