Multimedia Tools and Applications (2022) 81:42821–42847
https://doi.org/10.1007/s11042-022-13482-y

1221: DEEP LEARNING FOR IMAGE/VIDEO COMPRESSION AND VISUAL QUALITY ASSESSMENT

Multi-layer, multi-modal medical image intelligent fusion

Rekha R. Nair[1] · Tripty Singh[2] · Abhinandan Basavapattana[2] · Manasa M. Pawar[2]

Received: 24 June 2021 / Revised: 10 November 2021 / Accepted: 13 July 2022 /
Published online: 12 August 2022

Abstract
Recently, deep learning has high popularity in the field of image processing due to its unique feature extraction property. This paper, proposes a novel multi-layer, multi-tier system called Multi-Layer Intelligent Image Fusion(MLIIF) with deep learning(DL) networks for visually enhanced medical images through fusion. Implemented deep feature based multi-layer fusion strategy for both high frequency and low frequency components to obtain more informative fused image from the source image sets. The hybrid MLIIF consists of VGG-19, VGG-11, and Squeezenet DL networks for different layer deep feature extraction from approximation and detailed frequency components of the source images. The robustness of the proposed multi-layer, multi-tier fusion system is validated by subjective and objective analysis. The effectiveness of the proposed MLIIF system is evaluated by error image calculation with the ground truth image and thus accuracy of the system. The source images utilized for the experimentations are collected from the website www.med.harvard.edu and the proposed MLIIF system obtained an accuracy of 95%. The experimental findings indicate that the proposed system outperforms compared with existing DL networks.

Keywords Medical image fusion · Activity level map · Block based average operator · l_1-norm · VGG-19 · VGG-11 · SqueezeNet

1 Introduction

Off late, processing with the medical images has become an interesting field of research, with significant advancement in technology and instruments because of its crucial involvement in the health sector. The information from various perspectives about human organs

✉ Tripty Singh
tripty_singh@blr.amrita.edu

[1] Department of Computer Applications, School of Engineering, Dayananda Sagar University, Bengaluru, India

[2] Department of Computer Science and Engineering, Amrita School of Engineering, Bengaluru, Amrita Vishwa Vidyapeetham, India

is provided in areas of imaging in several medical modes, such as magnetic resonance imaging (MRI), positron emission tomography (PET), computed tomography (CT), and single-photon emission computed tomography (SPECT) [12, 22, 27] etc. CT represents the bone structure contour and complex bone structures, but without soft tissue characterization and lacks in identifying changes in physiology [1]. The MRI, more precisely represent the normal and abnormal structures of soft tissues of heart, brain, and eyes [23]. Detailed information about organ metabolism and blood supply in brain tissues is provided by PET images [23]. CT and MRI are usually paired with PET for details of anatomy and metabolism. SPECT discloses details about the functionality of the blood vessels [23]. Every category, therefore, possesses its characteristics, and the optimum solution is obtained by combining various modal images to attain the required details in a single image [3]. Fusion algorithms are the research focus topic in this field, to obtain better fused images while retaining input information. The goal of multimodal image fusion is to collect and disclose imperceptible information to humans in different modes [15]. A multi-modal fusion of medical images is the combination of various modality images and involves a variety of techniques dealing with medical problems depicted in organ images [25].

The fusion of medical images is typically performed in three phases i.e. pixel level, feature level, and decision level. Information related to the pixel intensity is the pixel level fusion [6]. Extracts different attributes like color, edges, and textures from the images, and then incorporated them into the final image in feature level [2]. The decision level is the highest level of fusion and deals with the initial identification and categorization of the fusion algorithm [33].

In the past three decades, image fusion analysis was the topic of discussion about significant social advantages [24]. The principal concern in the algorithms for image-fusion is the computation of a weight map that integrates information from various source modes to generate a final fused image. The study is focused on hand-craft features in traditional fusion methods that are very susceptible and effortlessly impacted by misalignment and noise. Therefore, its supremacy can not be developed as a robust method for extracting and fusing information at the activity level. Some DL approaches aim to explicitly and automatically learn from raw data without using manually developed features to achieve meaningful representations and features. However, data from health care are rising exponentially in the current scenario and can be very useful to evaluate and predict the same DL-based methods. Multi-spectral face image fusion carried out with discriminative deep feature learning [30, 31]. When the neural networks started working on superior methods on various levels of prioritized image analysis in computer vision science, the deep learning approach became an outstanding model. In the medical field, the importance of deep learning has begun with the introduction of the CNNs [17] and is helpful approach in learning essential image characteristics & other structured information. From the "content" image, generated image, and "style" image, VGG-network extracts [29] deep features at various layers. Iteration minimizes distinctions in profound characteristics from images and input images generated. But the discussed method based on VGG-network is extremely slow at execution on GPU system. A feed-forward network has been suggested to overcome these drawbacks to address the optimization in the real-time problem of the above [5]. Every network is connected with a fixed style based on feed-forward network. The networks with VGG network and adaptive instance normalization are limited in use of multi-layer network features [5, 9]. Later ImageNet trained fixed VGG-19 [29] is proposed to extract the parameters or features [20].

The main key feature of fusion is to help doctors and radiologists in diagnosis and treatment. The fusion of two or more medical images helps to obtain all the information in the source images into a single image, that are acquired from various sensors. The deep feature

based multilayer fusion strategy for both high frequency and low frequency components helps to obtain more informative fused image from the source image sets. VGG-11 and VGG-19 are good at detecting high-level features like edges, patterns, etc. A model with SqueezeNet can provide more accuracy and lesser storage space. It is able to extract more information compared to single-layer fusion with multilayer fusion. All these parameters encouraged to propose of a multi-layer, multi-modal fusion of medical images.

Taking account of the above fact, the most appropriate pre-trained frame for multimodal image fusion to improve image quality is considered in this work. The aim and motivation of the research is to achieve the following:

- Analyze, compare and identify the best-supported Deep Neural Network(DNN) for the novel MLIIF system with identified medical image data sets for existing pre-trained Squeezenet, Alexnet, VGG-11, VGG-13, VGG-16 and VGG-19 systems and other exiting published works.
- To achieve a visually enhanced & informative fusion quality, develop a multi-layer framework for medical images.
- The modal was quantitatively evaluated (13 performance measurements) with five different pairs of publicly available medical images to determine the robustness of the system. The system's effectiveness is identified by the calculation of the error image and accuracy of the model.

The contributions of this work are as follows:

- Design and implement a hybrid intelligent medical image fusion model for better visualization and accuracy of the model.
- Implement a multilayer fusion strategy based on activity level map with l1-norm and block based average operator for the fusion of source images.
- Identification of most suitable locations in the layer for the existing pre-trained Alexnet, VGG-19, VGG-16, VGG-13, VGG-11, and identify the best models among them for the proposed MLIF system.
- The proposed method is compared with existing pre-trained DNN(Deep Neural Network) such as Alexnet, VGG-19, VGG-16, VGG-13, VGG-11, and Squeezenet with MLIIF to observe the image interpretation and enhanced medical image information obtained by the fusion.
- To prove the state-of-art, the proposed MLIIF system is compared with published works.
- The robustness of the proposed multi-layer, multi-tier fusion system is validated by subjective and objective analysis.
- The effectiveness of the proposed MLIIF system is evaluated by error image calculation by comparison of resultant fused image with the ground truth image and thus accuracy of the system.
- The main contribution of the work is to help medicos for diagnosis and treatment.

The remaining part of the work is structured accordingly. The proposed methodology is explained in Section 2. The assessment methods are discourse in Section 3. The data set is listed in Section 4 and the experimental results and the analyses found in Section 5. The conclusion is in Section 6 and future scope of the work is in Section 7.

2 Proposed methodology

The proposed MLIIF system presents a 2-tier fusion technique with various DNN to attain the final fusion result as illustrated in Fig. 1. The system encompasses five phases, (i) Source image decomposition (ii) Proposed fusion outline (iii)Fusion of low or approximation(LF) and high or detailed(HF) frequency components (iv) Reconstruction of decomposed image (v) Error detection. The outline of the entire system is as follows.

2.1 Source image decomposition

The input images are decomposed into approximation and detailed components by an optimization method [19] for better time reduction and effectiveness. The input images are assumed to be registered. The low frequency or base components, and detailed contents, H_i^d are separated for each source image, I_i [19]. By solving this optimization problem, the base parts are obtained:

$$L_i^b = L_i^b \; agm||I_i - L_i^b||_F^2 + \lambda\left(||g_x * L_i^b||_F^2 + ||g_y * L_i^b||_F^2\right) \tag{1}$$

where,L_i^b is the low frequency or base components and *agm* represets *argmin*. The parameter "λ" is set to 5. The horizontal operator, g_x =[−1 1] and vertical operator, g_y

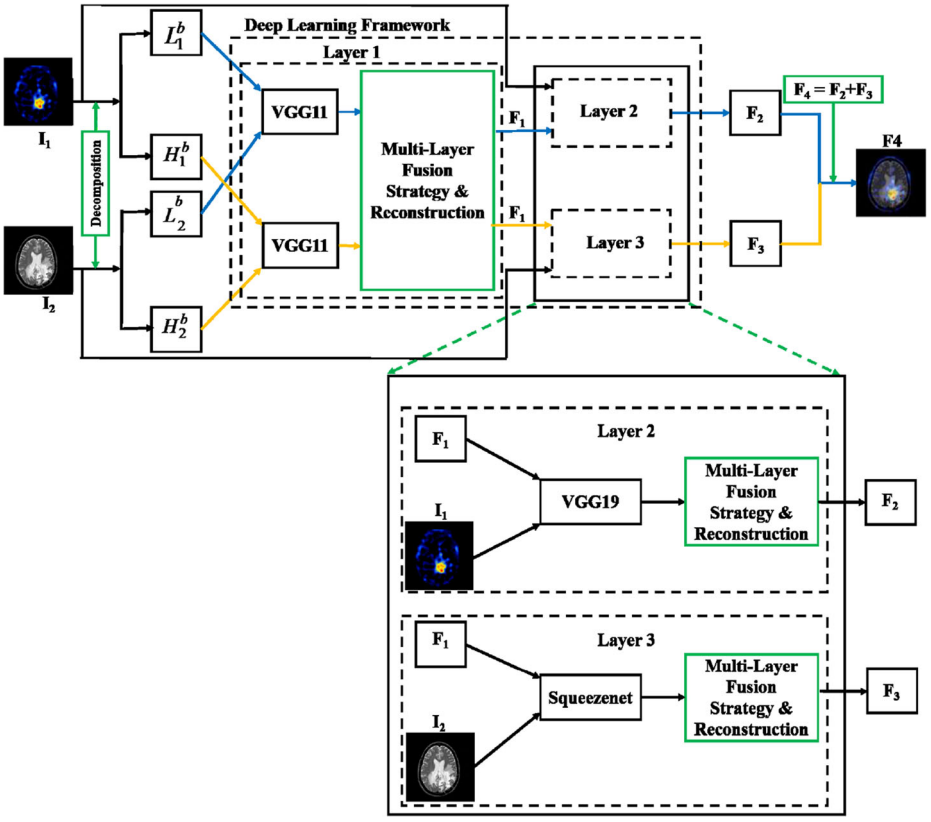

Fig. 1 Proposed block diagram

$=[-1 \quad 1]^T$. The high frequency component is derived by the (2),

$$high\ frequency\ component,\ H_i^d = I - L_i^b \tag{2}$$

The decomposed images are passed to the next stage for the further fusion process.

2.2 Proposed fusion outline

In the reconstruction process, a novel framework is adopted with deep learning and multi-layer strategy as shown in Fig. 1. The DNN models utilized in the work are VGG-19, VGG-11, and SqueezeNet. All the models are worked with default parameters.

1. VGG-19 [4] & VGG-11 [11]: The VGG accepts decomposed low frequency and high-frequency components as input and passes them through to a convolutionary stack layers of filter size 3 x 3 and stride 1. The input image can be downsampled via five maximum pooling filters embedded among convolutionary layers. The stack is accompanied by 3 completely linked layers, each of which comprises 4096, 4096, and 1000 channels. The network has a 224x224 image feed. VGG-11 and VGG-19 have 11 and 19 layers respectively.
2. SqueezeNet [10]: This generates a smaller neural network with fewer features or parameters and is effectively transmitted via a computer network. This improvisation is 3 times faster and 500 times smaller than AlexNet. It consists of a stack of pooling layers and fire modules. The squeezing layer and expanding layer retains the same size of the map as the former, which decreases the depth to a smaller level. Squeezing or bottleneck layer and expansion is common in neural architectures.

2.2.1 Responsibilities of each layer of the proposed model

- The source images $I1$ and $I2$ are decomposed by (1) and (2) to obtain L_i^b, H_i^d as illustrated in Fig. 1, where $i \in \{1, 2\}$.
- In $Layer1$, VGG-11 is used to extract deep features from the decomposed coefficients. The deep features are fused and reconstructed by multi-layer fusion strategy as displayed in Fig. 2 to obtain the fused image F_1.
- The fused image F_1 is passed to $Layer2$ and $Layer3$ as a source image 1 for each Layer.
- In $Layer2$, VGG-19 is used to extract deep features from the source images F_1 and I_1. The deep features are fused and reconstructed by multi-layer fusion strategy as displayed in Fig. 2 to obtain the fused image F_2.
- In $Layer3$, Squeezenet is used to extract deep features from the source images F_1 and I_2. The deep features are fused and reconstructed by multi-layer fusion strategy as displayed in Fig. 2 to obtain the fused image F_3.
- The final fused image, F_4 is obtained by the addition of F_2 and F_3. The entire work flow is demonstrated in Fig. 1.

2.3 Fusion of low and high-frequency components

The fusion strategy procedure for deep learning networks VGG-19, VGG-11 and SqueezeNet for extracting deep features from LF and HF components are illustrated in Fig. 2. The deep features of the components is extracted by VGG-11 for layer 1, VGG-19 for layer 2 and SqueezeNet for layer 3. Weight maps in Fig. 2 is obtained using multi-layer

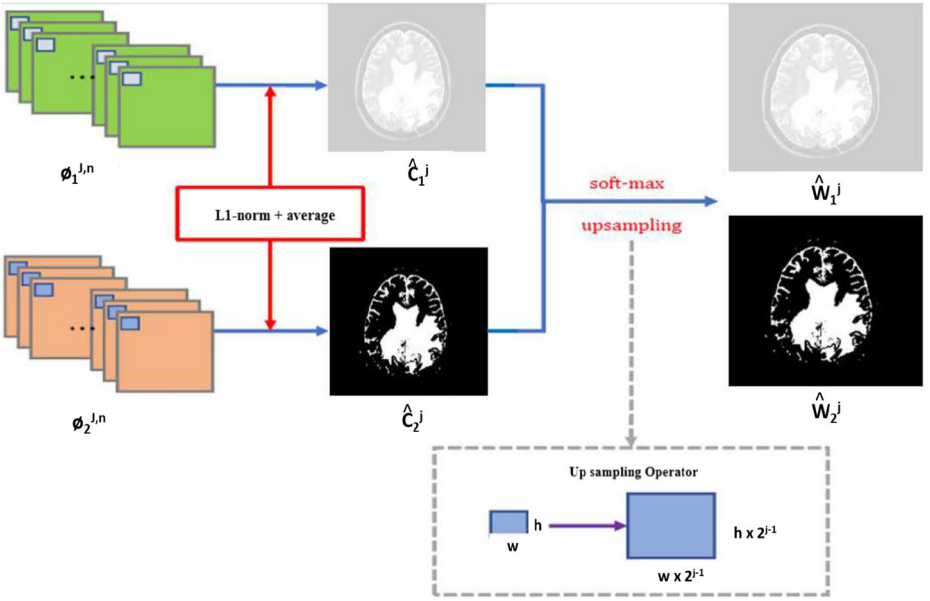

Fig. 2 Demonstration of multi-layer fusion strategy for approximation and detailed components

fusion strategy. The fused approximation and detailed comeponents are reconstructed back with weight maps and approximation and detailed comeponents [18].

The detailed explanation of multi-layer fusion strategy for the detailed component, H_i^d as follows:

In Fig. 2 $\phi_i^{j,n}$ be the feature map at k^{th} HF component from the layer j^{th} and the channel number is n. Let $n \in \{1, 2, ..., N\}$ and $N = 64x2^{j-1}$.

$$Deep\ feature,\ \phi_i^{j,n} = \Phi_i(H_i^d) \tag{3}$$

Where, $\Phi_i(.)$ be a layer in the each layer DL network. The $relu\_1\_1$, $relu\_2\_1$, $relu\_3\_1$ and $relu\_4\_1$ be represented by $j \in \{1, 2, 3, 4\}$.

The components of $\phi_i^{j,n}$ at (x, y) position in the feature map is represented by $\phi_i^{j,1:N}(x, y)$ and is an N-dimensional vector. After obtained the deep features, $\phi_i^{j,n}$, the activity level map, $\hat{C_i^j}$, calculated by $l_1 - norm$ and block based average operator. Let the value of $i \in \{1, 2\}$. The activity level measure of detailed source component be $l_1 - norm$ of $\phi_i^{j,1:N}(x, y)$. The activity level map at initial stage is calculated by:

$$C_i^j(x, y) = ||\phi_i^{j,1:N}(x, y)||_1 \tag{4}$$

where $||\phi_i^{j,1:N}(x, y)||_1$ is the $l_1 - norm$ of $\phi_i^{j,1:N}(x, y)$ at (x, y) position.

From the feature map, at (x, y) position the components of $\phi_i^{j,n}$ is $\phi_i^{j,1:N}$. Block-based average operator is used to calculate final activity level map $\hat{C_i^j}$. The final activity level map

helps to obtain a robust aligned image for fusion.

$$\hat{C}_i^j(x, y) = \frac{\sum_{\beta=-r}^{r} \sum_{\theta=-r}^{r} C_i^j(x+\beta, y+\theta)}{(2r+1)^2} \tag{5}$$

where, $\frac{\sum_{\beta=-r}^{r} \sum_{\theta=-r}^{r} C_i^j(x+\beta, y+\theta)}{(2r+1)^2}$ is the block based average operator. Where θ and β are the angles from origin of x, y coordinate and r is the block size.

Set the block size $r = 1$ for robust alignment and to reduce the loss in the HF component. If the block size, r, value is larger, some details may lost. In Fig. 2 W_i^j be the initial weight computed by the soft-max operator:

$$W_i^j(x, y) = \frac{\hat{C}_i^j(x, y)}{\sum_{n=1}^{K} \hat{C}_i^j(x, y)} \tag{6}$$

The number of activity level maps, K=2. $W_i^j(x, y)$ represents the initial value of weight map ranged [0,1]. Let pooling operator stride be s=2 and always resizes to $1/s$ of the original image. The map size of the feature at various levels is $1/2^{j-1}$ times the HF component size. After obtaining each weight map W_i^j, the weight map resizes to the source HF component size using the upsampling operator. Compute the final weight map by using (7).

$$\hat{W}_i^j(x + p, y + q) = W_i^j(x, y) \tag{7}$$

where, $p, q \in \{0, 1, ..., (2^{j-1} - 1)\}$ After this equation obtains four pairs of weight maps and in each pair, initial fused HF content is calculated by (8).

$$F_d^j(x, y) = \sum_{n=1}^{K} \hat{W}_n^j(x, y) * H_n^d(x, y), \quad K = 2 \tag{8}$$

The final fused HF content is obtained from the maximum among the four initial fused HF content.

$$F_d(x, y) = max[F_d^j(x, y) | i \in \{1, 2, 3, 4\}] \tag{9}$$

Similarly, Low-frequency or approximation component weight map is calculated.

2.4 Reconstruction of decomposed image

The final fused image F_4 is attained by the addition of F_2 and F_3 as demonstrated in (10).

$$F_4(x, y) = F_2(x, y) + F_3(x, y) \tag{10}$$

2.5 Error detection

The error image I_{error} is obtained by subtracting fused image from the ground truth or reference image as demonstrated in (11).

$$I_{error}(x, y) = I_{ref}(x, y) - F_4(x, y) \tag{11}$$

3 Quantitative evaluation measures

The performance evaluation is based on quantitative or objective evaluation measures for the proposed system. The non-reference performance measures includes mutual information(MI), entropy(Ef), standard deviation(SD), joint entropy(JE), Q_hnc, image spatial quality evalution(ISQE) [21], structural similarity index metric(SSIM), edge strength (Q_abf), spatial frequency(SF), fusion factor(FF) and Fusion Symmetry(FS) as tabulated in Table 1. The performance measurements based on the reference or ground truth, such as accuracy, root mean square error (RMSE) and peak signal to noise ratio(PSNR) are described below. The paper summarizes all relevant non-reference performance metrics used in the research. The non-reference performance metrics used in the research is described in papers [22] and [23]. The value of ISQE in the measured in 100's.

The reference or ground truth based performance measures are:

1. Accuracy: The accuracy of the fused image and the ground truth images are calculated based on MATLAB function called "bfscore".

$$Accuracy = bfscore(Fused\ image, Ground\ truth\ image) \tag{12}$$

2. Root mean square error(RMSE): It provides the amount of change per pixel between the reference and fused image due to processing [26]. The smaller the error value, the better is the fused image. The RMSE between the reference image and the image being fused is provided by:

$$RMSE = \sqrt{\frac{1}{XY}\sum_{i=1}^{X}\sum_{j=1}^{Y}\left(I_r(i,j) - I_r(i,j)\right)} \tag{13}$$

Where $I_r(i,j)$ and $I_f(i,j)$ are the intensity value of reference image and fused image at pixel location (i, j). The root mean square error should be lower for better quality images.

3. Peak signal to noise ratio (PSNR): The strength of signal is displayed by PSNR [26]. A higher value means improved fusion.

$$PSNR = 20log_{10}\left[\frac{L^2}{RMSE}\right] \tag{14}$$

where L is the number of gray level in the image.

The non-reference performance measures that are tabulated in Table 1 are:

1. The standard deviation(SD) is an estimation of contrast probability distribution of a picture element in the image. It is the aggregate of noise parts and signal and illustrates the contrast in the fused image. A high disparity image would have a high SD. If the SD is large, then the image greyscale is more scattered and the contrast is sharper. A higher value of standard deviation shows the high contrast of the image.
2. Entropy(Ef) is a statistical measure that shows the texture information of the fused image. The entropy of a fused image signifies the information content and denotes the correlation between the pixels of an image. The entropy of an image may target the difference between neighbour regions. The fused image offers more information in higher value entropy than entropy with lower values.
3. Spatial frequency(SF) provides information about spatial deviations such as fine details of an image, changes in edge, information about features etc. A higher value of SF fused image will be considered as the preferred result.

Table 1 Performance Metrics

Sl. No	Quality Metrics	Formula	Preferred Value
1	Entropy, Ef	$Ef = -\sum_{i=0}^{i=1} log_2(x_i) Pr(x_i)$	High
2	Joint Entropy, JE	$JE = H(X, Y) = \sum_{x=\chi} \sum_{y=\gamma} P(x, y) log P(x, y)$	Low
3	Standard Deviation, SD	$SD = \sqrt{\left(\frac{1}{XY}\right) \sum_{j=1}^{X} \sum_{i=1}^{Y} \left(I_{mf}(j, i) - \mu\right)}$	High
4	Mutual Information, MI	$MI = MI^{XF} + MI^{YF}$	High
5	Edge Strength, Q_{abf}	$Q_{abf} = \frac{\sum_{a=1}^{A} \sum_{b=1}^{B} Q^{XF}(a,b) W^{X}(a,b) + Q^{YF}(a,b) W^{Y}(a,b)}{\sum_{a=1}^{A} \sum_{b=1}^{B} W^{X}(a,b) + W^{Y}(a,b)}$	(0 - 1) Nearer to 1
6	Structural Similarity Index, SSIM	$SSIM = \frac{SSIMVAL1 + SSIMVAL2}{2}$	Range: (0 to 1) Prefer: Nearer to 1
7	Q_{hnc}	$Q_{hnc} = 0.2\left(\left(\frac{I_{(a,f)}}{H_a + H_f}\right) + \left(\frac{I_{(b,f)}}{H_b + H_f}\right)\right)$	(-1 to 1) Nearer to 1
8	Image Spatial Quality Evaluation	$ISQE = brisque(imf)$	Low
9	Spatial Frequency, SF	$SpatialFrequency, \quad SF = \sqrt{(RF^2 + CF^2)}$	High
10	Fusion Symmetry, FS	$FS = abs\left(\frac{I_{im1imf}}{I_{im1imf} + I_{im2imf}} - 0.5\right)$	High
11	Fusion factor, FF	$FF = I_{im1imf} + I_{im2imf}$	High
12	Fusion Quality Index, FQI	$FQI = \sum \left(c(w)\left(\lambda(w) QI(I_1, I_f \mid w)\right) + \left(1 - \lambda(w) QI(I_1, I_f \mid w)\right)\right)$	Range: (0 to 1) Prefer: Nearer to 1

4. Fusion factor(FF) provides information about the transfer of grayscale details from one image to another. High value FF provides more details about the image.
5. Fusion Symmetry(FS) denotes the dissimilarity with fused image and the source images. The lowest value of FS shows that the fused image is of better performance.
6. Mutual information (MI) depicts the total information of fused image acquired from the input images. A relatively high value of MI gives better performance than a relatively low value of MI.
7. Structural similarity index metric($SSIM$) shows the similarity between the source images and fused image and the range of value lies between 0 and 1. Value 1 depicts that all information from the input images contains in the fused image. The value nearer to 1 provides better results.
8. Image spatial quality evaluation($ISQE$) used to evaluate the image quality based on pixel deviations between images. The low value of $ISQE$ is preferred.
9. Edge strength (Q_{abf}) depicts the edge details related to the fused image and supports the human visual system. The higher value preferred in a range (0 to 1).
10. The $Qhnc$ based on mutual information and entropy and higher value in a range (-1 to 1) is preferred.
11. Joint entropy(Joint histogram, JE) is a measurement for the uncertainty related to a set of variables. For a digital image, entropy is a probability distribution of gray values. If joint entropy (uncertainty of two images) is low then the similarity of two images is high and vice versa.
12. Fusion quality index(FQI) provides information content transfer from source to fused image and ranges from (0 to 1). FQI value 1 shows that all information from the source images are contained in the fused image.

4 Data source images

The proposed model, MLIIF, was evaluated with different modality images (Neoplastic disease of brain tumor from the whole brain atlas) as shown in Fig. 3, and are collected from the website www.med.harvard.edu [13]. Tested with 5 sets of color images of brain organ that consists of normal and abnormal images. All the images are of size $224 \times 224 \times 3$ and in GIF format. Set 1 is of MR-Gad and PET, set 2 is of MR-Gad and SPECT-T1, Set 3 is of MR-T1 and PET, set 4 is MR-T2 and PET, and set 5 is MR-T2 and Spect. MRI images are T1-weighted(MR-T1) and T2-weighted(MR-T2) images [23]. The following Table 2 provides is a brief description of the dataset details and the evaluation slice number of each pair.

5 Experimental result and analysis

The proposed MLIIF system is validated and compared with the existing DNN on visualization and objective assessment. In comparison to six other networks of deep learning and four published work, the proposed work proves the efficiency of the system.

The proposed MLIIF system (Algorithm-7, A7) is compared with: (a) existing DNN and (b) published works as follows:

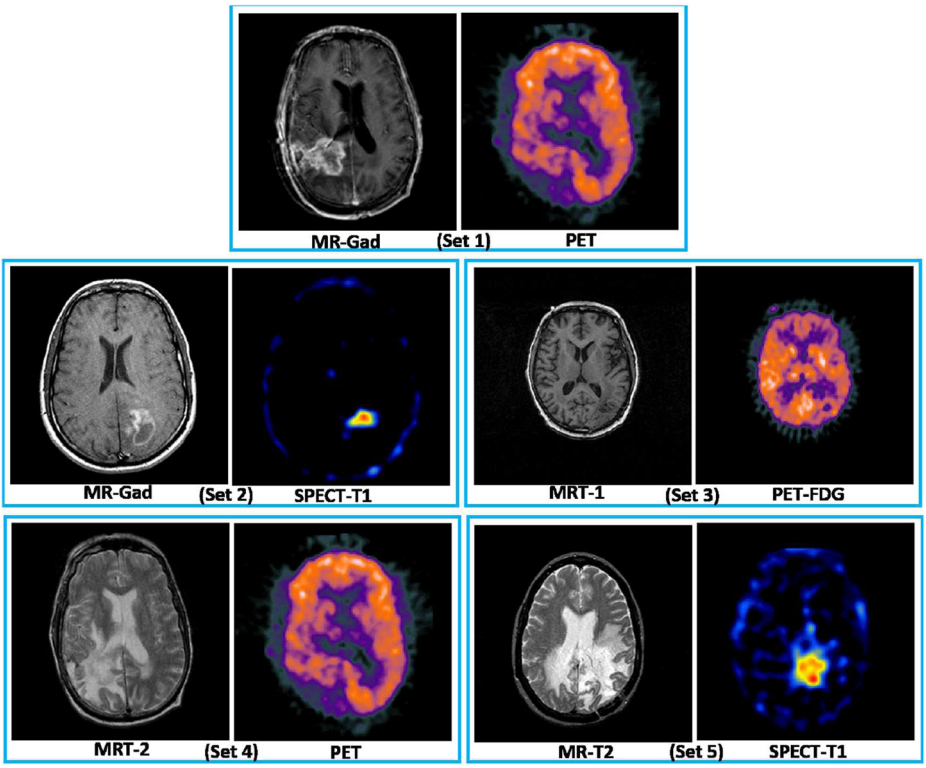

Fig. 3 Source Images utilized for the Proposed System

– Existing DNN networks such as VGG-19 [4] as Algorithm-1(A1), VGG-16 [28] as Algorithm-2 (A2), VGG-13 [29] as Algorithm-3(A3), VGG-11 [11] as Algorithm-4(A4), AlexNet [16] as Algorithm-5 (A5), and SqueezeNet [10] as Algorithm-6(A6).
– Published works such as AlgorithmRefered-1(Ar1) [14], AlgorithmRefered-2(Ar2) [32], AlgorithmRefered-3 (Ar3) [7], and AlgorithmRefered-4(Ar4) [8].

The proposed MLIIF system is evaluated with:

– Evaluation based on subjective assessment
– Evaluation based on error image

Table 2 Source dataset details

Dataset	Modality	Evaluation Slice no	Images total pairs (Normal & Malignant)	Body part
1	PET/MR-Gad	12	18	Brain
2	SPECT-T1/MR-Gad	17	44	Brain
3	PET/MR-T1	63	107	Brain
4	PET/MR-T2	13	19	Brain
5	SPECT/MR-T2	33	78	Brain

- Evaluation based on objective assessment
  - With ground truth images
  - Based on non-reference performance measures without ground truth images
- Evaluation based on accuracy

5.1 Experimental settings

The data set information is given in Section 4. In addition, the MLIIF system proposed uses VGG19, VGG-11 and SqueezeNet to extract deep features from LF and the HF components using the pre-trained layers of $relu\_1\_1$, $relu\_2\_1$, $relu\_3\_1$ and $relu\_4\_1$.

The hardware configuration of the system utilized for the proposed study is tabulated in Table 3. The test is conducted on Windows 10 operating system with Python 3.8 programming language on integrated development environment PyCharm 2021.1. 1n (Professional Edition). The machine learning libraries utilized for the proposed work is TensorFlow, NumPy, & Matplotlib.

5.2 Evaluation based on subjective assessment

In Fig. 9, images of fusion obtained through six deep learning networks and the proposed MLIIF system are displayed. The fusion results of published works and proposed MLIIF system is illustrated in Fig. 10. The various fused images obtained at different layers and sections of the proposed system, MLIIF is illustrated in Fig. 11. The fusion results of existing DNN and proposed MLIIF system is displayed in Fig. 9. In Fig. 10 the published works with the proposed MLIIF system is displayed. It can be perceived from the yellow window that, the resulting fused image generated by the proposed system conserves more informative information with less noise and better visualization of parenchyma or tissue structure. The assessment of subjective evaluation of the proposed system is validated with 5 radiologists from HCG, Bangalore and the details about radiologist are provided in Table 4.

The graphical user interphase(GUI) used for subjective evaluation is displayed in Fig. 4. The evaluation results for the existing DNN with proposed MLIIF is illustrated in Fig. 6 and published works and MLIIF is illustrated in Fig. 5. In the GUI, the top two images are the source images and the corresponding fusion result is in the bottom part of the Fig. 4. The radiologist evaluated the fused image and rated it in a range of 1 to 10. However, the algorithm used for results is not known to the radiologist. From Figs. 5 and 6, "TnE" represents the total number of evaluations carried out and "EvgE" represents the average of evaluations carried out by the radiologist. The visualization of the fused image as shown in Figs. 9 and 10, indicates that the proposed MLIIF system is displaying better tissue structure and parenchyma compared to other algorithms. The subjective evaluation result as illustrated in Figs. 5 and 6 displays the better performance of the MLIIF system (Figs. 7, 8, 9 and 10).

Table 3 Hardware configuration

Graphics Processing Unit:	4GB NVIDIA GeForce GTX 1650 Ti
Central Processing Unit:	Intel(R) Core (TM) i5-10300H CPU @ 2.50GHz
RAM size:	24.0 GB

Table 4 Details of the radiologist

Sl.No	Name	Age	Gender(F/M)	Work experience
1	Radiologist 1	53	M	25
2	Radiologist 2	35	M	6.8
3	Radiologist 3	42	M	10
4	Radiologist 4	47	F	21
5	Radiologist 5	56	M	29

5.3 Evaluation based on error image

The errors between the reference or the ground truth image, I_{ref}, and the fused image, F_4 are displayed in Fig. 11. It is observed from Fig. 11 that the proposed system fused image, F_4, is not showing much deviation from the ground truth image for all the sets of source images. The error image from the resultant fused image and the ground truth image calculated provides a negligence difference. This reveals the effectiveness and robustness of the proposed MLIIF system.

5.4 Evaluation based on accuracy

The accuracy of the proposed MLIIF system and the existing systems are calculated based on ground truth image and comparison of the accuracy is provided in Table 8. The proposed MLIIF system provided accuracy of 95.09% than existing DL networks even if execution time is slightly high based on milliseconds.

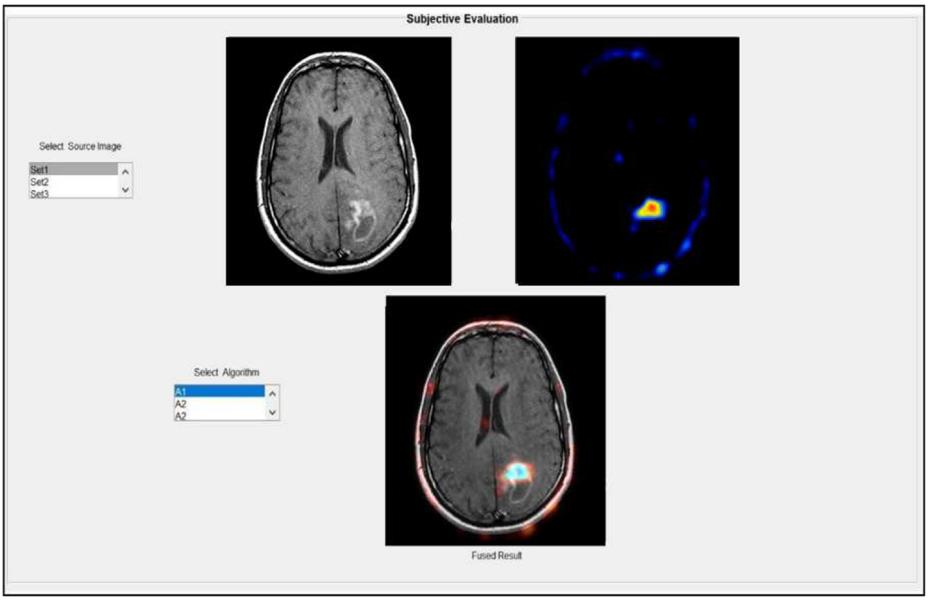

Fig. 4 GUI utilized for subjective evaluation

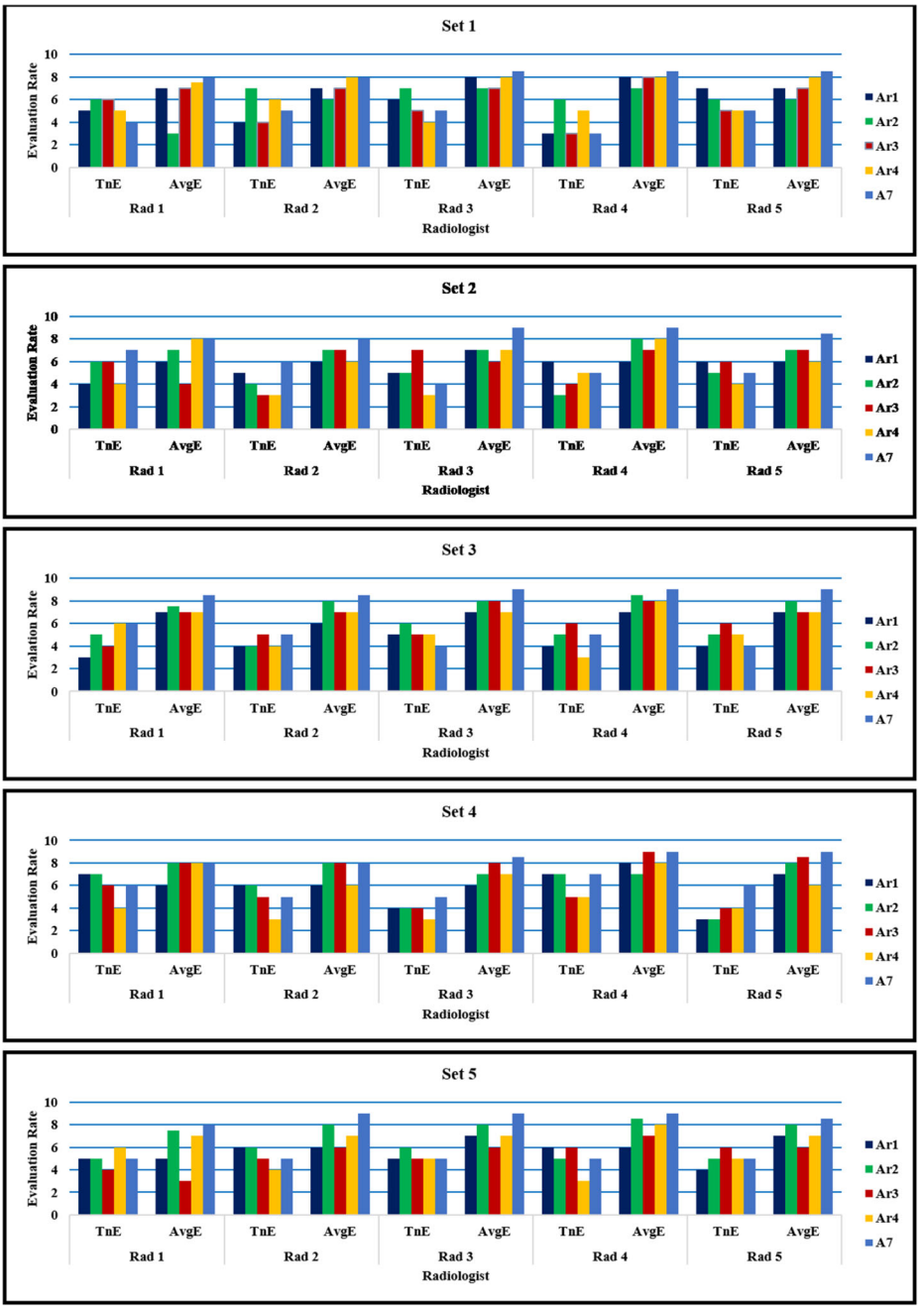

Fig. 5 Subjective evaluation of published works with proposed method. set1 is MR-Gad/PET, set1 is MR-Gad/SPECT-T1, set3 is MRT-1/PET-FDG, set4 is MRT-2/PET and set5 is MRT-2/SPECT-T1

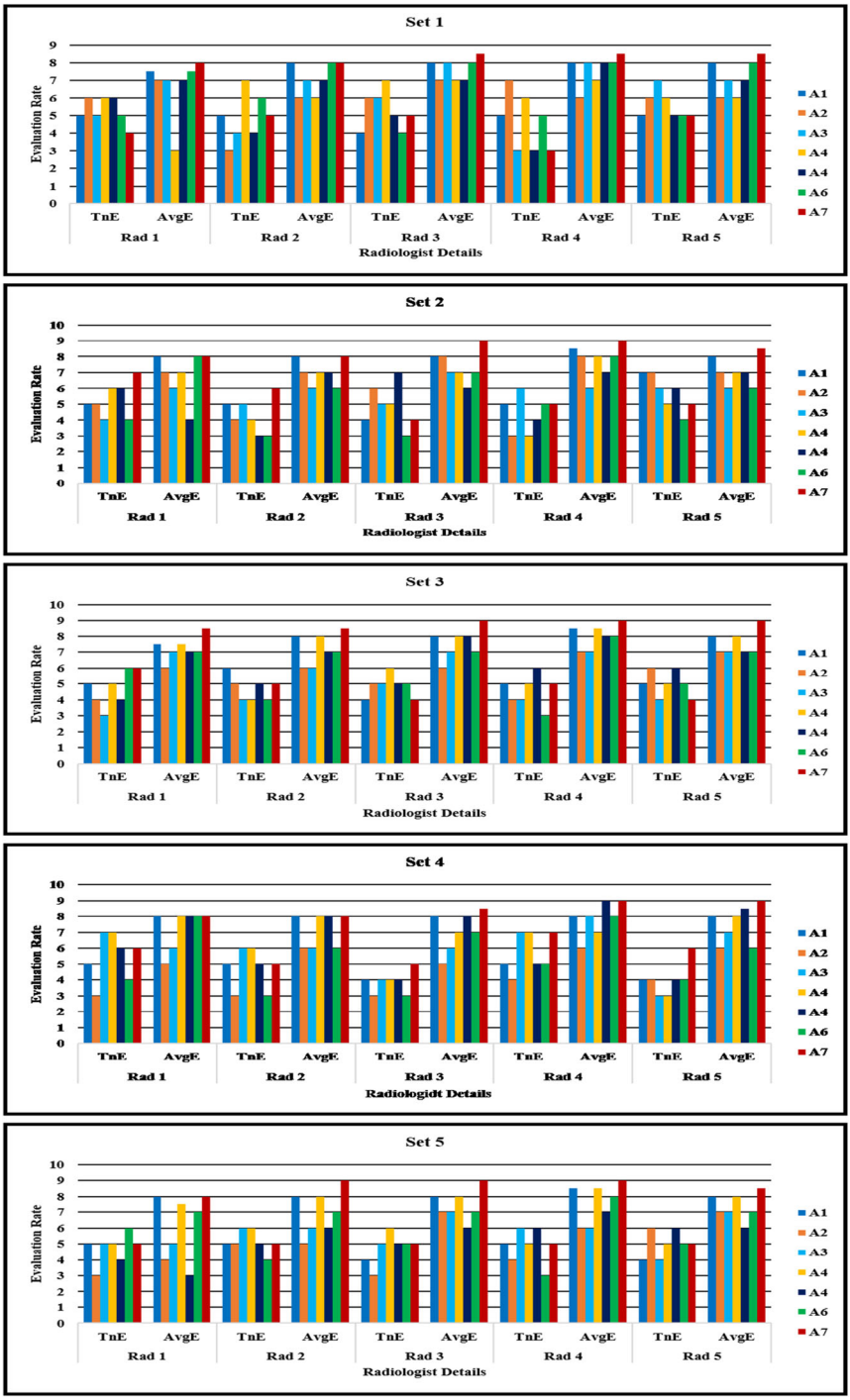

Fig. 6 Subjective evaluation of existing DNN with proposed method. set1 is MR-Gad/PET, set1 is MR-Gad/SPECT-T1, set3 is MRT-1/PET-FDG, set4 is MRT-2/PET and set5 is MRT-2/SPECT-T1

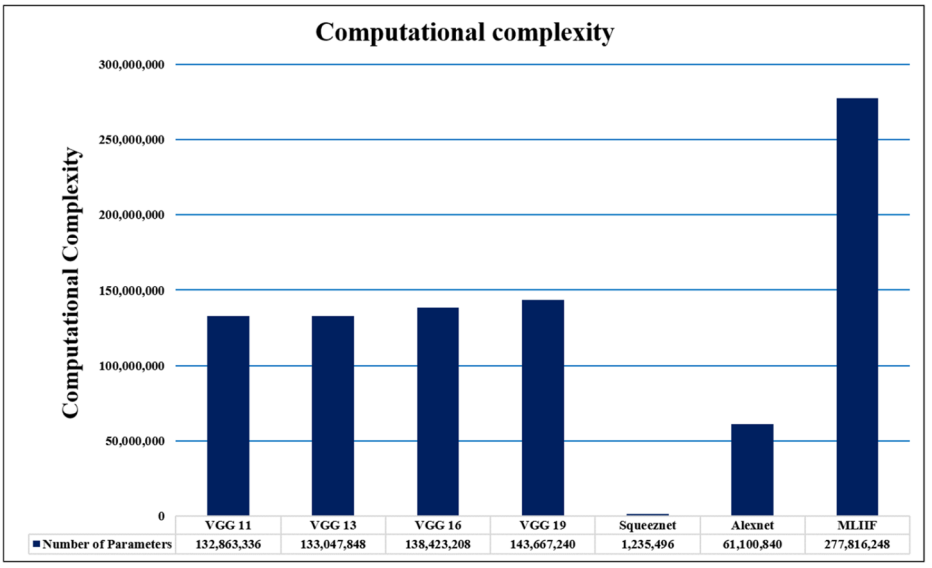

Fig. 7 Computational complexity of all the existing DNN and proposed MLIIF

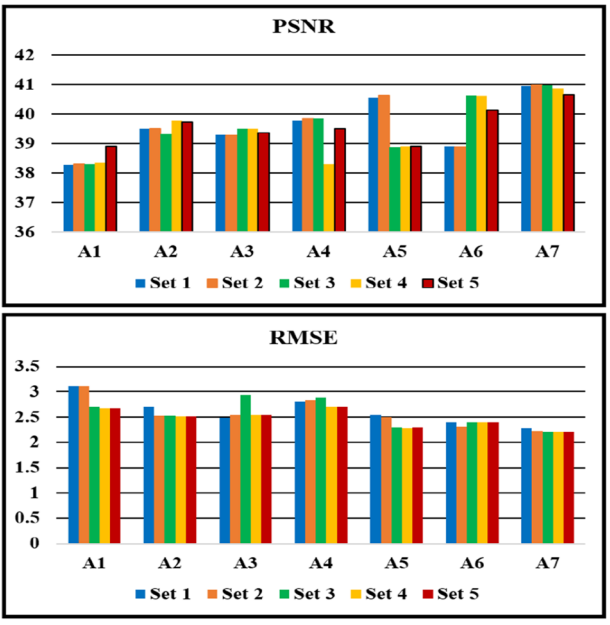

Fig. 8 PSNR and SNR values for all the sets of data

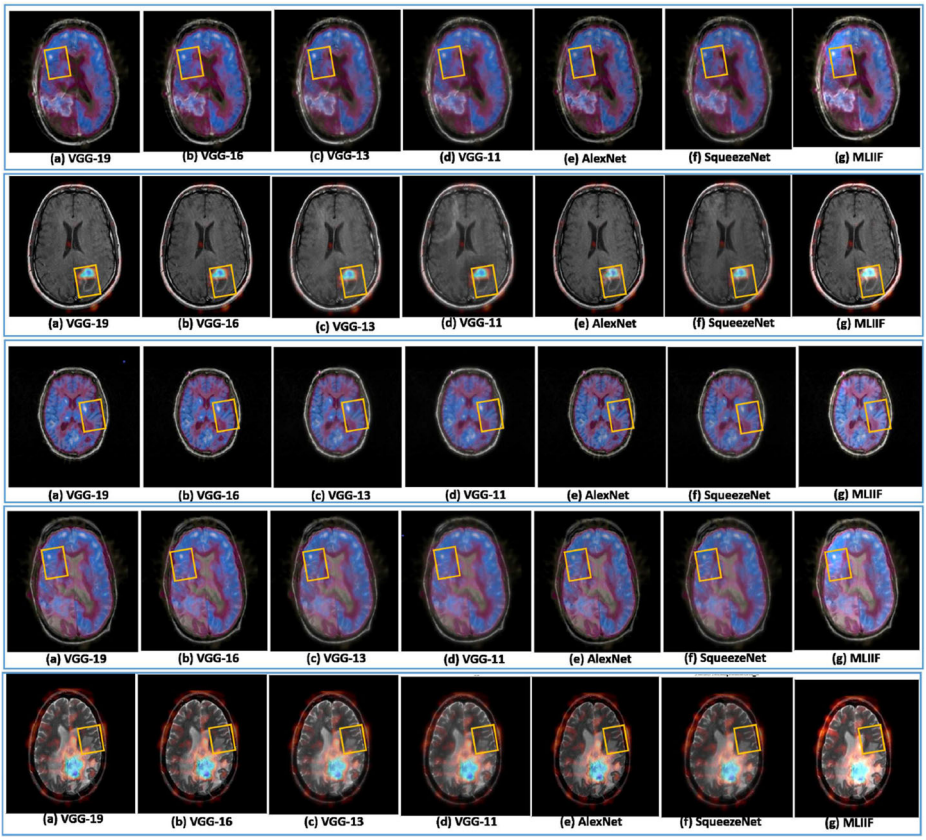

Fig. 9 Fusion results of all the datasets with existing DNN and proposed system

5.5 Evaluation based on objective assessment

The fused image assessment is performed using 11 measurements of the non-reference performance as shown in the Tables 1 and 2 measures for ground truth reference as shown in (13) and (14). The quantitative assessment outcomes of five pairs of image sets are tabulated in Table 5 for existing DNN with MLIIF and Table 6 for published works with MLIIF. In bold letters the significant results are shown in the table. JE and MI values are similar in all the source image pairs. For each pair of sets there are no changes in JE and MI values for the table.

The quantitative assessment analysis of every pair of data for the existing DNN and MLIIF techniques is the following:

– For set 1, the fused image entropy(Ef), $FFSD$, FQI, SF, Q_{hnc} and $ISQE$ are better for the proposed MLIIF as compared to other algorithms. A higher entropy value provides high information content in an image. A higher FF contains very robust information from input images. Sharpe contrast in the image shows a higher SD. A high value SF shows better details and information about the image features.Detailed information is transferred by FQI from the source image to the fused image. Low values in

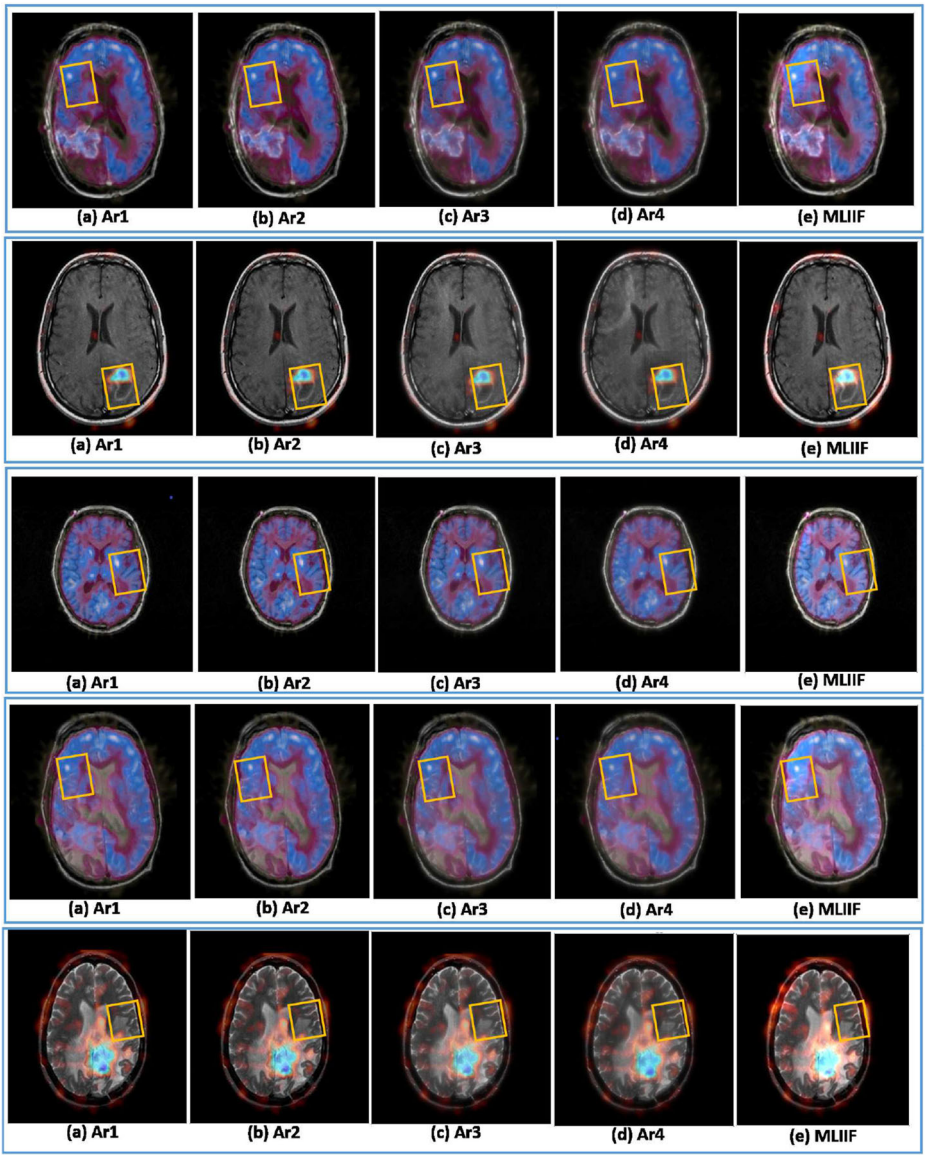

Fig. 10 Fusion results of proposed system and published works for all the sets of images

$ISQE$ provides the highest perception level. Information from source to fused image is transfered by $Qhnc$.

– The results in set 2 are better than other algorithms, for a proposed system, provided by the performances of Ef, SD, FF, $SSIM$, , FQI and FS. The structural information of the final resultant fused image is obtained by $SSIM$ values. Dissimilarity of the fused final image with the source images is indicated by FS.
– Set 3 produces better performance for MLIIF than other algorithms based on Ef, SD, FF, $SSIM$, SF and $Qhnc$ values.

Table 5 Comparison of performance measures with existing DNN and proposed MLIIF for all the five sets of images

Data set	DL N/W	Ef	SD	JE	MI	SSIM	FF	FS	SF	FQI	ISQE	Qhnc
MR-Gad/PET	A1 [4]	4.8893	49.7471	**6.6548**	**0.7424**	0.5145	1.7253	0.1119	8.8553	0.7794	0.4198	–0.1317
	A2 [28]	4.8788	50.3249	**6.6548**	**0.7424**	0.5148	1.7600	0.1140	8.6548	0.7798	0.4213	–0.1319
	A3 [29]	5.0182	48.1938	**6.6548**	**0.7424**	0.5181	1.7590	0.1006	7.5927	0.7691	0.4503	–0.1297
	A4 [11]	5.0780	48.9432	**6.6548**	**0.7424**	0.5204	1.8016	**0.1003**	6.9470	0.7644	0.4625	–0.1291
	A5 [16]	4.7205	45.4120	**6.6548**	**0.7424**	0.5204	1.7135	0.1123	7.2401	0.7856	0.4246	–0.1343
	A6 [10]	4.9814	46.8509	**6.6548**	**0.7424**	**0.5220**	1.7751	**0.1003**	6.4800	0.7686	0.4555	–0.1304
	A7	**5.1831**	**59.4811**	**6.6548**	**0.7424**	0.5075	**2.0231**	0.1152	**8.8747**	**0.9017**	**0.4196**	**–0.1286**
MR-Gad/SPECT-T1	A1 [4]	4.6514	46.9149	**4.4184**	**0.1070**	0.4890	0.9504	0.3460	**16.1083**	0.4734	**0.3982**	–0.1616
	A2 [28]	4.6033	45.5074	**4.4184**	**0.1070**	0.4906	0.9476	0.3428	15.7540	0.4734	0.3985	–0.1629
	A3 [29]	4.8054	47.5223	**4.4184**	**0.1070**	0.4891	0.9424	0.3458	13.3489	0.4756	0.4160	–0.1576
	A4 [11]	4.8133	45.7147	**4.4184**	**0.1070**	0.4915	0.9343	0.3462	11.9222	0.4744	0.4183	**–0.1574**
	A5 [16]	4.5668	43.1589	**4.4184**	**0.1070**	0.4923	0.9627	0.3348	12.9068	0.4748	0.4028	–0.1641
	A6 [10]	4.7105	41.3503	**4.4184**	**0.1070**	0.4974	0.9344	0.3359	11.1058	0.4740	0.4194	-0.1601
	A7	**4.9272**	**51.2925**	**4.4184**	**0.1070**	**0.5054**	**1.1025**	**0.3179**	14.8920	**0.4827**	0.4018	–0.1593
MRT1/PET	A1 [4]	4.0346	44.9115	**4.5147**	**0.7916**	0.6312	1.4174	0.0394	10.4373	0.2894	**0.3316**	–0.1723
	A2 [28]	4.0041	44.7665	**4.5147**	**0.7916**	0.6329	1.4188	0.0408	10.3293	0.2902	0.3323	–0.1731
	A3 [29]	4.3225	44.8279	**4.5147**	**0.7916**	0.6197	1.4362	0.0330	8.9676	0.2829	0.3856	–0.1645
	A4 [11]	4.3232	45.2946	**4.5147**	**0.7916**	0.6265	1.5119	**0.0275**	8.2315	**0.3043**	0.4023	–0.1655
	A5 [16]	3.9895	42.0239	**4.5147**	**0.7916**	0.6318	1.3565	0.0369	8.7309	0.2770	0.3515	–0.1728
	A6 [10]	4.2366	43.5322	**4.5147**	**0.7916**	0.6271	1.4130	0.0300	7.7037	0.2959	0.3812	–0.1668
	A7	**4.3607**	**56.3474**	**4.5147**	**0.7916**	**0.6364**	**1.6642**	0.0355	**10.4393**	0.3000	0.3695	**–0.1643**
MRT2/PET	A1 [4]	4.8337	52.7984	**6.4378**	**0.8580**	0.5259	1.7709	0.1072	8.1032	**0.7929**	**0.4178**	–0.1346
	A2 [28]	4.8264	53.2159	**6.4378**	**0.8580**	0.5265	1.7479	0.1063	7.9461	0.7921	0.4188	–0.1347
	A3 [29]	4.9649	52.098	**6.4378**	**0.8580**	0.5272	1.8028	0.1004	6.9327	0.7835	0.4454	–0.1328

Table 5 (continued)

Data set	DL N/W	Ef	SD	JE	MI	SSIM	FF	FS	SF	FQI	ISQE	Qhnc
	A4 [11]	5.0113	52.2526	**6.4378**	**0.8580**	0.5307	1.8400	0.1028	6.1975	0.7771	0.4566	–0.1323
	A5 [16]	4.7317	50.6491	**6.4378**	**0.8580**	0.5304	1.7212	0.0992	6.9396	0.7952	0.4197	–0.1366
	A6 [10]	4.9500	51.6816	**6.4378**	**0.8580**	0.5299	1.7850	**0.0970**	6.1333	0.7822	0.4498	–0.1333
	A7	**5.1519**	**66.1020**	**6.4378**	**0.8580**	**0.5345**	**2.0712**	0.1211	**8.2190**	0.7665	0.4380	**–0.1306**
MRT2/SPECT	A1 [4]	3.8933	49.4950	**4.9881**	**0.3111**	0.6172	1.2291	0.1371	9.8538	0.8103	0.4731	–0.1723
	A2 [28]	3.8670	48.6857	**4.9881**	**0.3111**	0.6184	1.2286	0.1365	9.6206	**0.8112**	0.4735	–0.1731
	A3 [29]	4.0597	50.1960	**4.9881**	**0.3111**	0.6203	1.2715	0.1314	8.4368	0.8053	0.4801	–0.1683
	A4 [11]	4.0759	48.0914	**4.9881**	**0.3111**	0.6242	1.2950	0.1267	7.4978	0.8031	0.4826	–0.1684
	A5 [16]	3.7031	45.7646	**4.9881**	**0.3111**	0.6238	1.2143	0.1337	7.7944	0.8103	0.4755	–0.1780
	A6 [10]	4.0226	47.7205	**4.9881**	**0.3111**	**0.6258**	1.2752	0.1226	7.1131	0.8017	0.4817	–0.1698
	A7	**5.0472**	**60.5767**	**4.9881**	**0.3111**	0.6106	**1.3717**	**0.0781**	**9.8672**	0.7985	**0.4625**	**–0.1681**

Table 6 Performance comparison between existing other published works with MLIIF proposed for five sets of data

Data	Algorithm	Ef	SD	JE	MI	SSIM	FF	FS	SF	FQI	ISQE	Qhnc
MR-GAD/PET	Ar1 [14]	4.8372	58.1472	**6.6548**	**0.7424**	0.4677	1.9037	0.1204	8.5638	0.6486	**0.3963**	–0.1306
	Ar2 [32]	4.3335	54.5793	**6.6548**	**0.7424**	0.4809	1.8318	**0.1038**	8.7881	0.8211	0.4327	–0.1472
	Ar3 [7]	4.5639	50.7826	**6.6548**	**0.7424**	0.4913	2.0064	0.0995	8.4319	0.7519	0.5609	–0.1603
	Ar4 [8]	5.0128	54.5837	**6.6548**	**0.7424**	0.4728	1.7462	0.1217	8.8205	0.5492	0.5382	–0.1291
	A7	**5.1831**	**59.4811**	**6.6548**	**0.7424**	**0.5075**	**2.0231**	0.1152	**8.8747**	**0.9017**	0.4196	**–0.1286**
MR-GAD/SPECT-T1	Ar1 [14]	4.6532	49.3817	**4.4184**	**0.1070**	0.5102	1.0462	0.3641	13.2705	0.4810	**0.3706**	–0.2063
	Ar2 [32]	4.0619	46.9543	**4.4184**	**0.1070**	0.4918	0.9452	**0.2990**	11.8857	0.4771	0.4851	–0.3097
	Ar3 [7]	3.9392	45.9856	**4.4184**	**0.1070**	**0.5534**	0.8845	0.4739	10.5794	0.4274	0.4158	–0.1689
	Ar4 [8]	3.7857	50.1268	**4.4184**	**0.1070**	0.5493	0.7951	0.4429	13.4851	0.4667	0.4261	–0.1972
	A7	**4.9272**	**51.2925**	**4.4184**	**0.1070**	0.5054	**1.1025**	0.3179	**14.8920**	**0.4827**	0.4018	**–0.1593**
MR-T1/PET	Ar1 [14]	4.0528	54.3312	**4.5147**	**0.7916**	0.6210	1.5407	0.1207	8.9278	0.2905	**0.3092**	–0.2309
	Ar2 [32]	4.1752	50.2490	**4.5147**	**0.7916**	0.6409	0.9827	0.0333	9.9163	0.2275	0.4202	–0.2230
	Ar3 [7]	4.3592	51.4891	**4.5147**	**0.7916**	0.5492	1.5606	**0.0282**	10.3017	0.2837	0.3572	**–0.1304**
	Ar4 [8]	4.3371	54.7094	**4.5147**	**0.7916**	0.5571	1.0817	0.0539	9.8673	0.2673	0.3348	–0.1671
	A7	**4.3607**	**56.3474**	**4.5147**	**0.7916**	**0.6364**	**1.6642**	0.0355	**10.4393**	**0.3000**	0.3695	–0.1643
MR-T2/PET	Ar1 [14]	5.0374	59.8423	**5.5481**	**0.8580**	0.4953	1.9836	**0.0318**	8.2205	0.6380	0.4484	–0.1274
	Ar2 [32]	4.9331	60.5725	**5.7273**	**0.8580**	0.4653	1.7582	0.1312	8.2309	0.7283	**0.2893**	–0.1403
	Ar3 [7]	4.6382	61.3327	**6.2581**	**0.8580**	0.5017	2.0183	0.1464	8.0528	0.6962	0.3871	**–0.1209**
	Ar4 [8]	4.5689	62.0384	**6.3362**	**0.8580**	0.5118	1.8865	0.1195	**8.3072**	0.7453	0.3576	–0.1357
	A7	**5.1519**	**66.1020**	**6.4378**	**0.8580**	**0.5345**	**2.0712**	0.1211	8.2190	**0.7665**	0.4380	–0.1306
MR-T2/SPECT	Ar1 [14]	4.7859	56.7317	**4.9881**	**0.3111**	0.5883	**1.3803**	0.0563	9.8702	0.5781	0.4712	–0.1758
	Ar2 [32]	4.9857	55.6371	**4.9881**	**0.3111**	0.5687	1.3452	**0.0530**	9.6461	0.6618	0.4736	–0.1851
	Ar3 [7]	5.0126	60.2315	**4.9881**	**0.3111**	0.5975	1.3375	0.0705	**9.8715**	0.7246	0.4653	–0.1779
	Ar4 [8]	4.4804	60.3316	**4.9881**	**0.3111**	0.6011	1.3725	0.0698	9.8359	0.7518	**0.4521**	–0.1690
	A7	**5.0472**	**60.5767**	**4.9881**	**0.3111**	**0.6106**	1.3717	0.0781	9.8672	**0.7985**	0.4625	**–0.1681**

- In set 4, $Ef, SSIM, SD, SF, FF$ and $Qhnc$ are shown preferred values for the proposed MLIIF compared to existing algorithms.
- The set 5 provides better values for the MLIIF system than other algorithms in $Ef, FS, SD, SF, FF, Q_{hnc}$and $ISQE$.

Lower RMSE values and higher PSNR values are preferred. It is observed from Fig. 8 that A7 is showing higher PSNR values and lower RMSE for all the sets of images compared to other algorithms.

The quantitative analysis of each pair of data with published works and the proposed MLIIF is:

- It can be observed that fused image entropy, Ef, $SD, SF, SSIM, Qhnc$ and FQI for the MLIIF system proposed is better than the published works for set 1. A higher entropy value provides high information content in an image. Sharpe contrast in the image shows a higher SD. The $SSIM$ provides structural information of the final resultant fused image. A high value of SF deals with information about feature and details of the image sets. The transfer rate of detailed information from the source image to a fused image is shown by FQI. Low value $ISQE$ provides highest perception level. $Qhnc$ transfers information to fused image from the source images.
- The performance measurements of $FF, Ef, SD, SF, FQI, Q_{hnc}$ and FS for the proposed system, MLIIF, provides better results than those published works in set 2 images. A higher FF contains very robust information from input images. Dissimilarity of the fused final image with the source images is indicated by FS.
- In set 3, significant results for $Ef, SD, FF, SSIM, SF$, and FQI for MLIIF than published works.
- Preferred values are shown by set 4 for MLIIF in $Ef, SD, SSIM, FF$, and FQI than published works.
- Set 5 shows better values in $Ef, SD, SSIM, FQI$, and Q_{hnc} for the MLIIF system compared to other published works.

From the objective evaluation mentioned above, most of the performance measures are showing better results for the proposed system MLIIF than other existing algorithms.

The advantages of the proposed system MLIIF over previously available methods are tabulated in Table 7.

A typical way to evaluate the performance of an algorithm is by computing the theoretical computational complexity which can often be used as proxies for performance. Figure 7 displays the overall time of the existing DNN and MLIIF. On the x-axis, existing DNN model, and MLIIF, the y-axis represents the computational complexity in millions of operations. The accuracy obtained by the algorithm is more important than complexity in the case of medical image fusion (Table 8 and Fig. 11).

The entropy value comparison of all the sets of images is illustrated in Fig. 12. It can be noted that the fused image entropy is higher than the source images for all the pairs and algorithms. It is notable that for all the pairs and algorithms the fused image entropy is higher than the source images. Better details or information are provided by the fused image compared to individual source images. From quantitative evaluation, it is observed that most of the values are showing prominent for the proposed MLIIF system. On the basis of all objective and subjective assessments, the MLIIF system is robust and appears to provide more information than existing DNN for all the image pair sets.

Table 7 Advantages of the proposed system MLIIF over previously available methods

Sl.No	Existing algorithm	MLIIF
1	Pre-trained existing VGG-11 [11], VGG-13 [29], VGG-16 [28], VGG-19 [4], Squeezenet [10] and Alexnet [16]	The hybrid MLIIF can provide more features and textured information than single DNN's
2	[Ar1] Fractional wavelet transform consumes time and only single modality pair is experienced [14]	Faster in computation and worked with different modality pairs. More performance measures are utilized
3	[Ar2] edge-preserving filter based decomposition is applied for the extraction of features [32]	Multiple fusion of the same pair provides more features and textures from the source image to fused image
4	[Ar3] CNN is applied only for the HF components which are obtained by the NSST decomposition. This work is time-consuming and experienced with only CT/MR pairs [7]	Faster in computation and worked with different modality pairs
5	[Ar4] CNN is applied only for the low-frequency components which are obtained by the NSST decomposition. This work is time-consuming and experienced with only CT/MR pairs [8]	Faster in computation and worked with different modality pairs

6 Conclusion

In diagnosis and treatment the major challenge is to create a visually enhanced medical image as medical images are a significant source of information for the ailment analysis. The fusion of medical images resolves the challenges by providing all the relevant information in a single image compared to the source images that are acquired from dissimilar sensors. This research work proposes a novel fusion system, MLIIF, with Multi-layer, multi-tier fusion on deep learning networks for multi-modal medical images that optimally decompose high and low-frequency components that are fused with different deep learning networks in each layer. Deep feature based multilayer fusion strategy helps to obtain more

Table 8 Comparitive analysis of the execution time of existing and proposed DL model & accuracy details of the models

DL N/W	Execution Time (Second)					Overall Accuracy (%)
	Set 1	Set 2	Set 3	Set 4	Set 5	
A1 [4]	**0.3209**	**0.3320**	**0.3292**	**0.3203**	**0.3197**	89.08
A2 [28]	0.8812	0.8772	0.8821	0.8728	0.8901	91.73
A3 [29]	0.9463	0.9367	0.9436	0.9436	0.9508	92.71
A4 [11]	0.8377	0.8720	0.8737	0.8368	0.8295	90.22
A5 [16]	0.8216	0.8319	0.8261	0.8126	0.8197	91.48
A6 [10]	0.6283	0.6309	0.6238	0.6328	0.5992	90.89
A7	0.9572	0.9534	0.9527	0.9505	0.9510	**95.09**

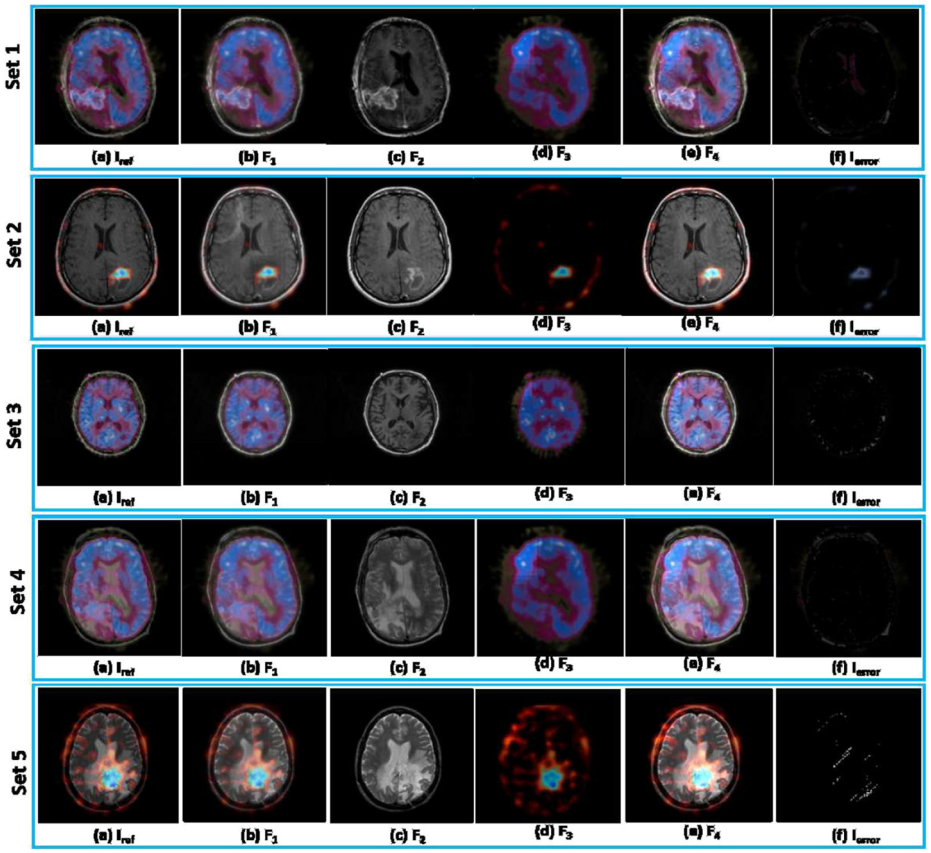

Fig. 11 Fusion results of proposed system

informative fused image from the source image sets. The deep features of the low frequency and high-frequency contents are obtained by VGG-19, VGG-11, and Squeezenet network. This helps in improved information and features from the coefficients to the final fused image. The robustness and effectiveness of the proposed multi-layer, multi-tier fusion system are validated by subjective and objective analysis, and the effectiveness of the proposed MLIIF system is evaluated by the error image calculation and accuracy from the ground truth image. The error image from the resultant fused image and the ground truth image calculated provides a negligence difference. The proposed MLIIF system obtained an accuracy of 95.09%. The experimental findings indicate that the system proposed is state-of-the-art even if computationally time-consuming.

7 Future scope

The work can be further applied with other sets and large data sets of medical images for multi-modal medical image fusion, multi-exposure image fusion and multi-focus image fusion.

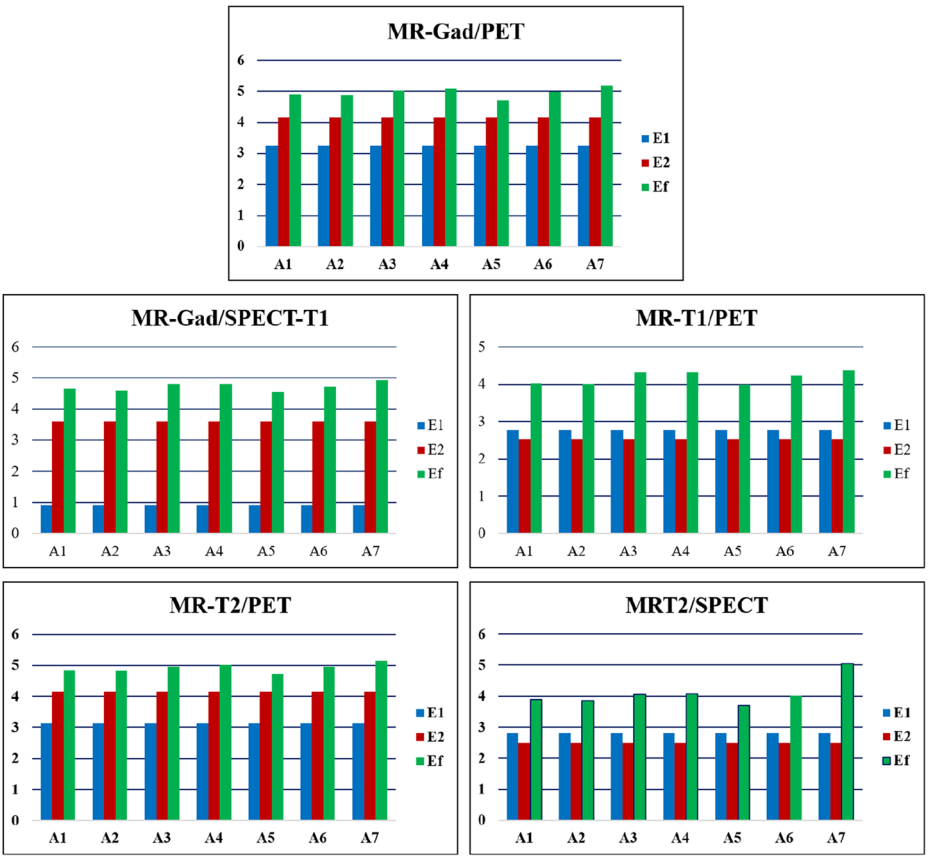

Fig. 12 Entropy comparison of source images and fused image for 5 sets of data

Acknowledgments The success of any research work is infinite, trustworthy and suitable information. HCG, Cancer research hospital, Bangalore is appreciated for its subjective assessment contributions. Appreciation is given to the valuable contribution by Dr. Ravi Nayar, ENT Surgeon and the dean, Dr. Shiv Kumar, head of radiology and the HCG team. For the valued advice and technical suggestions at all stages of our work, we thank Dr. VPS Naidu, Principal Scientist, and Assoc. Prof. (AcSIR), National Aerospace Laboratories (NAL).

Funding The authors did not receive support from any organization for the submitted work.

Declarations

Conflict of Interests The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. A. T, Parameswaran L (2013) A computationally efficient edge preserving mri-ct image fusion technique using complex wavelet transform and phase congruency fusion rule. Europ J Sci Res 112(4):469–483

2. Calhoun VD, Adali T (2008) Feature-based fusion of medical imaging data. IEEE Trans Inf Technol Biomed 13(5):711–720
3. Dasarathy BV (2012) Information fusion in the realm of medical applications-a bibliographic glimpse at its growing appeal. Inform Fus 13(1):1–9
4. Garcia-Gasulla D, Parés F, Vilalta A, Moreno J, Ayguadé E, Labarta J, Cortés U, Suzumura T (2018) On the behavior of convolutional nets for feature extraction. J Artif Intell Res 61:563–592
5. Gatys LA, Ecker AS, Bethge M (2016) Image style transfer using convolutional neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2414–2423
6. Haritha KC, Jeyakumar G, Thangavelu S (2017) Image fusion using evolutionary algorithms: a survey. In: 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, pp 1–7
7. Hermessi H, Mourali O, Zagrouba E (2018) Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain. Neural Comput Applic 30(7):2029–2045
8. Hou R, Zhou D, Nie R, Liu D, Ruan X (2019) Brain ct and mri medical image fusion using convolutional neural networks and a dual-channel spiking cortical model. Med Biol Eng Comput 57(4):887–900
9. Huang X, Belongie S (2017) Arbitrary style transfer in real-time with adaptive instance normalization. In: Proceedings of the IEEE International Conference On Computer Vision, pp 1501–1510
10. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size. arXiv:1602.07360
11. Iglovikov V, Shvets A (2018) Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation. arXiv:1801.05746
12. James AP, Dasarathy BV (2014) Medical image fusion: a survey of the state of the art. Inform Fus 19:4–19
13. Johnson. KA, Becker JA (2021) The whole brain atlas. https://www.med.harvard.edu/aanlib/home.html. Accessed 30 Dec 2018
14. Kaur J, Shekhar C (2020) Multimodal medical image fusion using deep learning. In: Advances in computational techniques for biomedical image analysis. Elsevier, pp 35–56
15. Krishnamoorthy S, Soman KP (2010) Implementation and comparative study of image fusion algorithms. Int J Comput Applic 9(2):25–35
16. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
17. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324
18. Li H, Wu X-J, Kittler J (2018) Infrared and visible image fusion using a deep learning framework. In: 2018 24th International conference on pattern recognition (ICPR). IEEE, pp 2705–2710
19. Li S, Kang X, Hu J (2013) Image fusion with guided filtering. IEEE Trans Image Process 22(7):2864–2875
20. Ma J, Zhou Z, Wang B, Zong H (2017) Infrared and visible image fusion based on visual saliency map and weighted least square optimization. Infrared Phys Technol 82:8–17
21. Moushmi S, Sowmya V, Soman KP (2016) Empirical wavelet transform for multifocus image fusion. In: Proceedings of the international conference on soft computing systems. Springer, pp 257–263
22. Nair RR, Singh T (2019) Multi-sensor medical image fusion using pyramid-based dwt: a multi-resolution approach. IET Image Process 13(9):1447–1459
23. Nair RR, Singh T (2020) Multi-modal based msmif using hybrid fusion with 1-d wavelet transform. IJAST 29(5):5353–5368
24. Nair RR, Singh T (2021) Mamif: multimodal adaptive medical image fusion based on b-spline registration and non-subsampled shearlet transform. Multimed Tools Appl 80(12):19079–19105
25. Nair RR, Singh T (2021) An optimal registration on shearlet domain with novel weighted energy fusion for multi-modal medical images. Optik 225:165742
26. Nair S, Elias B, Naidu VPS (2007) Pixel level image fusion using fuzzylet fusion algorithm. IJAREEIE An ISO, 3297
27. Parvathy VS, Pothiraj S, Sampson J (2020) Optimal deep neural network model based multimodality fused medical image classification. Phys Commun, 101119
28. Qassim H, Feinzimer D, Verma A (2017) Residual squeeze vgg16. arXiv:1705.03004
29. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556
30. Wu F, Jing X-Y, Dong X, Hu R, Yue D, Wang L, Ji Y-M, Wang R, Chen G (2018) Intraspectrum discrimination and interspectrum correlation analysis deep network for multispectral face recognition. IEEE Trans Cybern 50(3):1009–1022

31. Wu F, Jing X-Y, Feng Y, Ji Y, Wang R (2021) Spectrum-aware discriminative deep feature learning for multi-spectral face recognition. Pattern Recogn 111:107632
32. Zhang Y-D, Dong Z, Wang S-H, Yu X, Yao X, Zhou Q, Hu H, Li M, Jiménez-Mesa C, Ramirez J et al (2020) Advances in multimodal data fusion in neuroimaging: overview, challenges, and novel orientation. Inform Fus 64:149–187
33. Zhao Y, Yin Y, Fu D (2008) Decision-level fusion of infrared and visible images for face recognition. In: 2008 Chinese control and decision conference. IEEE, pp 2411–2414