ORIGINAL RESEARCH

Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks

Manjit Kaur[1,3] · Dilbag Singh[2,3]

Abstract

The advancements in automated diagnostic tools allow researchers to obtain more and more information from medical images. Recently, to obtain more informative medical images, multi-modality images have been used. These images have significantly more information as compared to traditional medical images. However, the construction of multi-modality images is not an easy task. The proposed approach, initially, decomposes the image into sub-bands using a non-subsampled contourlet transform (NSCT) domain. Thereafter, an extreme version of the Inception (Xception) is used for feature extraction of the source images. The multi-objective differential evolution is used to select the optimal features. Thereafter, the coefficient of determination and the energy loss based fusion functions are used to obtain the fused coefficients. Finally, the fused image is computed by applying the inverse NSCT. Extensive experimental results show that the proposed approach outperforms the competitive multi-modality image fusion approaches.

Keywords Fusion · Diagnosis · CNN · Multi-modality · Differential evolution

1 Introduction

Biomedical images are extensively utilized for automated diagnosis of various diseases such as COVID-19, pneumonia, tuberculosis, cancer, etc. Biomedical imaging systems play an efficient role to monitor and diagnose the internal body organs without utilizing any kind of surgery. Biomedical images come up with various modalities to understand the internal body organs (Du et al. 2016). These modalities are positron emission tomography (PET), magnetic resonance imaging (MRI), computerized tomography (CT), X-ray, ultrasound, etc. (James and Dasarathy 2014). Every modality comes up with its own significance during its usage in the diagnosis process. Due to a single modality, these images are limited to certain diseases or issues. Therefore, multi-modality biomedical images are desirable. A multi-modality image can be obtained by using the efficient fusion approach (Daniel et al. 2017). Thus, these images have more information as compared to classical images and also more helpful to diagnose various kind of diseases (Du et al. 2016).

Recently, many researchers have designed and implemented fusion approaches to obtain efficient multi-modality biomedical images (Ravi and Krishnan 2018; Hu et al. 2020). However, many researchers have utilized the existing image fusion approaches. Therefore, the obtained multi-modality fused images may suffer from various issues such as gradient and texture distortion, especially for the infected region. To overcome the issues associated with the existing approaches, many researchers have utilized deep learning and dictionary learning kind of approaches. These are dictionary learning (Hu et al. 2020), local-features fuzzy sets (Ullah et al. 2020), deep learning (Algarni 2020; Xia et al. 2019; Zhou et al. 2019), deep learning and NSST (Wang et al. 2019), etc. that found to be the best tool to obtain efficient multi-modality fused images.

The design and development of an efficient multi-modality biomedical image fusion approach is still an open area for research. Deep learning is found to be one of the best fusion approaches to obtain promising results. Additionally,

✉ Dilbag Singh
  dilbag@ieee.org

1   Department of Computer and Communication Engineering, Manipal University Jaipur, Jaipur, India

2   Department of Computer Science and Engineering, Manipal University Jaipur, Jaipur, India

3   Computer Science Engineering, School of Engineering and Applied Sciences, Bennett University, Greater Noida 201310, India

a deep transfer learning-based multi-modality biomedical fusion approach can provide better results.

The main contributions are summarized as:

- A multi-objective differential evolution and Xception model based multi-modality biomedical fusion model is proposed.
- The proposed approach, initially, decomposes the image into subbands using a non-subsampled contourlet transform (NSCT).
- An extreme version of the Inception (Xception) is then used for feature extraction of the source images.
- The multi-objective differential evolution is used to select the optimal features.
- To obtain the fused coefficients, a coefficient of determination and the energy loss based fusion functions are used.
- Finally, a fused image is computed by applying the inverse NSCT.
- The proposed and the competitive approaches are compared by considering the benchmark multi-modality image fusion dataset.

The remaining structure of this paper is organized as: Existing literature is presented in Sect. 2. The proposed medical image fusion is illustrated in Sect. 3. Experimental results and comparative analyses are discussed in Sect. 4. Section 5 concludes the proposed work.

2 Literature review

Shu et al. (Zhu et al. 2019) implemented local laplacian energy and phase congruency based fusion approach in the NSCT domain (LEPN). Local laplacian energy utilized weighted local energy and sum of laplacian coefficients to obtain the regulated details and features of input images. Zhu et al. (2020) designed a diffusion-based approach by using the synchronized-anisotropic operators (DSA). A maximum absolute value constraint was also utilized for base layers fusion. The fusion decision map was computed by considering the sum of the modified anisotropic Laplacian approach by using the similar corrosion sub-bands obtained from the anisotropic diffusion. Kumar et al. (2020) proposed a co-learning based fusion maps for obtained more efficient multi-modality fused biomedical images. A convolutional neural network (CNN) was also used for the prediction and segmentation of potential objects.

Lu et al. (2014) designed an edge-guided dual-modality (EGDM) approach to obtain the multi-modality images. It performs significantly better even on highly under-sampled data. Lifeng et al. (2001) utilized wavelet and quality analysis to obtain the biomedical multi-modality fused

images. The pyramid wavelet used for the fusion process. Ma et al. (2020) designed a dual-discriminator conditional generative adversarial network (DDcGAN) to obtain multi-modality fused images. It obtained a real-like fused image by using the content loss to dupe both discriminators. Two discriminators were also considered intention to differentiate the composition variations between the fused and source images, respectively.

Wang et al. (2019) developed a 3D auto-context-based locality adaptive multi-modality GANs (3GANs) to obtain more efficient multi-modality fused images. A non-unified kernel was also used along with the adaptive approach for multimodality fusion. Gai et al. (2019) utilized pulse coupled neural network (PCNN) by considering the edge preservation and enhanced sparse representation in the non-subsampled shearlet transform (NSST). It completely utilized the features of various modalities that can handle edge details well and improves the results.

Liu et al. (2020) utilized a VGG16 based deep transfer learning approaches for image fusion to improve the classification process. VGG16 can obtain more efficient features as compared to the traditional deep learning approaches. The obtained features were then used to fuse the multi-modality biomedical images. Tavard et al. (2014) designed a multi-modal registration and fusion approach to improve the cardiac re-synchronization. The approach helps in improving therapy optimization.

Zhu et al. (2016) designed a novel dictionary learning-based image fusion approach for multi-modality biomedical images. Due to the use of dictionary learning, this approach achieves higher accuracy. But it is computationally complex in nature. Liu et al. (2018) proposed the biomedical image decomposition approach using NSST to fuse the multi-modality images. This approach has shown significant results over the existing approaches, but, suffer from the edge degradation issue. Wang et al. (2020) utilized CNN and contrast pyramid to fuse the biomedical multi-modality images. CNN can fuse the images efficiently. However, it is a computationally extensive approach, and some time also not provide promising results if the biomedical images are very similar to each other.

From the literature review, it has been observed that multi-modality image fusion is still an open area for research. The deep learning is found to be the best promising techniques to obtain better multi-modality fused biomedical images. But these approaches can provide better results if pre-trained deep transfer learning approaches are used. Additionally, the initial parameter selection of the deep learning and deep transfer learning approach is also a challenging issue (Pannu et al. 2019; Kaur and Singh 2019; Pannu et al. 2018). Therefore, in this paper, a well-known multi-objective differential evolution is used to enhance the results.

3 Proposed multi-objective differential evolution based deep transfer learning model for multi-modality image fusion

Assume that I_1 and I_2 show source images. Initially, NSCT is used to decompose I_1 and I_2 into sub-bands. Our primary objective is to fuse the respective sub-bands of both source images. The fusion of high sub-band is achieved by using the extreme version of the Inception (Xception) model. The coefficient of determination is utilized to evaluate the significance of the computed fused high sub-bands. Low sub-bands fusion is achieved by using the local energy function. Inverse NSCT is utilized to compute the multi-modality fused image. Figure 1 shows the step by step methodology of the proposed model.

3.1 Nonsubsampled contourlet transform

Nonsubsampled Contourlet Transform (NSCT) is a well-known transform used to decompose the images into the wavelet domain. It is a shift-invariant transform which can provide rich directional details. This directionality is effective to convert the transformed images to the actual one with minimum root mean square error (for more details please see Da Cunha et al. 2006).

3.2 Feature extraction using deep Xception model

CNN may suffer from the under-fitting issue, as many potential features may not be extracted. To overcome this issue, an extreme version of the Inception (Xception) model is used. Figure 2 represents the block diagram of the xception model (for mathematical and other information please see Chollet 2017).

Both high sub-bands of source images are placed in parallel fashion in the Xception model. Consider $\eta I_1(p,q)$ and $\eta I_2(p,q)$ are the obtained features from respective high sub-bands by using the Xception model.

3.3 Feature selection using multi-objective differential evolution

In this step, the optimal features are selected from the features obtained from the Xception model. The fusion factor and entropy metrics are used as the fitness function to select the optimal features. A multi-objective differential evolution can solve many computationally complex problems (Babu et al. 2005). It can significantly balance the fast convergence and population diversity. It can be described in the following steps:

I. Initialization: First of all, various parameters related to differential evolution are defined such as population size (t_p), crossover rate (c_r), mutation rate (m_r), etc. Random distribution is used to generate the random solutions $\beta_\alpha^0 (\alpha = 1, 2, \ldots, t_p)$. h defines the number of function evaluations. It is used to control the iterative process of differential optimization with maximum function evaluations (h_M).

II. Iterative step Mutation and crossover operators are used to obtain the optimal number of features.

Mutation is implemented on a β_α^h to evaluate a child vector Π_α^h. In this paper, following mutation is used:

$$\Pi_\alpha^h = \beta_{d1}^h + m_r \cdot (\beta_{d2}^h - \beta_{d3}^h) \qquad (1)$$

Here, α shows index values. $d_i \neq \alpha \, \forall i = 1 : 3$. d_1, d_2, and d_3 are random numbers selected from $[1, \, t_p,]$.

Crossover is used to obtain the news solutions. A child ϵ_α^h can be obtained from $\forall \, \beta_\alpha^h$, as:

$$\epsilon_{\alpha_\kappa}^h = \begin{cases} \Pi_{\alpha_\kappa}^h, & \beta_\kappa \leq c_r \quad \text{or} \quad \kappa = \kappa_{d_n} \\ \beta_{\alpha_\kappa}^h, & otherwise \end{cases} \kappa = 1, 2, \ldots, D, \quad (2)$$

where D shows dimensions of the problem. $\beta_\kappa \in [0, \, 1]$ and $\kappa_{d_n} \in [1, \, D]$.

III. Selection: A child vector ϵ_α^h can be drawn by considering its parent vector β_α^h as:

$$\beta_\alpha^{h+1} = \begin{cases} \epsilon_\alpha^h, & f(\epsilon_\alpha^h) \leq f(\beta_\alpha^h) \\ \beta_\alpha^h, & otherwise \end{cases} \qquad (3)$$

IV. Stopping condition: If number of functional evaluations are lesser than the total available evaluation then Steps II and III will be repeated.

3.4 Fusion of high sub-bands

The extracted and selected features using the Xception model from high sub-bands are then fused by using the coefficient of determination (R). R between $\eta I_1(p,q)$ and $\eta I_2(p,q)$ can be computed as:

$$R_N(\eta I_1, \eta I_2)$$
$$= \frac{\left(\sum_{p=1}^m \sum_{q=1}^n (\eta I_1(p,q) - \overline{\eta I_1})(\eta I_2(p,q) - \overline{\eta I_2}) \right)^2}{\sum_{p=1}^m \sum_{p=1}^n (\eta I_1(p,q) - \overline{\eta I_1})^2 \times \sqrt{\sum_{p=1}^m \sum_{p=1}^n (\eta I_2(p,q) - \overline{\eta I_2})^2}}$$
$$(4)$$

Here, $\overline{\eta I_1}$ and $\overline{\eta I_2}$ shows the average of high sub-bands, respectively.

The dominated features are preserved in the obtained feature maps as:

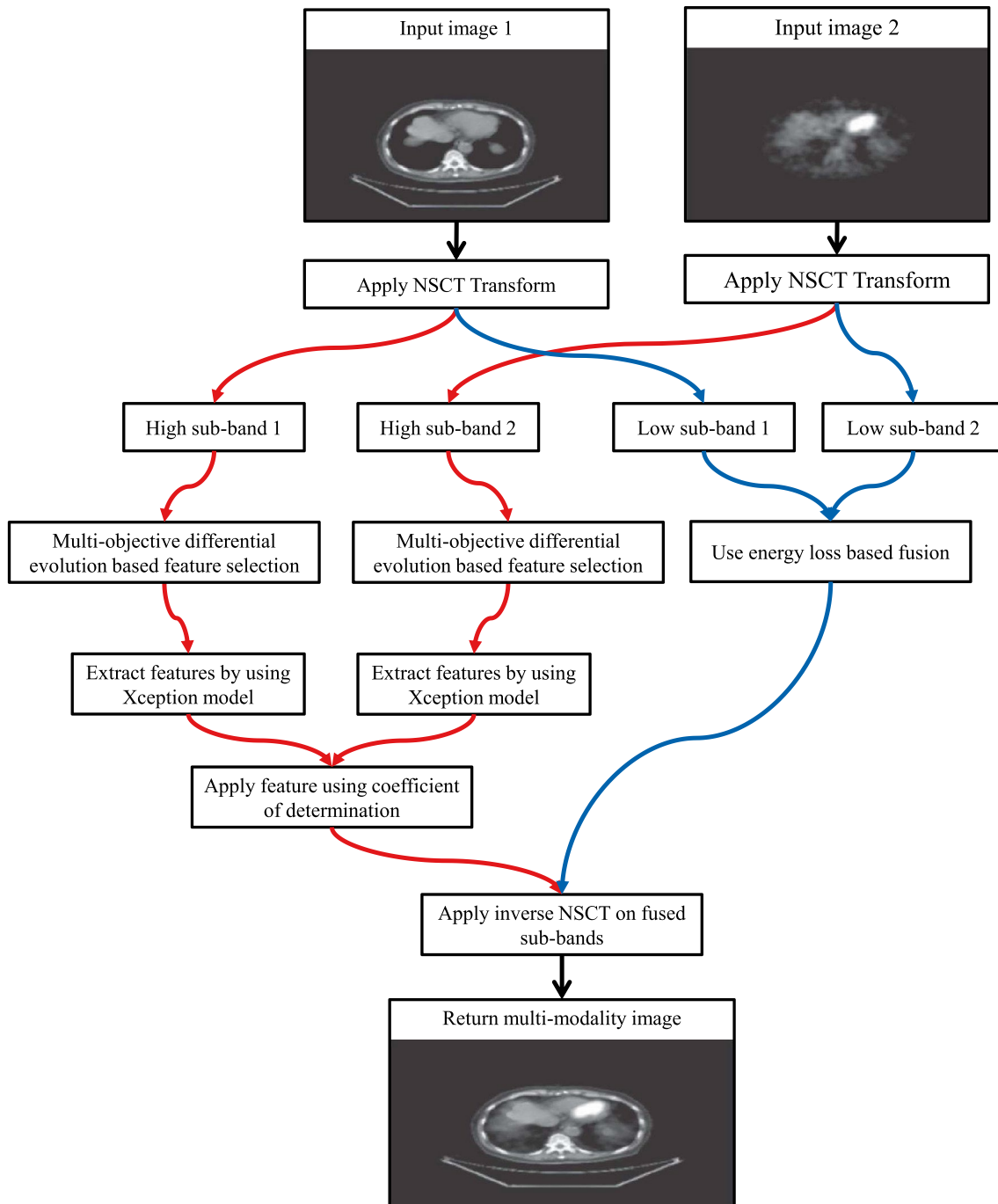$$F_s(p,q) = max(sI_1 \times R_N + sI_2 \times (1 - R_N)) \qquad (5)$$

Fig. 1 Diagrammatic flow of the proposed multi-objective differential evolution based deep transfer learning model for multi-modality image fusion

Here, sI_1 and sI_2 show high sub-bands of I_1 and I_2, respectively.

3.5 Fusion of low sub-bands

Motivated from (Hermessi et al. 2018), local energy is used to fuse the low sub-bands as:

$$\chi_I(p, q) = \sum_{p' \in \gamma} \sum_{q' \in \delta} |I(p + p', q + q')| \tag{6}$$

Here, $I = I_1$ or I_2. $\gamma \times \delta$ represents neighbors of patch placed at (p, q). size of local patch is assigned as 5×5. Fused coefficients of low sub-bands can be computed as:
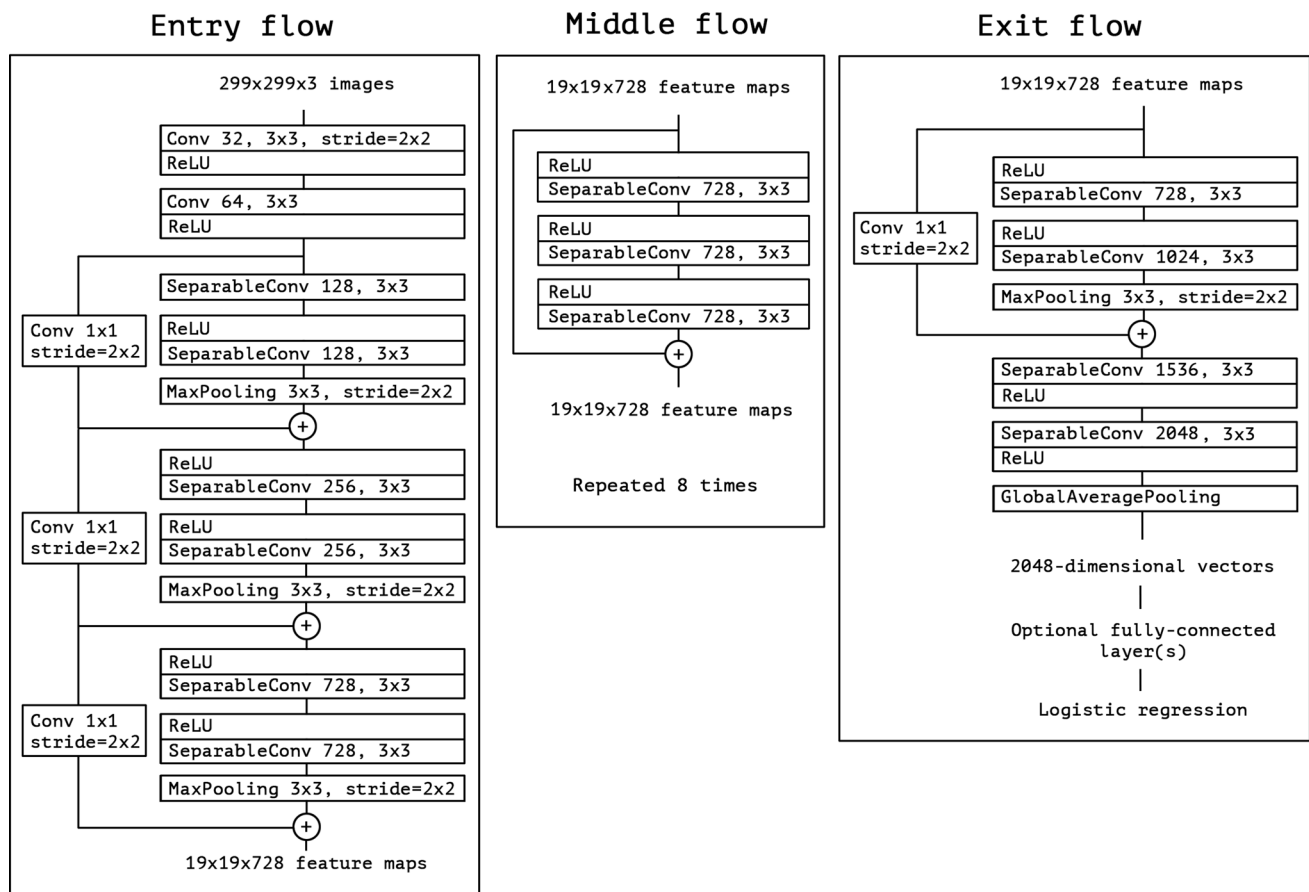
Fig. 2 Architecture of Xception model (obtained from Chollet 2017)

$$\psi_f(p,q) = \begin{cases} \psi I_1(p,q) & |\chi I_1(p,q)| \geq |\chi I_2(p,q)| \\ \psi I_2(p,q) & |\chi I_1(p,q)| < |\chi I_2(p,q)| \end{cases} \qquad (7)$$

4 Experimental results

To evaluate the performance of the proposed approach, benchmark multi-modality biomedical images dataset is obtained from Ullah et al. (2020). Fifteen different pairs of modality images are taken for comparative purposes. The main goal is to fuse these images to obtain multi-modality fused images. To draw the comparisons, six competitive multi-modality biomedical fusion approaches such as LEPN (Zhu et al. 2019), DSA (Zhu et al. 2020), CNN (Kumar et al. 2020), EGDM (Lu et al. 2014), DDcGAN (Ma et al. 2020), and 3GANs (Wang et al. 2019) are also implemented on the same set of images. The hyper-parameters of these approaches are assigned as mentioned in their respective papers.

4.1 Visual analysis

Figures 3 and 4 represent the source images and their respective multi-modality fused biomedical images obtained from the LEPN (Zhu et al. 2019), DSA (Zhu et al. 2020), CNN (Kumar et al. 2020), EGDM (Lu et al. 2014), DDcGAN (Ma et al. 2020), 3GANs (Wang et al. 2019) and the proposed approach. It is clearly shown that the obtained results have better modality as compared to the competitive approaches. Although, the existing approaches such as LEPN (Zhu et al. 2019), DSA (Zhu et al. 2020), CNN (Kumar et al. 2020), EGDM (Lu et al. 2014), DDcGAN (Ma et al. 2020), and 3GANs (Wang et al. 2019) provide significant visual results but have little edge and texture distortion. Figures 3i and 4j show the obtained results from the proposed approach. These images prove that the proposed approach provides a better visual appearance of the obtained multi-modality fused images.

Fig. 3 Analysis of multi-modality biomedical fusion approaches: **a** MRI, **b** CT, **c** LEPN (Zhu et al. 2019), **d** DSA (Zhu et al. 2020), **e** CNN (Kumar et al. 2020), **f** EGDM (Lu et al. 2014), **g** DDcGAN (Ma et al. 2020), and **h** 3GANs (Wang et al. 2019) and **i** proposed approach
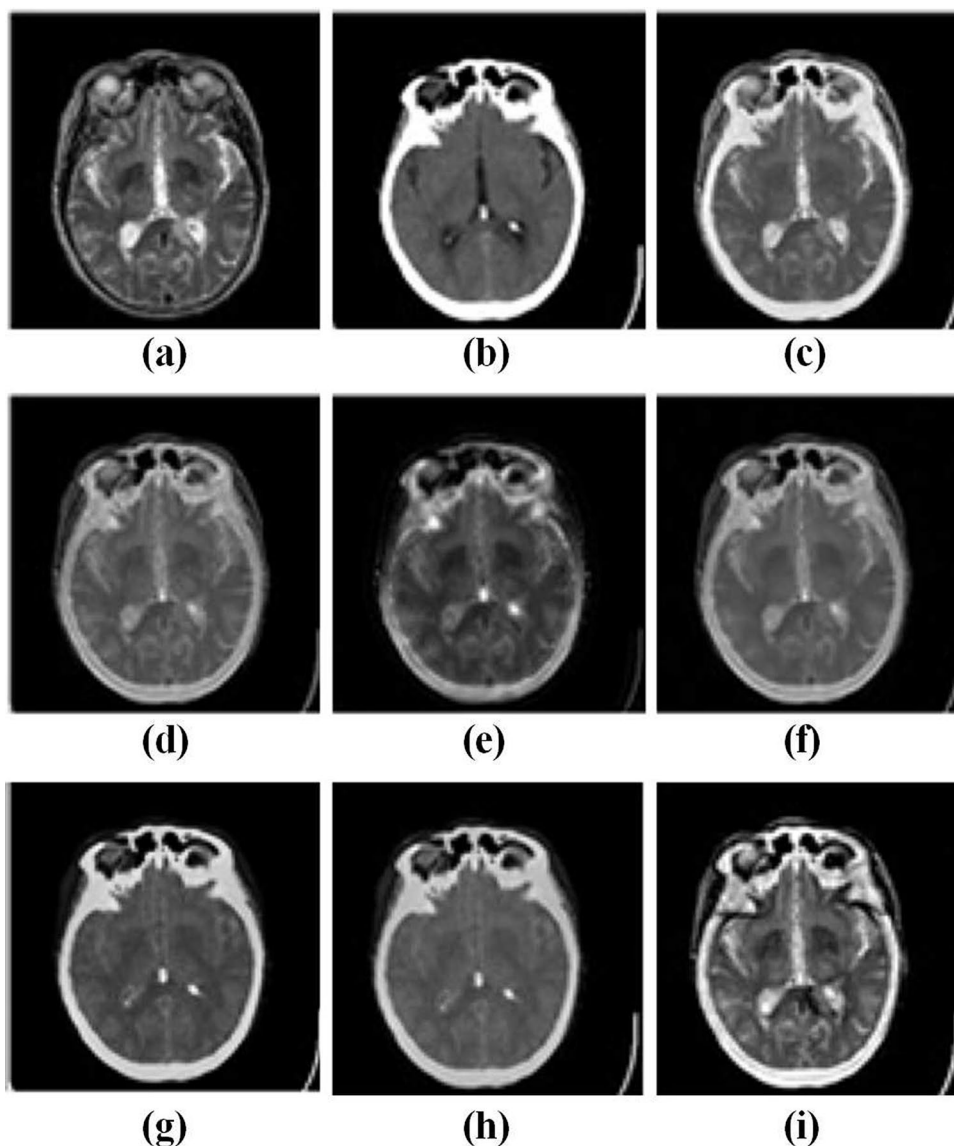
4.2 Quantitative analysis

In this section, we have compared the proposed approach with the existing approaches such as LEPN (Zhu et al. 2019), DSA (Zhu et al. 2020), CNN (Kumar et al. 2020), EGDM (Lu et al. 2014), DDcGAN (Ma et al. 2020), and 3GANs (Wang et al. 2019), by considering the some well-known performance metrics. The selected performance measures are as edge strength, fusion symmetry, entropy, and fusion factor (for mathematical information see Prakash et al. 2019).

A multi-modality biomedical image fusion approach generally provides significant entropy values. Table 1 depicts entropy analysis of the proposed deep transfer learning-based multi-modality biomedical image fusion approach. It shows that the proposed approach provides significantly more entropy values than the existing multi-modality

biomedical image fusion approaches. It is found that the proposed approach provides 1.8343% improvement over the best available approaches.

Mutual information represents the preserved details from the sourced image in the fused image. Therefore, it is desirable to be maximum. Table 2 depicts the mutual information analysis of the proposed approach over the competitive approaches. The proposed approach shows an average improvement of 1.8373%.

The fusion factor is a well-known performance metric that shows the strength of the fusion process. It is desirable to be maximum. Table 3 shows the fusion factor analysis of the proposed and competitive approaches. The proposed approach shows an average improvement of 1.3928% over the competitive fusion models.

Fusion symmetry evaluates the symmetric details between source and fused images. It is desirable to be maximum.

Fig. 4 Analysis of multi-modality biomedical fusion approaches: **a** MRI, **b** CT, **c** LEPN (Zhu et al. 2019), **d** DSA (Zhu et al. 2020), **e** CNN (Kumar et al. 2020), **f** EGDM (Lu et al. 2014), **g** DDcGAN (Ma et al. 2020), and **h** 3GANs (Wang et al. 2019) and **i** proposed approach
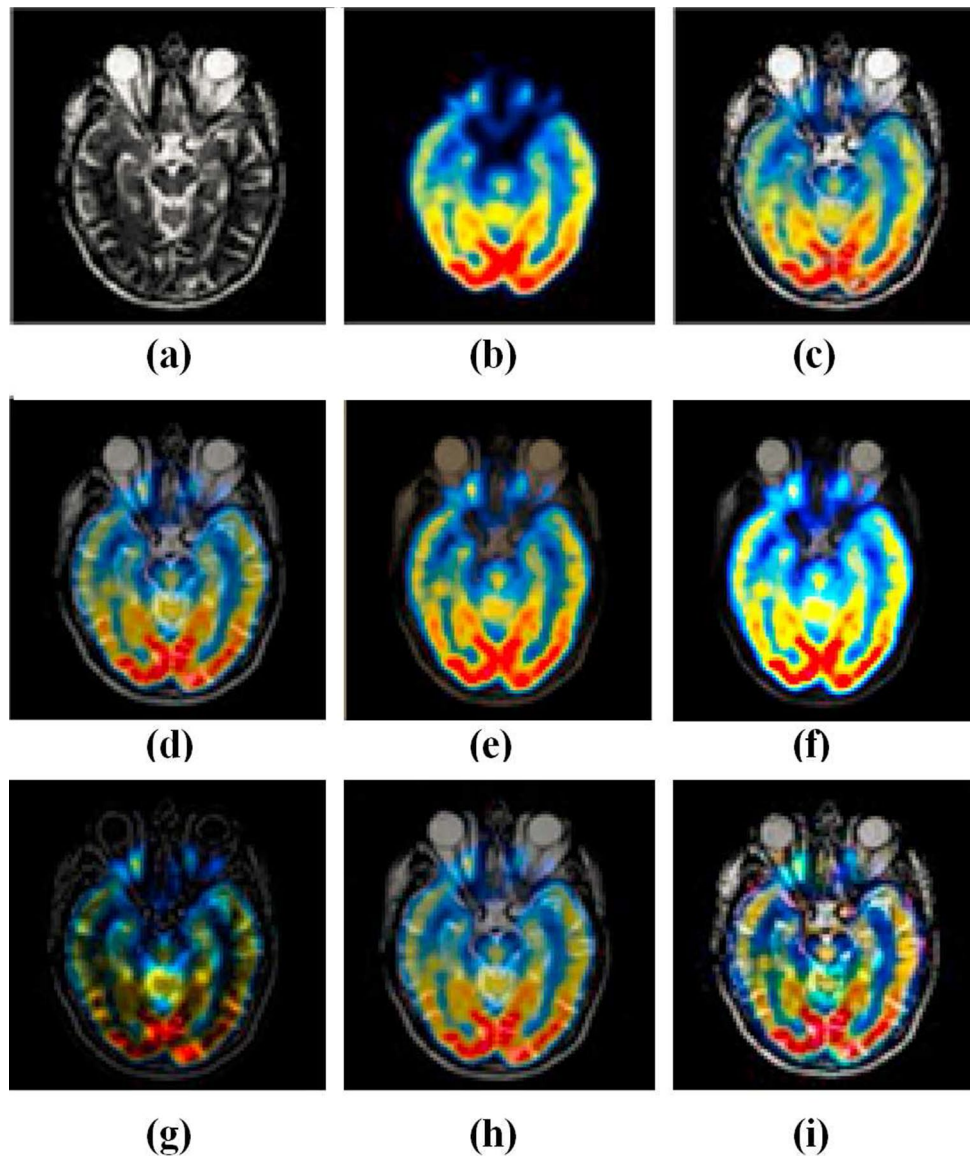
Table 4 shows fusion symmetry analysis of the proposed deep transfer learning model based multi-modality fusion model. It is found that the proposed model achieves an average improvement of 1.1974% over the competitive models.

Edge strength evaluates the degree of edge preservation and it is desirable to be maximum (Xydeas and Petrovic 2000). Table 5 shows the edge strength analysis of the proposed deep transfer learning-based multi-modality image fusion model. The proposed model achieves an average improvement of 1.6928% over the competitive approaches.

5 Conclusion

From the literature review, it has been found that multi-modality image fusion is still an open area for research. The deep learning-based fusion approaches are found to be one of the best promising techniques to obtain better multi-modality fused biomedical images. However, these approaches are computationally complex in nature and also still suffer from the under-fitting issue. The proposed approach, initially, decomposes the image into sub-bands

Table 1 Comparative analysis among the proposed deep transfer learning based multi-modality image fusion and the competitive approaches in terms of entropy (maximum is better)

Group	LEPN (Zhu et al. 2019)	DSA (Zhu et al. 2020)	CNN (Kumar et al. 2020)	EGDM (Lu et al. 2014)	DDcGAN (Ma et al. 2020)	3GANs (Wang et al. 2019)	Proposed
G_1	6.22 ± 0.75	6.33 ± 0.69	6.25 ± 0.67	6.56 ± 0.55	6.43 ± 0.68	6.73 ± 0.09	6.77 ± 0.35
G_2	6.18 ± 0.64	6.04 ± 0.62	6.18 ± 0.54	6.23 ± 0.45	6.73 ± 0.58	6.33 ± 0.08	6.77 ± 0.43
G_3	6.43 ± 0.75	6.35 ± 0.55	5.95 ± 0.45	6.48 ± 0.54	6.68 ± 0.55	6.42 ± 0.45	6.71 ± 0.41
G_4	6.09 ± 0.55	5.93 ± 0.55	6.29 ± 0.55	6.46 ± 0.45	6.47 ± 0.58	6.62 ± 0.57	6.69 ± 0.45
G_5	6.38 ± 0.72	6.44 ± 0.75	6.64 ± 0.64	6.26 ± 0.56	6.66 ± 0.47	6.57 ± 0.45	6.79 ± 0.39
G_6	6.28 ± 0.73	5.94 ± 0.65	6.39 ± 0.45	6.31 ± 0.44	6.53 ± 0.55	6.66 ± 0.35	6.73 ± 0.45
G_7	6.17 ± 0.66	5.96 ± 0.75	6.26 ± 0.55	6.58 ± 0.45	6.69 ± 0.68	6.65 ± 0.35	6.72 ± 0.35
G_8	5.91 ± 0.55	5.96 ± 0.55	6.06 ± 0.49	6.46 ± 0.46	6.28 ± 0.48	6.57 ± 0.35	6.81 ± 0.43
G_9	5.96 ± 0.76	6.19 ± 0.65	6.33 ± 0.47	6.74 ± 0.62	6.26 ± 0.52	6.75 ± 0.48	6.71 ± 0.41
G_{10}	5.93 ± 0.75	5.97 ± 0.53	6.38 ± 0.45	6.36 ± 0.54	6.79 ± 0.68	6.22 ± 0.53	6.66 ± 0.45
G_{11}	6.05 ± 0.66	6.33 ± 0.75	6.16 ± 0.47	6.77 ± 0.69	6.55 ± 0.43	6.77 ± 0.54	6.69 ± 0.39
G_{12}	6.18 ± 0.75	5.98 ± 0.55	6.82 ± 0.55	6.31 ± 0.64	6.44 ± 0.47	6.46 ± 0.57	6.75 ± 0.43
G_{13}	6.12 ± 0.65	6.32 ± 0.75	6.19 ± 0.66	6.39 ± 0.45	6.37 ± 0.55	6.72 ± 0.45	6.65 ± 0.41
G_{14}	6.31 ± 0.55	5.94 ± 0.65	6.14 ± 0.45	6.39 ± 0.67	6.77 ± 0.62	6.23 ± 0.45	6.75 ± 0.45
G_{15}	5.96 ± 0.75	6.38 ± 0.75	5.93 ± 0.53	6.64 ± 0.64	6.44 ± 0.69	6.25 ± 0.57	6.78 ± 0.39

Table 2 Comparative analysis among the proposed deep transfer learning based multi-modality image fusion and the competitive approaches in terms of mutual information (maximum is better)

Group	LEPN (Zhu et al. 2019)	DSA (Zhu et al. 2020)	CNN (Kumar et al. 2020)	EGDM (Lu et al. 2014)	DDcGAN (Ma et al. 2020)	3GANs (Wang et al. 2019)	Proposed
G_1	0.63 ± 0.013	0.62 ± 0.011	0.61 ± 0.012	0.59 ± 0.012	0.61 ± 0.008	0.62 ± 0.009	0.63 ± 0.009
G_2	0.61 ± 0.014	0.61 ± 0.012	0.62 ± 0.012	0.61 ± 0.009	0.61 ± 0.011	0.64 ± 0.009	0.61 ± 0.009
G_3	0.63 ± 0.014	0.62 ± 0.011	0.62 ± 0.008	0.63 ± 0.008	0.59 ± 0.011	0.63 ± 0.008	0.63 ± 0.007
G_4	0.61 ± 0.013	0.62 ± 0.012	0.63 ± 0.012	0.61 ± 0.012	0.63 ± 0.012	0.62 ± 0.008	0.63 ± 0.007
G_5	0.61 ± 0.011	0.63 ± 0.012	0.63 ± 0.011	0.63 ± 0.012	0.64 ± 0.011	0.64 ± 0.012	0.63 ± 0.008
G_6	0.64 ± 0.014	0.61 ± 0.013	0.63 ± 0.012	0.59 ± 0.012	0.64 ± 0.011	0.61 ± 0.008	0.64 ± 0.009
G_7	0.63 ± 0.014	0.59 ± 0.012	0.59 ± 0.012	0.63 ± 0.009	0.64 ± 0.009	0.59 ± 0.007	0.64 ± 0.009
G_8	0.64 ± 0.014	0.62 ± 0.013	0.59 ± 0.012	0.61 ± 0.008	0.61 ± 0.009	0.61 ± 0.012	0.64 ± 0.007
G_9	0.62 ± 0.011	0.59 ± 0.011	0.59 ± 0.009	0.63 ± 0.009	0.62 ± 0.011	0.59 ± 0.008	0.63 ± 0.009
G_{10}	0.63 ± 0.015	0.63 ± 0.015	0.63 ± 0.008	0.63 ± 0.009	0.59 ± 0.011	0.61 ± 0.007	0.63 ± 0.009
G_{11}	0.62 ± 0.012	0.62 ± 0.013	0.59 ± 0.011	0.62 ± 0.012	0.62 ± 0.008	0.62 ± 0.009	0.62 ± 0.007
G_{12}	0.59 ± 0.011	0.59 ± 0.013	0.61 ± 0.008	0.62 ± 0.012	0.62 ± 0.013	0.62 ± 0.009	0.62 ± 0.007
G_{13}	0.61 ± 0.012	0.63 ± 0.012	0.62 ± 0.012	0.62 ± 0.012	0.63 ± 0.008	0.63 ± 0.008	0.63 ± 0.008
G_{14}	0.61 ± 0.013	0.63 ± 0.011	0.61 ± 0.008	0.61 ± 0.011	0.62 ± 0.011	0.64 ± 0.012	0.63 ± 0.007
G_{15}	0.59 ± 0.015	0.64 ± 0.011	0.63 ± 0.008	0.63 ± 0.011	0.64 ± 0.011	0.59 ± 0.008	0.64 ± 0.008

using a non-subsampled contourlet transform (NSCT). Thereafter, an extreme version of the Inception (Xception) has been used for feature extraction of the source images. The multi-objective differential evolution has been used to select the optimal features. Thereafter, to obtain the fused coefficients, a coefficient of determination and the energy loss functions are used. Finally, the fused image has been computed by applying the inverse NSCT. Extensive experimental results have shown that the proposed

approach outperforms the competitive multi-modality image fusion approaches in terms of various performance metrics. In near future, one may use the proposed model for other applications such as remote sensing images (Singh et al. 2018; Singh and Kumar 2019a), medical images, etc. Additionally, the proposed hyper-parameters tuning approach can be used to tune the hyper parameter of the other approaches such as visibility restoration models

Table 3 Comparative analysis among the proposed deep transfer learning based multi-modality image fusion and the competitive approaches in terms of fusion factor (maximum is better)

Group	LEPN (Zhu et al. 2019)	DSA (Zhu et al. 2020)	CNN (Kumar et al. 2020)	EGDM (Lu et al. 2014)	DDcGAN (Ma et al. 2020)	3GANs (Wang et al. 2019)	Proposed
G_1	1.26 ± 0.026	1.22 ± 0.022	1.22 ± 0.022	1.22 ± 0.031	1.24 ± 0.022	1.26 ± 0.026	1.26 ± 0.022
G_2	1.22 ± 0.022	1.26 ± 0.028	1.18 ± 0.032	1.22 ± 0.029	1.22 ± 0.026	1.28 ± 0.025	1.29 ± 0.022
G_3	1.26 ± 0.022	1.24 ± 0.025	1.21 ± 0.033	1.24 ± 0.022	1.26 ± 0.022	1.24 ± 0.028	1.26 ± 0.022
G_4	1.23 ± 0.021	1.22 ± 0.024	1.22 ± 0.022	1.22 ± 0.025	1.28 ± 0.031	1.26 ± 0.022	1.28 ± 0.021
G_5	1.22 ± 0.03	1.24 ± 0.022	1.21 ± 0.026	1.28 ± 0.027	1.24 ± 0.024	1.24 ± 0.021	1.28 ± 0.021
G_6	1.26 ± 0.024	1.22 ± 0.035	1.28 ± 0.028	1.22 ± 0.021	1.28 ± 0.028	1.28 ± 0.028	1.28 ± 0.021
G_7	1.22 ± 0.028	1.24 ± 0.025	1.21 ± 0.026	1.22 ± 0.022	1.24 ± 0.031	1.24 ± 0.022	1.24 ± 0.022
G_8	1.24 ± 0.021	1.22 ± 0.028	1.22 ± 0.022	1.26 ± 0.021	1.26 ± 0.026	1.24 ± 0.021	1.26 ± 0.021
G_9	1.24 ± 0.031	1.26 ± 0.024	1.21 ± 0.024	1.24 ± 0.022	1.18 ± 0.022	1.27 ± 0.022	1.26 ± 0.022
G_{10}	1.24 ± 0.022	1.28 ± 0.028	1.24 ± 0.023	1.21 ± 0.024	1.24 ± 0.024	1.27 ± 0.026	1.28 ± 0.022
G_{11}	1.22 ± 0.026	1.28 ± 0.028	1.26 ± 0.026	1.21 ± 0.032	1.26 ± 0.022	1.24 ± 0.024	1.28 ± 0.022
G_{12}	1.24 ± 0.031	1.26 ± 0.028	1.26 ± 0.022	1.26 ± 0.028	1.24 ± 0.032	1.26 ± 0.021	1.26 ± 0.021
G_{13}	1.28 ± 0.021	1.28 ± 0.026	1.24 ± 0.022	1.24 ± 0.026	1.18 ± 0.024	1.27 ± 0.024	1.28 ± 0.021
G_{14}	1.22 ± 0.022	1.24 ± 0.021	1.26 ± 0.023	1.22 ± 0.022	1.24 ± 0.024	1.28 ± 0.022	1.28 ± 0.021
G_{15}	1.24 ± 0.028	1.22 ± 0.022	1.24 ± 0.023	1.24 ± 0.026	1.18 ± 0.024	1.22 ± 0.026	1.24 ± 0.022

Table 4 Comparative analysis among the proposed deep transfer learning based multi-modality image fusion and the competitive approaches in terms of fusion symmetry (maximum is better)

Group	LEPN (Zhu et al. 2019)	DSA (Zhu et al. 2020)	CNN (Kumar et al. 2020)	EGDM (Lu et al. 2014)	DDcGAN (Ma et al. 2020)	3GANs (Wang et al. 2019)	Proposed
G_1	0.31 ± 0.014	0.31 ± 0.014	0.31 ± 0.012	0.31 ± 0.012	0.32 ± 0.008	0.32 ± 0.012	0.33 ± 0.008
G_2	0.31 ± 0.012	0.32 ± 0.013	0.32 ± 0.012	0.32 ± 0.013	0.31 ± 0.014	0.32 ± 0.012	0.33 ± 0.007
G_3	0.29 ± 0.012	0.31 ± 0.012	0.31 ± 0.012	0.31 ± 0.014	0.32 ± 0.015	0.33 ± 0.008	0.34 ± 0.008
G_4	0.32 ± 0.013	0.31 ± 0.012	0.29 ± 0.009	0.31 ± 0.011	0.31 ± 0.008	0.31 ± 0.007	0.33 ± 0.007
G_5	0.31 ± 0.015	0.32 ± 0.012	0.31 ± 0.011	0.31 ± 0.011	0.32 ± 0.008	0.31 ± 0.007	0.33 ± 0.007
G_6	0.31 ± 0.013	0.31 ± 0.013	0.31 ± 0.013	0.33 ± 0.008	0.31 ± 0.008	0.31 ± 0.009	0.34 ± 0.008
G_7	0.29 ± 0.015	0.32 ± 0.013	0.32 ± 0.012	0.32 ± 0.014	0.31 ± 0.011	0.32 ± 0.007	0.33 ± 0.007
G_8	0.31 ± 0.011	0.31 ± 0.014	0.32 ± 0.009	0.31 ± 0.012	0.31 ± 0.011	0.33 ± 0.007	0.34 ± 0.007
G_9	0.29 ± 0.012	0.32 ± 0.012	0.31 ± 0.009	0.31 ± 0.013	0.32 ± 0.008	0.32 ± 0.009	0.33 ± 0.008
G_{10}	0.31 ± 0.014	0.31 ± 0.015	0.31 ± 0.008	0.32 ± 0.014	0.31 ± 0.011	0.32 ± 0.008	0.33 ± 0.008
G_{11}	0.28 ± 0.015	0.31 ± 0.012	0.31 ± 0.014	0.32 ± 0.015	0.32 ± 0.008	0.31 ± 0.012	0.33 ± 0.008
G_{12}	0.29 ± 0.011	0.32 ± 0.015	0.32 ± 0.012	0.32 ± 0.012	0.32 ± 0.013	0.31 ± 0.008	0.33 ± 0.008
G_{13}	0.31 ± 0.013	0.32 ± 0.015	0.31 ± 0.009	0.31 ± 0.012	0.33 ± 0.009	0.33 ± 0.008	0.34 ± 0.008
G_{14}	0.32 ± 0.012	0.31 ± 0.012	0.32 ± 0.012	0.32 ± 0.011	0.32 ± 0.011	0.31 ± 0.007	0.33 ± 0.007
G_{15}	0.31 ± 0.014	0.31 ± 0.013	0.32 ± 0.012	0.31 ± 0.013	0.31 ± 0.008	0.32 ± 0.012	0.33 ± 0.008

(Osterland and Weber 2019; Singh and Kumar 2018, 2019b; Wang et al. 2019; Singh et al. 2019a, 2019b), filtering models (Gupta et al. 2019; Kaur et al. 2020; Wiens 2019), deep learning models (Jaiswal et al. 2020; Basavegowda and Dagnew 2020; Kaur et al. 2019, 2020; Ghosh et al. 2020), etc.

Table 5 Comparative analysis among the proposed deep transfer learning based multi-modality image fusion and the competitive approaches in terms of edge strength (maximum is better)

Group	LEPN (Zhu et al. 2019)	DSA (Zhu et al. 2020)	CNN (Kumar et al. 2020)	EGDM (Lu et al. 2014)	DDcGAN (Ma et al. 2020)	3GANs (Wang et al. 2019)	Proposed
G_1	0.61 ± 0.012	0.62 ± 0.014	0.64 ± 0.012	0.63 ± 0.013	0.59 ± 0.012	0.63 ± 0.012	0.64 ± 0.008
G_2	0.62 ± 0.014	0.62 ± 0.015	0.62 ± 0.012	0.62 ± 0.012	0.62 ± 0.011	0.61 ± 0.009	0.62 ± 0.009
G_3	0.61 ± 0.012	0.64 ± 0.015	0.62 ± 0.008	0.63 ± 0.009	0.62 ± 0.012	0.62 ± 0.012	0.64 ± 0.008
G_4	0.61 ± 0.013	0.63 ± 0.014	0.63 ± 0.012	0.63 ± 0.008	0.63 ± 0.012	0.63 ± 0.009	0.63 ± 0.008
G_5	0.63 ± 0.013	0.62 ± 0.013	0.64 ± 0.012	0.62 ± 0.011	0.64 ± 0.012	0.62 ± 0.007	0.64 ± 0.007
G_6	0.61 ± 0.014	0.61 ± 0.011	0.62 ± 0.011	0.62 ± 0.011	0.63 ± 0.008	0.62 ± 0.012	0.63 ± 0.008
G_7	0.62 ± 0.011	0.64 ± 0.012	0.64 ± 0.011	0.63 ± 0.012	0.62 ± 0.009	0.59 ± 0.007	0.64 ± 0.007
G_8	0.63 ± 0.015	0.61 ± 0.011	0.62 ± 0.011	0.62 ± 0.009	0.64 ± 0.009	0.63 ± 0.007	0.64 ± 0.007
G_9	0.62 ± 0.014	0.62 ± 0.012	0.61 ± 0.008	0.62 ± 0.012	0.63 ± 0.011	0.62 ± 0.008	0.63 ± 0.008
G_{10}	0.59 ± 0.013	0.64 ± 0.012	0.63 ± 0.008	0.62 ± 0.011	0.63 ± 0.012	0.59 ± 0.008	0.64 ± 0.008
G_{11}	0.62 ± 0.015	0.64 ± 0.015	0.62 ± 0.011	0.61 ± 0.008	0.63 ± 0.009	0.61 ± 0.007	0.64 ± 0.007
G_{12}	0.62 ± 0.011	0.61 ± 0.013	0.61 ± 0.013	0.63 ± 0.012	0.64 ± 0.011	0.62 ± 0.008	0.64 ± 0.008
G_{13}	0.61 ± 0.013	0.63 ± 0.013	0.63 ± 0.008	0.61 ± 0.009	0.63 ± 0.012	0.63 ± 0.009	0.63 ± 0.008
G_{14}	0.63 ± 0.014	0.64 ± 0.014	0.59 ± 0.012	0.64 ± 0.012	0.63 ± 0.009	0.63 ± 0.012	0.64 ± 0.009
G_{15}	0.63 ± 0.012	0.63 ± 0.015	0.61 ± 0.008	0.62 ± 0.009	0.63 ± 0.012	0.59 ± 0.009	0.63 ± 0.008

References

Algarni AD (2020) Automated medical diagnosis system based on multi-modality image fusion and deep learning. Wirel Pers Commun 111:1033–1058

Babu B, Chakole PG, Mubeen JS (2005) Multiobjective differential evolution (mode) for optimization of adiabatic styrene reactor. Chem Eng Sci 60(17):4822–4837

Basavegowda HS, Dagnew G (2020) Deep learning approach for microarray cancer data classification. CAAI Trans Intell Technol 5(1):22–33

Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1251–1258

Da Cunha AL, Zhou J, Do MN (2006) The nonsubsampled contourlet transform: theory, design, and applications. IEEE Trans Image Process 15(10):3089–3101

Daniel E, Anitha J, Kamaleshwaran K, Rani I (2017) Optimum spectrum mask based medical image fusion using gray wolf optimization. Biomed Signal Process Control 34:36–43

Du J, Li W, Lu K, Xiao B (2016) An overview of multi-modal medical image fusion. Neurocomputing 215:3–20

Gai D, Shen X, Cheng H, Chen H (2019) Medical image fusion via pcnn based on edge preservation and improved sparse representation in nsst domain. IEEE Access 7:85413–85429. https://doi.org/10.1109/ACCESS.2019.2925424

Ghosh S, Shivakumara P, Roy P, Pal U, Lu T (2020) Graphology based handwritten character analysis for human behaviour identification. CAAI Trans Intell Technol 5(1):55–65

Gupta B, Tiwari M, Lamba SS (2019) Visibility improvement and mass segmentation of mammogram images using quantile separated histogram equalisation with local contrast enhancement. CAAI Trans Intell Technol 4(2):73–79

Hermessi H, Mourali O, Zagrouba E (2018) Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain. Neural Comput Appl 30(7):2029–2045

Hu Q, Hu S, Zhang F (2020) Multi-modality medical image fusion based on separable dictionary learning and gabor filtering. Signal Process Image Commun 83:115758

Jaiswal A, Gianchandani N, Singh D, Kumar V, Kaur M (2020) Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning. J Biomol Struct Dyn. https://doi.org/10.1080/07391102.2020.1788642

James AP, Dasarathy BV (2014) Medical image fusion: a survey of the state of the art. Inf Fusion 19:4–19

Kaur M, Singh D (2019) Fusion of medical images using deep belief networks. Clust Comput. https://doi.org/10.1007/s10586-019-02999-x

Kaur M, Gianey HK, Singh D, Sabharwal M (2019) Multi-objective differential evolution based random forest for e-health applications. Mod Phys Lett B 33(05):1950022

Kaur M, Singh D, Kumar V, Sun K (2020) Color image dehazing using gradient channel prior and guided l0 filter. Inf Sci 521:326–342. https://doi.org/10.1016/j.ins.2020.02.048

Kaur M, Singh D, Uppal R Singh (2020) Parallel strength pareto evolutionary algorithm-II based image encryption. IET Image Process 14(6):1015–1026

Kumar A, Fulham M, Feng D, Kim J (2020) Co-learning feature fusion maps from PET-CT images of lung cancer. IEEE Trans Med Imaging 39(1):204–217. https://doi.org/10.1109/TMI.2019.2923601

Lifeng Y, Donglin Z, Weidong W, Shanglian B (2001) Multi-modality medical image fusion based on wavelet analysis and quality evaluation. J Syst Eng Electron 12(1):42–48

Liu X, Mei W, Du H (2018) Multi-modality medical image fusion based on image decomposition framework and nonsubsampled shearlet transform. Biomed Signal Process Control 40:343–350

Liu Z, Wu J, Fu L, Majeed Y, Feng Y, Li R, Cui Y (2020) Improved kiwifruit detection using pre-trained VGG16 with RGB and NIR information fusion. IEEE Access 8:2327–2336. https://doi.org/10.1109/ACCESS.2019.2962513

Lu Y, Zhao J, Wang G (2014) Edge-guided dual-modality image reconstruction. IEEE Access 2:1359–1363. https://doi.org/10.1109/ACCESS.2014.2371994

Ma J, Xu H, Jiang J, Mei X, Zhang X (2020) Ddcgan: a dual-discriminator conditional generative adversarial network for multi-resolution image fusion. IEEE Trans Image Process 29:4980–4995. https://doi.org/10.1109/TIP.2020.2977573

Osterland S, Weber J (2019) Analytical analysis of single-stage pressure relief valves. Int J Hydromech 2(1):32–53

Pannu HS, Singh D, Malhi AK (2018) Improved particle swarm optimization based adaptive neuro-fuzzy inference system for benzene detection. CLEAN-Soil Air Water 46(5):1700162

Pannu HS, Singh D, Malhi AK (2019) Multi-objective particle swarm optimization-based adaptive neuro-fuzzy inference system for benzene monitoring. Neural Comput Appl 31:2195–2205

Prakash O, Park CM, Khare A, Jeon M, Gwak J (2019) Multiscale fusion of multimodal medical images using lifting scheme based biorthogonal wavelet transform. Optik 182:995–1014

Ravi P, Krishnan J (2018) Image enhancement with medical image fusion using multiresolution discrete cosine transform. In: Materials today: proceedings 5 (1, part 1) 1936 – 1942, international conference on processing of materials, minerals and energy (July 29th–30th) 2016, Ongole, Andhra Pradesh, India

Singh D, Kumar V (2018) Dehazing of outdoor images using notch based integral guided filter. Multimed Tools Appl 77(20):27363–27386

Singh D, Kumar V (2019a) A comprehensive review of computational dehazing techniques. Arch Comput Methods Eng 26(5):1395–1413

Singh D, Kumar V (2019b) Single image defogging by gain gradient image filter. Sci China Inf Sci 62(7):79101

Singh D, Kaur M, Singh H (2018) Remote sensing image fusion using fuzzy logic and gyrator transform. Remote Sens Lett 9(10):942–951

Singh D, Kumar V, Kaur M (2019a) Single image dehazing using gradient channel prior. Appl Intell 49(12):4276–4293

Singh D, Kumar V, Kaur M (2019b) Image dehazing using window-based integrated means filter. Multimed Tools Appl. https://doi.org/10.1007/s11042-019-08286-6

Tavard F, Simon A, Leclercq C, Donal E, Hernández AI, Garreau M (2014) Multimodal registration and data fusion for cardiac resynchronization therapy optimization. IEEE Trans Med Imaging 33(6):1363–1372. https://doi.org/10.1109/TMI.2014.2311694

Ullah H, Ullah B, Wu L, Abdalla FY, Ren G, Zhao Y (2020) Multi-modality medical images fusion based on local-features fuzzy sets and novel sum-modified-laplacian in non-subsampled shearlet transform domain. Biomed Signal Process Control 57:101724

Wang C, Zhao Z, Ren Q, Xu Y, Yu Y (2019) Multi-modality anatomical and functional medical image fusion based on simplified-spatial frequency-pulse coupled neural networks and region energy-weighted average strategy in non-sub sampled contourlet transform domain. J Med Imaging Health Inform 9(5):1017–1027

Wang Y, Zhou L, Yu B, Wang L, Zu C, Lalush DS, Lin W, Wu X, Zhou J, Shen D (2019) 3d auto-context-based locality adaptive multi-modality gans for pet synthesis. IEEE Trans Med Imaging 38(6):1328–1339. https://doi.org/10.1109/TMI.2018.2884053

Wang R, Yu H, Wang G, Zhang G, Wang W (2019) Study on the dynamic and static characteristics of gas static thrust bearing with micro-hole restrictors. Int J Hydromech 2(3):189–202

Wang K, Zheng M, Wei H, Qi G, Li Y (2020) Multi-modality medical image fusion using convolutional neural network and contrast pyramid. Sensors 20(8):2169

Wiens T (2019) Engine speed reduction for hydraulic machinery using predictive algorithms. Int J Hydromech 2(1):16–31

Xia K-J, Yin H-S, Wang J-Q (2019) A novel improved deep convolutional neural network model for medical image fusion. Clust Comput 22(1):1515–1527

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

Zhou T, Ruan S, Canu S (2019) A review: deep learning for medical image segmentation using multi-modality fusion. Array 3:100004

Zhu Z, Chai Y, Yin H, Li Y, Liu Z (2016) A novel dictionary learning approach for multi-modality medical image fusion. Neurocomputing 214:471–482

Zhu Z, Zheng M, Qi G, Wang D, Xiang Y (2019) A phase congruency and local laplacian energy based multi-modality medical image fusion method in NSCT domain. IEEE Access 7:20811–20824. https://doi.org/10.1109/ACCESS.2019.2898111

Zhu R, Li X, Zhang X, Ma M (2020) Mri and ct medical image fusion based on synchronized-anisotropic diffusion model. IEEE Access 8:91336–91350. https://doi.org/10.1109/ACCESS.2020.2993493