ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

Brain Tumor Detection through Image Fusion Using Cross Guided Filter and Convolutional Neural Network

Srikanth M.V.[1], Sivalal Kethavath[2], Srinivas Yerram[3], SivaNagiReddy Kalli[4], Nagasirisha.B[5] and Jatothu Brahmaiah Naik[6]

ABSTRACT

This Data fusion has become a significant issue in diagnostic imaging, particularly in medical applications like radiation and guided image surgery. Medical image fusion aims to enhance the precision of tumor diagnosis, by preserving the salient information and characteristics of the original images in the fused image. It has been shown that guided filters are capable of maintaining edges well. In this paper, we propose a novel cross-guided filter-based fusion approach for multimodal medical images utilizing convolutional neural networks. The cross-guided filter is used in the proposed algorithm to extract the detailed features from the source images. Convolutional neural networks are used to generate the feature weights of source images derived from the detail layers. The weighted average rule is used to merge the source images based on these weights. We used thirty distinct types of medical images from diverse sources to compare the effectiveness of the proposed strategy to that of existing methods, both numerically and visually. The experimental findings demonstrated that, in terms of both objective evaluation and qualitative image quality, the suggested system performs better than other standard methods already in use. The quantitative results show that compared to existing methods under consideration for comparison, the proposed algorithm improves mutual information by 25%, image entropy by 9.5%, spatial frequency by 21%, standard deviation by 18.1%, structural similarity index by 30%, and edge strength of the fused image by 39%.

Article information:

Keywords: Medical Image Fusion, Cross-guided Filter, Convolutional Neural Networks, Mutual Information, Image Entropy, Spatial Frequency, Standard Deviation, Structural Similarity Index, Edge Strength

Article history:
Received: May 5, 2024
Revised: September 19, 2024
Accepted: October 3, 2024
Published: October 19, 2024 (Online)

DOI: 10.37936/ecti-cit.2024184.256650

1. INTRODUCTION

Medical imaging is essential to many patients' medical experiences and provides a solid foundation for clinical decision-making. Many clinical applications, including computer-assisted detection, therapy, action planning, diagnosis, and planning, can benefit from medical imaging. While medical imaging systems play a significant role in many clinical tasks, there is an increasing need for trustworthy automated ways to assist healthcare personnel in interpreting complex medical pictures. For the purpose of interpreting structured data, such images, medical imaging research thus benefits from the development of advanced computer tools [1,2]. New technologies for capturing, processing, and interpreting images are driving innovation, particularly in the fields of registration, segmentation, reconstruction, fusion, detection, modeling, and tracking. Understanding medical imagery ultimately requires prior knowledge and might be challenging at times. Biomedical pictures can have noise and a variety of modality-specific artifacts depending on the techniques and settings used during acquisition. The topic of healthcare image

[1]The author is with the Usha Rama College of Engineering and Technology, Andhra Pradesh,India, E-mail: sree.02476@gmail.com
[2]The author is with the Rajiv Gandhi University of Knowledge Technologies, Nuzvid, Andhra Pradesh, India. E-mail: shiv.rathod@rguktn.ac.in
[3]The author is with the Vignana Bharathi Institute of Technology,Hyderabad, India, E-mail: srinivasyerram06@gmail.com
[4]The author is with the Sridevi Women's Engineering College, Hyderabad, India, E-mail: sivanagireddykalli@gmail.com
[5]The author is with the Gudlavalleru Engineering College, Andhra Pradesh, India, E-mail: nagasirishab@gmail.com
[6]The author is with the Narasaraopeta Engineering College, Andhra Pradesh, India, E-mail: brahmaiahnaik@gmail.com
[1]Corresponding author: sree.02476@gmail.com

processing has seen a growth in study over the past 20 years. At first, it focused on standard image processing tasks including registration, contrast enhancement, and segmentation. Nevertheless, imaging biomarker identification has focused on converting functional data into pertinent biomarkers that can provide information about various medical disorders as medical image processing has developed [3-6].

Image fusion combines information from multiple multimodal shots of the same scene to create a single fused image that is much more informative than the individual multimodal images. This work focuses on multimodality medical image fusion, which is the fusion of different imaging modalities used to fuse multimodal medical pictures from the same region of the human body. The medical images created using these approaches include features that depict the current state of the human body, such as metabolic rate and bone architecture. However, a single image might only depict one type of human body. For instance, high spatial resolution anatomical information is provided by MRI imaging. Magnetic resonance imaging (MRI) can only display soft tissues and cannot show bone information; in contrast, computed tomography (CT) scans can more efficiently distinguish between tissues with variable densities, such as blood vessels and complex structures. Poor spatial resolution positron emission tomography (PET) pictures and single photon emission computed tomography (SPECT) images display similar patterns. The notable differences in images produced by various imaging modalities complement one another. Fusing multiple images from different modalities can result in a single fused image that provides a multitude of information, enhances image quality, and dramatically increases image spatial resolution.

Radiologists can better treat a patient's illness after using this composite image to provide a thorough diagnosis. In conclusion, the diagnosis and treatment of all diseases depend on integrating many sets of medical images. Discrete Wavelet Transformation (DWT), Curvelet Transformation (CVT), Non-Subsampled Contourlet Transformation (NSCT), and other techniques are examples of conventional image fusion techniques [9, 10]. These methods are not well suited for combining specific input image information without producing artifacts in the images. Deep learning approaches have greatly succeeded in various computer vision applications due to recent advancements in machine learning. Deep learning research has been increasingly important in image fusion in recent years [11].

This research proposes a multimodal image fusion technique that uses a cross-guided filter to extract high- frequency information from the source images. The feature weights produced by convolutional neural networks are then fused to enhance the fusion performance of medical images further.

2. LITERATURE REVIEW

Many image fusion methods have been proposed in the last few decades. For integration of medical images, Agarwal *et al.* [12] have proposed a hybrid approach that combines wavelet transform (WT) and curvelet transform (CVT). The segmented input images are then fused into sub-bands using the WT approach with the CVT, which divides the bands into overlapping tiles and effectively converts curves in images to straight lines. To provide a more comprehensive fusion result, these tiles are blended using the inverse wavelet transform. The results demonstrate that the fusion outcome produces higher-quality results with fewer errors. A sophisticated hybrid technique for the fusing of multi-focus images has been developed. It combines principal component analysis (PCA) with stationary wavelet transform (SWT)[13]. By splitting the final image into four sub-bands, SWT helps create features from both the source and resultant images. The greatest eigenvector for each of these sub-bands is determined using PCA-based fusion algorithms, and eigenvectors are preserved for the purpose of optimal image representation. By improving image information and assisting in removing artifacts, this strategy may improve visual perception, according to the evaluation criteria.

Bavirisetti & Dhuli [14] provided an edge-preserving fusion method for visible & infrared sensor images with an objective of greatly enhancing image quality and minimizing artifacts in fused images. The source images are first anisotropically diffused into layers of approximation and detail. After that, the Karhunen-Loeve transform and linear superposition are used to calculate the final approximation and detail layers. The final approximation and detail layer images are combined linearly to accomplish fusion. This technique dramatically increases contrast while preserving the most critical aspects of the image. A multiscale fusion methodology based on weighted least squares (WLS) optimization and visual saliency map (VSM) was also developed to solve some of the shortcomings of conventional fusion approaches [15]. This technique divides the input image into basic and detail layers using a multiscale decomposition that combines a Gaussian filter with rolling guidance filtering (RGF). It can accomplish the unique capacity to lessen haloes around edges while preserving the exact scale attributes. This method also yields fusion result details that look more natural and are more appropriate for human visualization. For a range of multimodal pictures, a framework for image fusion with quick spatial filtering has been proposed [16]. The gradient's magnitude is first used to assess an image's sharpness and contrast. Next, utilizing the image gradient magnitude, a brief morphological closing operation is performed to close gaps and fill in areas. The gradient magnitude of the multiple images is converted to a weight map using a fast structure-preserving filter. Finally, a weighed-sum rule is used to generate the fusion result. This method helps make multimodal images appear more realistic.

Numerous types of coupled fusion image techniques, such as coupled image factorization optimization & the modified flexible coupled technique, are proposed by Lu *et al.* [17] about coupled matrix & tensor factorization optimization & flexible coupling technique, respectively. According to experimental findings, the CIF-OPT approach's performance improves when different types of noise are present. In particular, the CIF-OPT approach allows for accurate image reconstruction without sacrificing crucial features. The drawbacks of earlier techniques include color distortion, blurring, and noise. Laplacian redecomposition (LRD) technique is proposed for multimodality medical image fusion [18] to tackle these problems. This approach combines two technological developments. This approach originally suggested a Laplacian decision graph decomposition strategy with picture augmentation to provide complementary details, redundant details, and low frequency sub-band images. To account for the diverse properties of redundant and complimentary information, this method then introduces the principle of overlapping as well as non-overlapping domain. To improve target recognition accuracy while laying the groundwork for clinical applications, Goyal *et al.* [19] recently proposed a multimodality medical image fusion approach that integrates low-resolution multimodal medical pictures with minimal processing complexity. Jose *et al.*'s [20] unique multimodal approach for the Non-Subsampled Shearlet transform (NSST) of images linked to teenage identification search is proposed. Generally, NSST is a multi-scale, multidirectional, multidimensional, and multi-scale wavelet transform. Using the NSCT domain, Kaur and Singh [21] present a medical image fusion that divides images into sub-bands. The features from the input photographs are then extracted using the most advanced version of inception. Through multi-objective differential evolution, this approach makes decisions. The fundamental fusion function for the coefficients of determination, as well as energy loss, is then utilized to produce the fused coefficients. Finally, the fusion result is obtained using an inverse NSCT. Srikanth *et al.* [7,8, 22-23] presented brain tumor identification through image fusion by using metaheuristic algorithms. This article presents a mechanism of image fusion using cross-guided filters and convolutional neural networks.

1. Detail layers of each source image are extracted from its blurred version obtained through cross-guided filter
2. Weight maps for detail images are generated from convolutional neural networks based on feature relevance.
3. Based on the weights of detail layers, source images are fused using the weighted average combination rule.
4. The performance of the proposed mechanism is tested with all types of multimodal sensor datasets subjectively & objectively.

3. CONCEPTS OF PROPOSED METHODS

3.1 Cross-Guided Image Filter

An explicit image filter called a "guided image filter" (GIF) calculates an image's output for each pixel by considering the statistics of that pixel's surrounding neighborhood [24,25]. Between the output and the guidance images, there is a local linear model. Edge information is preserved while the input image is smoothed with this filter. For a guidance image P centered at pixel 'k' in a square window w_k, the filter output F at a pixel 'i' is given as follows:

$$F_i = m_k P_i + n_k, \forall i \in w_k \tag{1}$$

Where m_k and n_k are the coefficients which are linear in the window m_k that are used to minimize the objective function as follows:

$$E(m_k, n_k) = \sum\nolimits_{i \in w_k} \left((m_k P_i + n_k - I_i)^2 + \varepsilon m_k^2\right) \tag{2}$$

Where ϵ is the regularization control parameter. The optimal solution of equation (2) is given by the following values of m_k and n_k

$$m_k = \frac{\frac{1}{|w|}\sum_{i \in w_k} P_i I_i - \mu_k E[I_k]}{\sigma_k^2 + \varepsilon} \tag{3}$$

$$n_k = E[I_k] - m_k \mu_k \tag{4}$$

Where $|w|$ indicates the total count of pixels in w_k and μ_k, σ_k^2 are the mean and variance in w_k, $E[I_k]$ is the expectation of I in wk. After obtaining the linear coefficients, the output F_i can be solved using Equation (1); however, distinct overlapping windows w_k centered at k share pixel i. Take the average of all F_i estimates to address this issue, as given in equation (5), which is the filtered output.

$$F_i = \bar{m}_i P_i + \bar{n}_i \tag{5}$$

If the guidance for the first source image is generated from the second image, and vice versa, a cross-guided image filter is created. Cross guidance causes the filter output to thoroughly smooth the filter inputs since guided filters can retain guidance details in their output. It means that when the filter output is subtracted from the source images, rich detail information is produced.

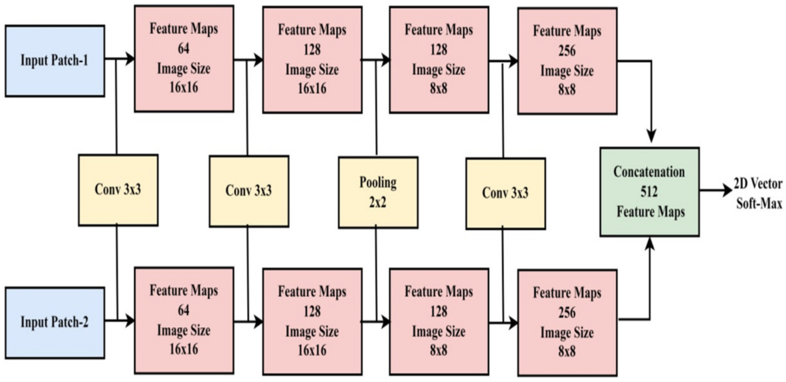

Fig.1: *Schematic representation of a Siamese network of CNN.*

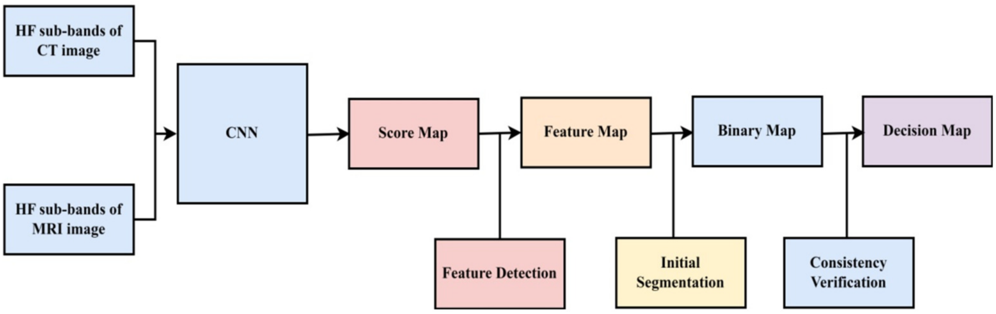

Fig.2: *Decision map computation from high-frequency layers.*

3.2 Convolutional Neural Network (CNN)

As seen in Figure 1, the CNN is used in this work to integrate source images based on the feature map of detail layers. The high frequency subbands of both images are received by two different CNN branches that mimic weights and structures. These kinds of networks are called Siamese networks. Using this network primarily aims to make classification easier because it trains two images simultaneously, allowing it to distinguish between the high-frequency subbands of the two source images. Figure 1 shows the schematic representation of the two branches of the Siamese network [26, 27]. Each branch comprises one max-pooling layer, three convolution layers (64, 128, and 256 filters, respectively), and one filter. While the max-pooling comprises 2 × 2 with a step of 2, each convolution layer has a 3 ⋆ 3 filter responsive with a stride of 1. We have reduced the number of fully linked layers in our method to reduce computation time and memory use. Rather than feeding only fixed-size images, it can also be advantageous to take images of any size. Following concatenation, a 2D vector soft-max layer receives the 512 feature maps and generates a probability distribution for each class. Stochastic gradient descent (SGD) is used for optimization, and its soft-max loss function is employed. Furthermore, the batch size is 128, the momentum is 0.9, and the weight decay is 0.00001.

Three steps make up the CNN decision map generation process from high-frequency images: feature recognition, initial segmentation & consistency checking, as shown in figure 2. A discussion on each of these steps is presented in the subsequent sections.

(*a*) *Feature recognition:* To create a score map, the high -frequency subbands of both images are supplied to two different network branches separately. The feature properties are indicated by each coefficient that appears in the score map. Each coefficient weights between 0 and 1, signifying the two 16 × 16 equivalent blocks from the two high-pass sub-images.

As a result, the score map's overlapping sections are averaged to create a feature map M(p,q) that is comparable in size.

(*b*) *Initial segmentation:* The binary map B(p,q) must be obtained next, following the acquisition of the feature map M. To recover more valuable features for creating binary maps, the choose-max selection approach with an acceptable threshold of 0.5 is applied. Equation (9) is used to calculate it.

$$B(p,q) = \begin{cases} 1 & if \quad M(p,q) \geq 0.5 \\ 0 & if \quad M(p,q) < 0.5 \end{cases} \tag{6}$$

The following step involves processing the binary map B(x,y) obtained by verifying consistency.

(*c*) *Consistency checking:* Singularity points are instances in the binary map where specific misclassified pixels have values that are not near to those of the surrounding pixels. We used an

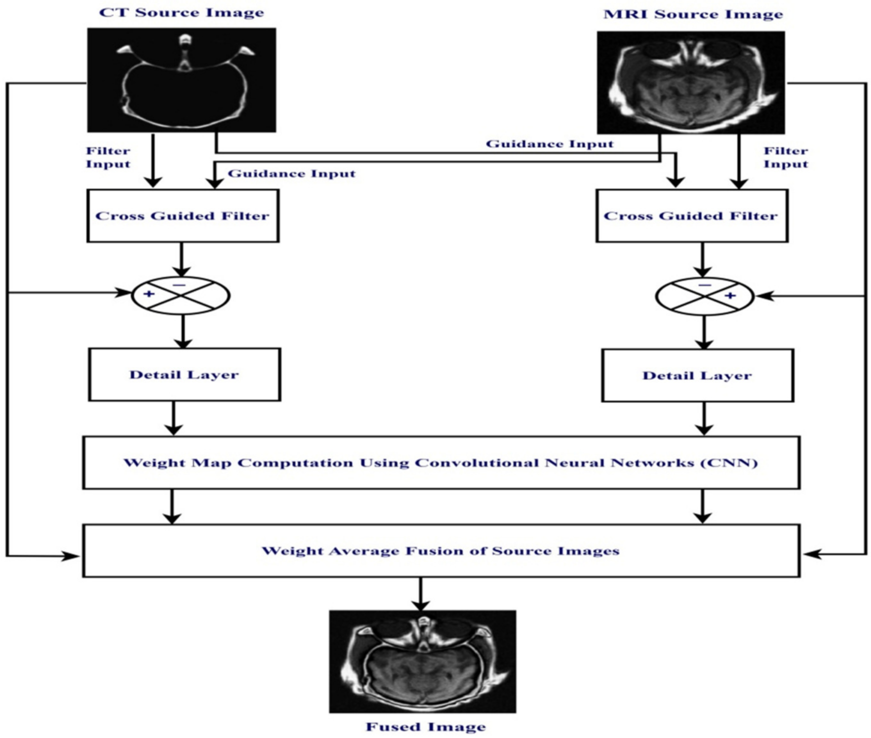

Fig.3: *Proposed flow diagram of image fusion.*

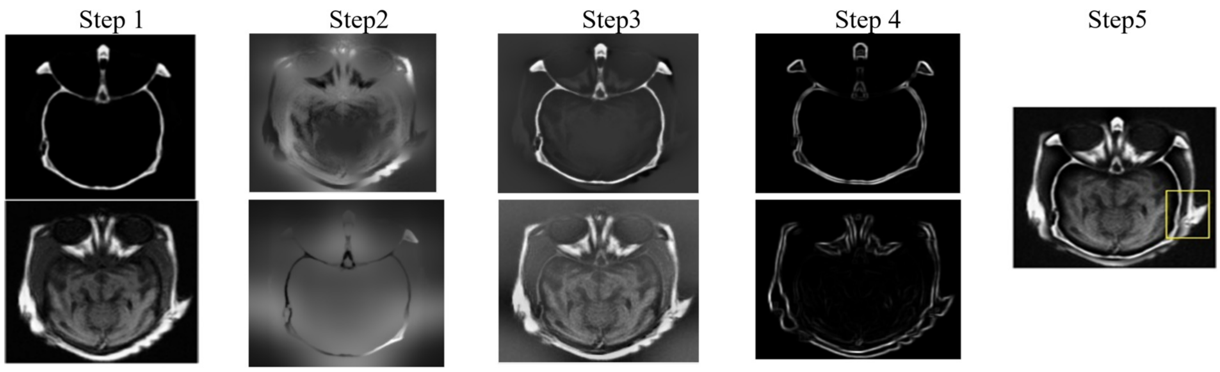

Fig.4: *Illustration of step-by-step results of the proposed method.*

8×8 window consistency check in our technique to eliminate the singularity spots from the focus map. In addition to the singularity point problem, the decision map produces artifacts in the fused image. The guided filtering technique, which preserves edges, is utilized to tackle the issue of artifacts. It is necessary to assign the size of the window 'r' & normalization parameter 'ε', which are selected as 5 and 0.1, respectively.

4. PROPOSED METHODOLOGY

The flow diagram of our proposed fusion technique is presented in figure 3. The step-by-step procedure of the method is presented below:

Step 1: Read the source images to be fused and label them as I_1 and I_2.

Step 2: Apply cross-guided image filtering on each source image to produce their base layers I_{B1} and I_{B2}.

Step 3: Generate high-frequency layers with rich details by subtracting base layers from the source images,

$$I_{D1} = I_1 - I_{B1} \tag{7}$$

$$I_{D2} = I_2 - I_{B2} \tag{8}$$

Step 4: Apply the detail layers to CNN discussed in section 3.2 to produce decision maps D_1 and D_2

Step 5: Combine the source images using decision maps D_1 and D_2 to produce the fused image F.

$$F(p,q) = \frac{(D_1(p,q) + I_1(p,q) + D_2(p,q) + I_2(p,q))}{(D_1(p,q) + D_2(p,q))} \tag{9}$$

The step-by-step results of the proposed method are illustrated in figure 4.

5. SUBJECTIVE ANALYSIS

Five sets of brain-related MRI and CT scans (called “Dataset-1,” “Dataset-2,” “Dataset-3,” “Dataset-4,” and “Dataset-5”) are selected to demonstrate the effectiveness of the proposed fusion procedure. A healthy patient’s brain is shown in Dataset-1; a patient who had a fatal stroke is shown in Dataset-2;

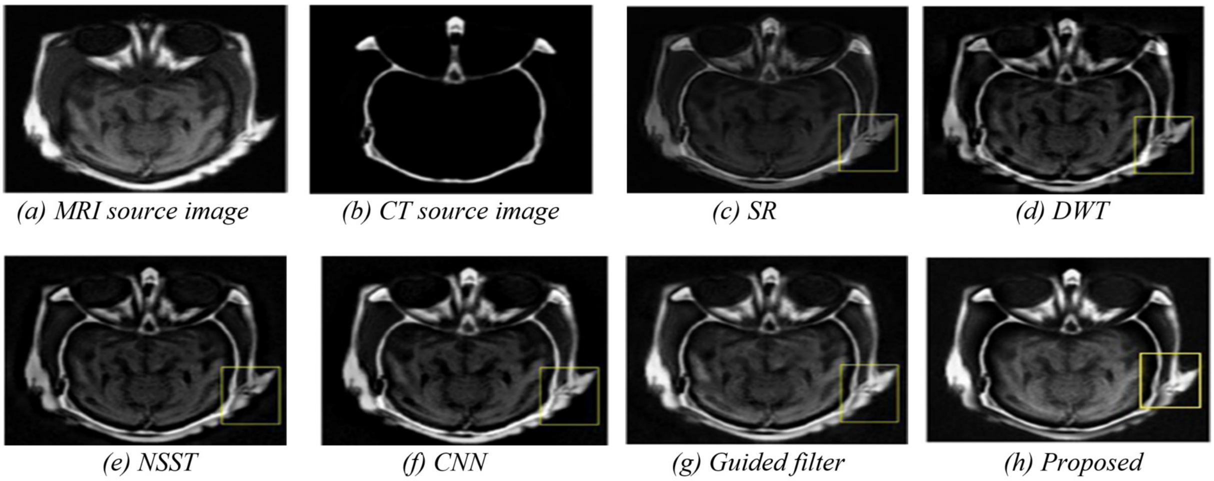

Fig.5: *Fusion results of CT-MRI of healthy brain.*

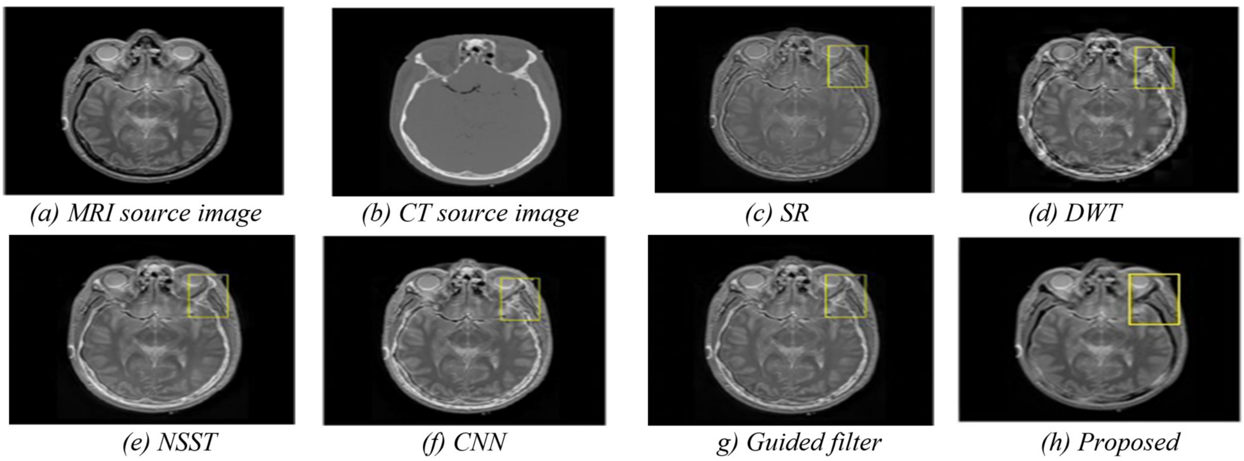

Fig.6: *Fusion results of CT-MRI of Fatal stroke.*

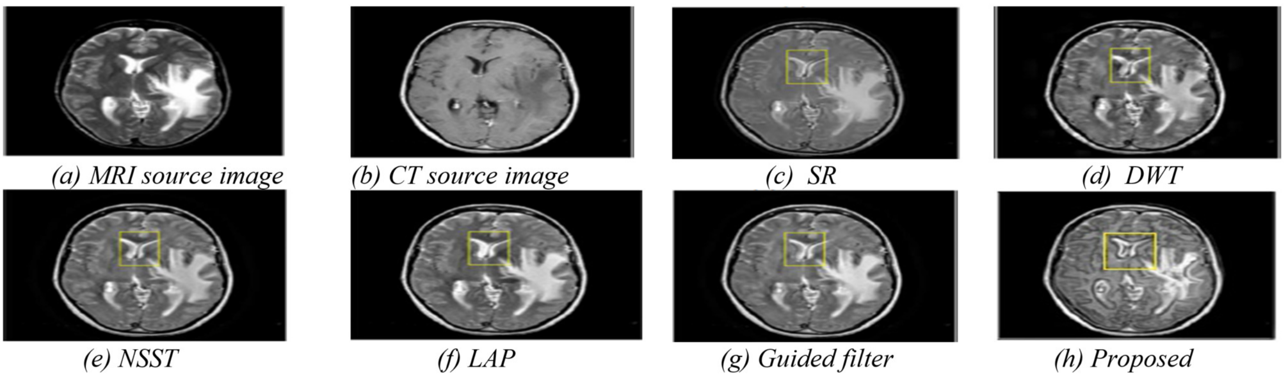

Fig.7: *Fusion results of CT-MRI of neoplastic tumor.*

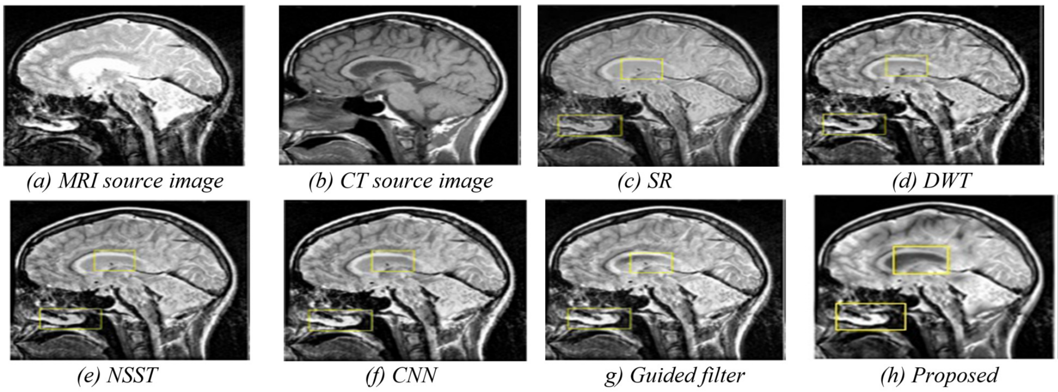

Fig.8: *Fusion results of CT-MRI of brain skull.*

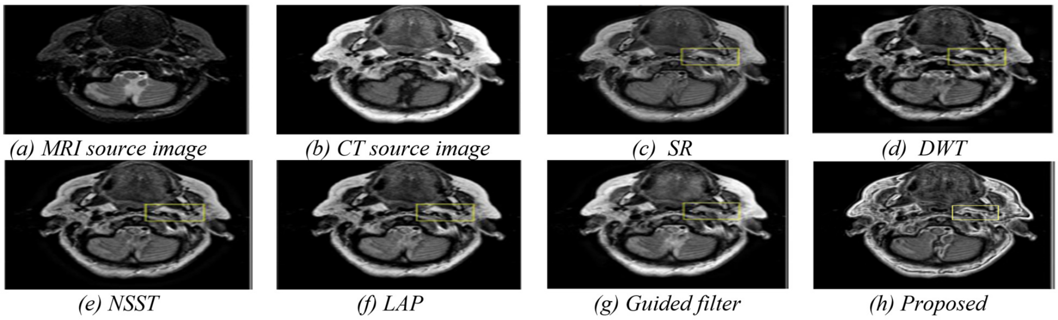

Fig.9: *Fusion results of CT-MRI of Cerebella metastasis.*

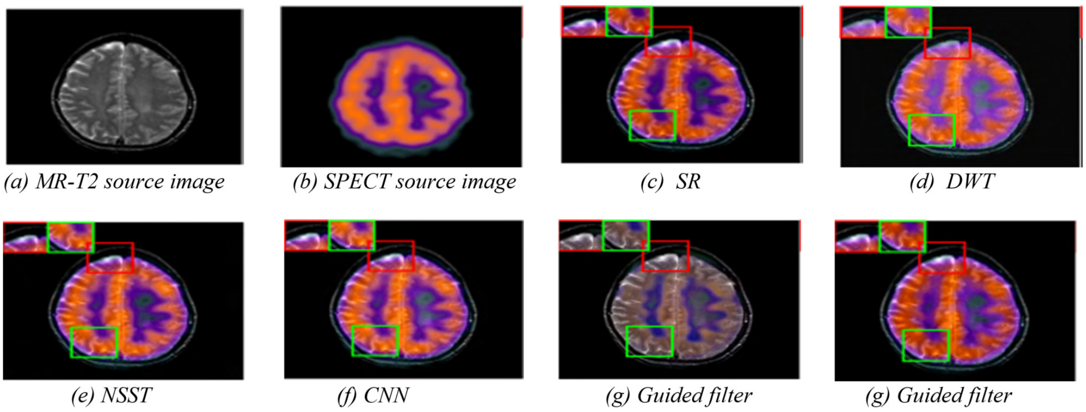

Fig.10: *Fusion results of MR-T2 and SPECT of Metastatic Bronchogenic.*

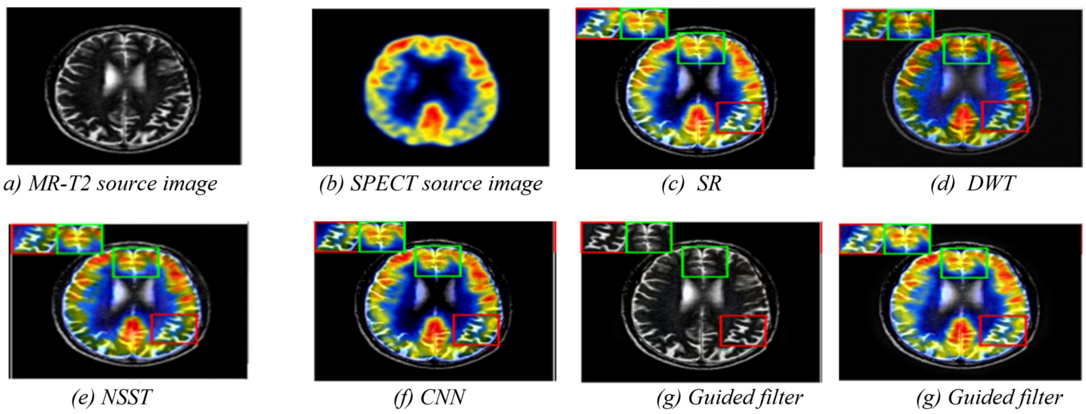

Fig.11: *Fusion results of MR-T2 and PET of Alzheimer's disease.*

a patient who had a neoplastic tumor is shown in Dataset-3; sagittal plane images of the brain skull are shown in Dataset-4; and a patient who had a Cerebella metastasis is shown in Dataset-5. Moreover, SPECT and MR-T2 datasets relevant to Metastatic Bronchogenic disease and MR-T2 and PET datasets associated with Alzheimer's disease are considered. Each one is 256 by 256 in size and features 256 different shades of gray. The Benchmark Brain Atlas, available at `http://www.med.harvard.edu/aanlib/home.html`, was the source of all the datasets. [28]. The Dataset-1 image fusion results utilizing various techniques are shown in Fig. 5. Images of the MRI and CT are shown in Figs. 5(a) and (b), respectively. The fusion results from guided filter approaches, the suggested method, Convolution Neural Network (CNN), Non-Subsampled Shearlet Transformation (NSST), Sparse Representation (SR), and Discrete Wavelet Transformation (DWT) are displayed in Figs. 5(c)–(h). Together with the bone structures from the CT scan, the soft tissues from the MRI image are mostly maintained in the fused results. How the techniques preserve detail and contrast, however, differ only a little. A yellow rectangle

Table 1: *Statistical measures of the proposed method for multimodal datasets.*

Dataset type	Method	Mutual Information (MI)	Image Entropy (H)	Spatial Frequency (SF)	Standard Deviation (SD)	MSSIM	Edge Strength ($Q_{AB/F}$)
Dataset1	SR	2.57	5.80	11.68	30.82	0.5122	0.5756
	DWT	1.92	6.17	17.13	44.71	0.5246	0.6073
	NSST	2.05	6.20	17.05	44.16	0.5366	0.6816
	CNN	2.43	6.07	17.40	52.89	0.5518	0.7184
	Guided filter	2.31	6.52	16.97	52.89	0.5634	0.7210
	Proposed	**3.93**	**6.66**	**18.16**	**54.28**	**0.9908**	**0.8736**
Dataset2	SR	3.42	4.94	17.76	51.40	0.8248	0.5178
	DWT	3.19	5.19	**22.01**	55.73	0.7915	0.5051
	NSST	3.34	5.12	20.95	54.56	0.8160	0.5887
	CNN	3.34	4.89	21.93	59.92	0.8146	0.5888
	Guided filter	3.79	**5.20**	20.25	55.68	0.8207	0.6028
	Proposed	**5.20**	4.84	21.86	**59.71**	**0.9898**	**0.8290**
Dataset3	SR	3.18	4.52	20.19	61.50	0.7640	0.5157
	DWT	3.12	4.86	25.11	66.53	0.7489	0.5473
	NSST	3.20	4.88	24.52	65.89	0.7733	0.5971
	CNN	3.38	4.39	25.99	69.60	0.7775	0.6042
	Guided filter	3.34	5.05	24.39	69.63	0.7762	0.6119
	Proposed	**4.12**	**5.26**	**26.45**	**73.14**	**0.9766**	**0.8260**
Dataset4	SR	3.33	7.56	28.98	69.84	0.6532	0.4964
	DWT	3.08	7.41	35.94	76.80	0.6263	0.4699
	NSST	3.23	7.44	34.60	79.49	0.6628	0.5349
	CNN	3.26	7.31	**37.03**	**79.84**	0.6462	0.5171
	Guided filter	3.52	**7.76**	34.30	75.36	0.6602	0.5510
	Proposed	**4.64**	7.65	30.38	79.40	**0.9851**	**0.8558**
Dataset5	SR	3.19	5.24	17.58	51.71	0.7427	0.4823
	DWT	2.80	5.36	22.28	55.72	0.7098	0.4573
	NSST	2.94	5.44	21.47	53.79	0.7311	0.5226
	CNN	3.18	4.83	23.06	61.11	0.7448	0.5214
	Guided filter	3.23	**5.78**	21.56	66.98	0.7342	0.5330
	Proposed	**4.03**	5.53	**25.43**	**68.08**	**0.9795**	**0.8326**
Dataset6 MR-T2 SPECT	SR	4.02	4.21	18.86	47.02	0.7617	0.6056
	DWT	3.37	3.63	18.74	41.36	0.7239	0.5783
	NSST	4.07	4.13	19.28	48.53	0.7451	0.7120
	CNN	3.86	4.04	20.06	45.82	0.7604	0.7248
	Guided filter	3.65	3.77	18.56	41.27	0.7462	0.6761
	Proposed	**4.47**	**4.39**	**24.62**	**50.99**	**0.9934**	**0.9434**
Dataset7 MR-T2 PET	SR	4.65	5.18	22.19	63.45	0.7572	0.5660
	DWT	3.91	4.43	23.11	58.61	0.7189	0.5371
	NSST	4.32	4.98	24.52	**68.94**	0.7411	0.6708
	CNN	4.17	4.76	**27.96**	62.47	0.7584	0.6836
	Guided filter	3.97	4.45	24.39	58.26	0.7432	0.6449
	Proposed	**4.84**	**5.64**	26.61	66.28	**0.9901**	**0.8710**

was used to highlight how different the comparing methods were from one another. The fused images show a slightly low intensity in the highlighted area, as shown in Figs. 5(c) and (d). The NSST and CNN fusion images are shown in Figs. 5(e) and (f) tend to preserve more information from the CT image, but they also tend to lose some information from the MRI image. The visual clarity of the guided filter is comparable to that of our proposed algorithm, which can preserve all of the original image's features. However, as Figs. 5(g)–(h) demonstrates that the contrast of the guided filter is less than that of our method.

Fig. 6 displays the fusion findings of Dataset-2, the second medical image set. The application of DWT results in a poor visual effect and a loss of information regarding hard tissues, such as bone structures, as demonstrated in Fig. 6(d); Fig. 6(c) likewise exhibits this low contrast problem. The results of the other three techniques already in use did not differ much. Since our method, in this case, cannot identify the row-and column-wise frequency variations of pixels in the CT scan, the suggested method's retention of CT information is not as strong as that of the guided filter. The recommended methodology yielded a fused image with solid contrast and complete preservation of soft tissue information, as demonstrated in Fig. 7(h), which compares the suggested method to other comparison techniques using the third medical data set (Dataset-3).In Figs. 8 and 9, which include the fusion results of Dataset-4 and Dataset-5, the NSST, DWT, and SR methods lose details from the source images and don't give adequate information on the bone structure. The recommended procedure yields results with stronger contrast, more sharp edges, and finer details.

The metastatic bronchogenic images from MR and SPECT-TC are fused, as shown in Figure 10. The chopped sections are shown separately for a more readable view. GFF is not able to maintain the original hue. On DWT, contrast is inadequate. Although the NSST and SR approaches appear to have a high contrast, they are not as successful in preserving the originality of the MR-T2 picture's structural elements. The color, energy, and structural aspects of the original images are preserved in the proposed method.

A combination of MR and PET images with moderate Alzheimer's disease are fused, as shown in Figure 11. Parts that have been cropped are shown separately to provide a better image. GFF is unable to maintain the original color. DWT has a low contrast. While CNN and SR methods successfully restore the characteristics, the original color is gone. NSST has unduly improved the image. In terms of retaining structure and color information, the proposed method works incredibly well to preserve uniqueness.

6. OBJECTIVE ANALYSIS

It is necessary to use both qualitative and quantitative evaluation criteria while measuring fusion performance. Quantitative assessment metrics, including mutual information (MI), image entropy (IE), mean structural similarity (MSSIM), spatial frequency (SF), standard deviation (SD), and margin information retention ($Q_{AB/F}$), are used in this work to evaluate the efficacy of different fusion procedures [29–31].

Table 1 displays the quantitative analyses based on image evaluation metrics. The top results are shown in bold. The suggested method of fusing the images performed significantly well in terms of MI, IE, QAB/F, & MSSIM, and the other metrics are barely comparable to one another, demonstrating that it can preserve edge features and saliency information. Figure 12 presents a visual comparison of several ap-

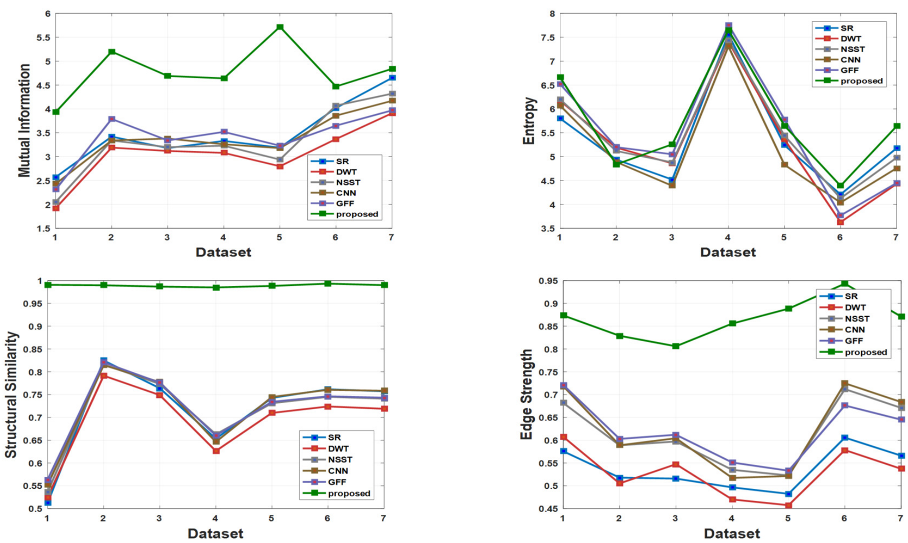

Fig.12: *Comparative analysis of average value of quality assessment metrics.*

proaches based on the average values of four key relative assessment metrics (MI, IE, QAB/F & MSSIM) studied over thirty different slices of seven datasets.

7. RESULT ANALYSIS

The experimental results show that low intensity and insufficient bone structure knowledge produce unsatisfactory fusion outcomes for the Sparse Representation (SR) & DWT techniques. While NSST, CNN, and guided filter techniques have a reasonable visual impact, MR-CT does not entirely preserve the edge and texture within the yellow highlighted zone.

Furthermore, the color could not be maintained by the Guided Filter approach in the MR-SPECT and MR-PET fusion situations. However, the suggested method maintains saliency features, which offer the most information about the soft tissues and bone structure, resulting in fused images that are sharper and more vivid. Among the six measures considered are image entropy (IE), spatial frequency (SF), and standard deviation (SD). These measures are frequently used to demonstrate the inherent qualities of a single image and assess the quality of fused images. The data entropy of the fused image is represented by the IE. The SF makes the image transparent. The SD characterizes the contrast of the combined image. A higher SD results in a more widely scattered grey level dispersion of the image, whereas a higher contrast facilitates the perception of the fused image. The additional data that some existing approaches contain are redundant, increasing the value of each of these metrics.

To facilitate a more thorough objective examination, this work adds three more metrics: QAB/F, MSSIM, and MI. The MI indicates how similar the intensity distributions of the linked image pairings are, as well as providing an estimate of the amount of information recovered from the source images. As more information is extracted from the original images and the combined image's clarity and activity level increase, the MI value also increases. The composite image's level of distortion is determined by the MSSIM. The quantity of edge information carried over from the original images into the fused image is also evaluated by $\mathrm{Q}_{\mathrm{AB/F}}$. With the incorporation of more edge information, including texture and bone structure, this metric becomes more significant for clinical image fusion the more values it possesses. This is because it enables correct pathological assessment of edges. Based on a statistical examination of experimental findings from both the existing and proposed techniques, it was shown that the proposed algorithm improves mutual information by 25%, image entropy by 9.5%, spatial frequency by 21%, standard deviation by 18.1%, structural similarity index by 30%, and edge strength of the fused image by 39% compared to existing methods.This shows minimal distortion in the fused image and an adequate number of prominent features, soft tissue details, denser structures, and significant edge details.

8. CONCLUSION

In this paper, a novel fusion technique for MRI and CT medical images is proposed by combining convolutional neural networks and cross-guided filters in the spatial domain. The proposed method uses a cross-guided filter to decompose the input images and extract detail layers. Convolutional neural networks are used to determine the strength of details in the source images, which are then used to fuse the input images using the weighted average rule to create the fused image. Comprehensive contrast experiments have been carried out on many pairs of CT and MR images, proving that the proposed method is superior in both qualitative and objective assessment.

ACKNOWLEDGEMENT

The authors would like to extend their sincere gratitude to the anonymous reviewers for their thoughtful feedback and constructive comments, which have greatly improved the quality of this manuscript. We are also thankful to the editor for professional handling of the review process.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: Conceptualization and Methodology: Dr.Srikanth M.V, Sivalal Kethavath; Data collection: Dr.Jatothu Brahmaiah Naik; Software: Dr.Nagasirisha.B, Analysis and interpretation of results: Dr.Nagasirisha.B , Dr.Srikanth M.V, Dr.Srinivas Yerram; Draft manuscript preparation: Dr.SivaNagiReddy Kalli, Dr.Jatothu Brahmaiah Naik; All authors reviewed the results and approved the final version of the manuscript.

References

[1] L. Ali, I. Wajahat, N. A. Golilarz, F. Keshtkar and S. A. C. Bukhari, "LDA-GA-SVM: improved hepatocellular carcinoma prediction through dimensionality reduction and genetically optimized support vector machine," *Neural Computing and Applications*, vol. 33, no. 7, pp. 2783-2792, 2021.

[2] F. S. Ahmad, L. Ali, H. A. Khattak, T. Hameed, I. Wajahat, S. Kadry and S. A. C. Bukhari, "A hybrid machine learning framework to predict mortality in paralytic ileus patients using electronic health records (EHRs)," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 3, pp. 3283-3293, 2021.

[3] T. Meraj, H. T. Rauf, S. Zahoor, A. Hassan, M. I. Lali, L. Ali, S. A. C. Bukhari and U. Shoaib,"Lung nodules detection using semantic

segmentation and classification with optimal features," *Neural Computing and Applications*, vol. 33, pp. 10737-10750, 2021.

[4] S. Kadry, Y. Nam, H. T. Rauf, V. Rajinikanth and I. A. Lawal, "Automated detection of brain abnormality using deep-learning-scheme: a study," in *Proceedings of 2021 7th International Conference On Bio Signals, Images, And Instrumentation (ICBSII)*, Chennai, India, pp. 1-5, 2021.

[5] M. A. Khan, I. Ashraf, M. Alhaisoni, R. Damasevicius, R. Scherer, A. Rehman, and S. A. C. Bukhari, "Multimodal brain tumor classification using deep learning and robust feature selection: a machine learning application for radiologists," Diagnostics, vol. 10, no. 8:565, 2020.

[6] V. Rajinikanth, S. Kadry, R. Damasevicius, D. Taniar and H. T. Rauf, "Machine-learning-scheme to detect choroidal-neovascularization in retinal OCT image," in *Proceedings of 2021 7th International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*, Chennai, India, pp. 1-5, 2021.

[7] M. Srikanth, V. Prasad and K.Satya Prasad, "Brain Tumor Detection Through Modified Optimization Algorithm by Region-based Image Fusion," *ECTI-CIT Transactions*, vol. 17, no. 1, pp. 117–127, Mar. 2023.

[8] M. V. Srikanth, V. V. K. D. V. Prasad and K. S. Prasad, "Application of Novel Improved Firefly Algorithm for Image Fusion To Detect Brain Tumor," *ECTI-CIT Transactions*, vol. 17, no. 2, pp. 292–307, Jun. 2023.

[9] B. Meher, S. Agrawal, R. Panda and A. Abraham, "A survey on region based image fusion methods," *Information Fusion*, vol. 48, pp. 119-132, 2019.

[10] S. Li, X. Kang, L. Fang, J. Hu and H. Yin, "Pixel-level image fusion: a survey of the state of the art," *Information Fusion*, vol. 33, pp. 100-112, 2017.

[11] H. Zhang, H. Xu, X. Tian, J. Jiang and J. Ma, "Image fusion meets deep learning: a survey and perspective," *Information Fusion*, vol. 76, pp. 323-336, 2021.

[12] J. Agarwal and S. S. Bedi, "Implementation of hybrid image fusion technique for feature enhancement in medical diagnosis," *Human-centric Computing and Information Sciences*, vol. 5, no. 3, 2015.

[13] S. Aymaz and C. Kose, "Multi-focus image fusion using stationary wavelet transform (SWT) with principal component analysis (PCA)," in *Proceedings of 2017 10th International Conference on Electrical and Electronics Engineering (ELECO)*, Bursa, Turkey, pp. 1176-1180, 2017.

[14] D. P. Bavirisetti and R. Dhuli, "Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-Loeve transform," *IEEE Sensors Journal*, vol. 16, no. 1, pp. 203-209, 2015.

[15] J. Ma, Z. Zhou, B. Wang and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," *Infrared Physics & Technology*, vol. 82, pp. 8-17, 2017.

[16] K. Zhan, Y. Xie, H. Wang, and Y. Min, "Fast filtering image fusion," *Journal of Electronic Imaging*, vol. 26, no. 6:063004, 2017.

[17] L. Lu, X. Ren, K. H. Yeh, Z. Tan, and J. Chanussot,"Exploring coupled images fusion based on joint tensor decomposition," *Human-centric Computing and Information Sciences*, vol. 10, no. 10, 2020.

[18] X. Li, X. Guo, P. Han, X. Wang, H. Li and T. Luo, "Laplacian Redecomposition for Multimodal Medical Image Fusion," in *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 9, pp. 6880-6890, Sep. 2020.

[19] B. Goyal, D. C. Lepcha, A. Dogra, V. Bhateja and A. Lay-Ekuakille, "Measurement and analysis of multimodal image fusion metrics based on structure awareness using domain transform filtering," *Measurement*, vol. 182, no. 109663, 2021.

[20] J. Jose, N. Gautam, M. Tiwari, T. Tiwari, A. Suresh, V. Sundararaj and M. R. Rejeesh, "An image quality enhancement scheme employing adolescent identity search algorithm in the NSST domain for multimodal medical image fusion," *Biomedical Signal Processing and Control*, vol. 66, article no. 102480, 2021.

[21] M. Kaur and D. Singh, "Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks," Journal of Ambient Intelligence and Humanized Computing, vol. 12, no. 2, pp. 2483-2493.(2021)

[22] M. V. Srikanth, V. V. K. D. V. Prasad and K. Satya Prasad, "A 2D Histogram-Based Image Thresholding Using Hybrid Algorithms for Brain Image Fusion," *International Journal of System Dynamics Applications (IJSDA)*, vol. 11, no. 6, 2021.

[23] M. V. Srikanth, V. V. K. D. V. Prasad and K. Satya Prasad, "An improved firefly algorithm-based 2-D image thresholding for brain image fusion," *International Journal of CognitiveInformatics and Natural Intelligence (IJCINI)*, vol. 14, no. 3, 2020.

[24] S. Li, X. Kang and J. Hu, "Image Fusion With Guided Filtering," in *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2864-2875, Jul. 2013.

[25] H. Zhang, X. Han and R. Zhang, "Multimodal Image Fusion Method Based on Guided Filter," *International Journal of Pattern Recog-*

nition and Artificial Intelligence, vol. 36, no. 1:2254003, 2022.

[26] Z. Wang, X. Li, H. Duan, Y. Su, X. Zhang and X. Guan, "Medical image fusion based on convolutional neural networks and non-subsampled contourlet transform," *Expert Systems with Applications*, vol. 171, no. 114574, Jun. 2021.

[27] K. Wang, M. Zheng, H. Wei, G. Qi and Y. Li, "Multi-Modality Medical Image Fusion Using Convolutional Neural Network and Contrast Pyramid," *Sensors*, vol. 20, no.8:2169, Apr. 2020.

[28] K. A. Johnson and J. Alex Becker, "The whole brain atlas," [Online]. Available: `https://www.med.harvard.edu/aanlib/`.

[29] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[30] V. Petrović and C. Xydeas, "Evaluation of image fusion performance with visible differences," *8th European Conference on Computer Vision, ECCV 2004, Lecture Notes in Computer Science*, vol. 3023, pp. 380–391, 2004.

[31] C. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electronics Letters*, vol. 36, no. 4, pp. 308–309, 2000.

Srikanth M.V. received B.Tech. (ECE) degree from JNTU, Hyderabad, India in 2006,M.E (Communication Engineering) from Osmania University in 2009, Ph. D in the field of Image Processing from Jawaharlal Nehru Technological University, Kakinada,Andhra Pradesh, India in 2024. He has been working as an Academician in the capacity of an Associate Professor in Electronics and Communication Engineering department of Usharama College of Engineering and Technology, a reputed Autonomous Engineering College in Andhra Pradesh, India. He has 17 years of experience in the field of Education and research at various levels. He has published more than 25 articles in reputed International journals and conferences. His keen interests are inclined more towards Signal and Image Processing and Embedded Systems.

Sivalal Kethavath is a Research Scholar pursuing his Ph.D. in the field of medical image analysis from NIT Warangal, Telangana, India. He is currently serving as an Assistant Professor in the Electronics and Communication Department at RGUKT Nuzvid, Andhra Pradesh. His research interests include Deep learning, signal processing, and advanced medical imaging techniques. His passion for integrating technology with healthcare drives his innovative approach to academic research. He is dedicated to making impactful contributions to both academia and the medical imaging field.

Srinivas Yerram obtained M.E degree from Osmania University, Hyderabad, India and Ph.D. from Jawaharlal Nehru Technological University Hyderabad, India. He is currently working as Professor in the Department of Electronics and Communication Engineering at Vignana Bharathi Institute of Technology at Hyderabad affiliated to JNTU Hyderabad. He has 24 years of experience in the field of Education, research and administration at various levels. He has published more than 15 articles in reputed International journals and conferences. He is recipient of Best teacher award in 2023 and 2011.Apart from Academics, he is an active involvement in Professional Societies as life member in IETE. His area of interest is Antennas and Microwaves.

Siva NagiReddy Kalli obtained is B.E Degree from Bharathidasan University, Trichy, India, M.Tech Degree from JNTUK Kakinada, India, Ph.D. from Jawaharlal Nehru Technological University, Hyderabad, India. He is currently designated as Professor in the Department of Electronics and Communication Engineering at Sridevi Women's Engineering College at Hyderabad affiliated to JNTU Hyderabad. He has 19 Years of Experience in the field of Education, Research and Administration at various levels. He has obtained Grant for one Australian Patents and published two Indian Patents. He has published more than 35 Articles in IEEE/SCI/Scopus/WoS Journals, conferences and also Reviewer in various leading Journals. He has filed4 Patents, Published 3 Book Chapters and authored 4 Books of Contemporary Models of Wireless Communications and Networks, Wireless Communications and Networks, Digital design Through Verilog, Verilog: Digital VLSI. He is a Recipient of Faculty Excellency Award (2021). Apart from Academics, he is an active involvement in Professional Societies as a Life Member in IEEE, Asian Council of Science Editors, IAOE, and IAENG.

Nagasirisha.B received B.Tech. (ECE) degree from JNTU, Hyderabad, India in 2005, M.Tech. (Digital systems and computer electronics) from JNTUK, ananthapuramu in 2012, Ph. D in the field of Speech processing from Jawaharlal Nehru Technological University, Kakinada,Andhra Pradesh in 2023. She has more than 16 years of experience in teaching and 8 years in R&D. Her current research interests include Speech denoising, adaptive filters, noise modeling in EEG and ECG signals,Signals & Systems, Communications, Digital signal processing, RADAR and Telemetry. She is working as Assistant Professor of Electronics & Communication Engineering in Gudlavalleru Engineering College, an Autonomous NBA accredited College in Andhra Pradesh.

Jatothu Brahmaiah Naik received Ph.D in Image Processing from JNTU, Kakinada in 2020 and his M.Tech in Instrumentation control system from UCEK, Kakinada in 2010. He is currently designated as Professor in the Department of Electronics and Communication Engineering at Narasaraopeta Engineering College, at Guntur, affiliated to JNTU Kakinada. He has 17 Years of Experience in the field of Education, Research and Administration at various levels. He published many papers in conferences and journals in the area of Signal processing and Image processing.