Brain MRI and CT Image Fusion Using Multiscale Local Extrema and Image Statistics

M. V. Srikanth[1†], **A. Suneel Kumar**[1],
B. Nagasirisha[2], and **T. Venkata Lakshmi**[2], Non-members

ABSTRACT

In medical applications such as radiotherapy and guided-image surgery, data fusion for diagnostic imaging has emerged as a critical issue. Because the objective of medical image fusion is to improve patient diagnosis accuracy, the fused image is created by preserving the source images' prominent details and features. It has been demonstrated that the Multi-Level Local Extrema (MLE) representation has numerous advantages over traditional image modeling approaches. We propose an innovative MLE-based fusion method for multimodal medical images in this paper. In the MLE schema, the proposed algorithm decomposes source images into coarse and detailed layers, then fuses the source images using weights calculated from these detail images using image statistics. We visually and quantitatively compared the efficacy of the suggested approach to that of existing methods using five different types of medical images from various sources. The experimental results showed that the proposed scheme outperforms other current typical schemes in terms of both qualitative image quality and objective evaluation. From the quantitative results it is concluded that the proposed algorithm provides an improvement of 64.2% in mutual information,12.3% improvement in image entropy, 26.5% improvement in spatial frequency, 21.8% improvement in standard deviation, 44% improvement in Structural similarity index, 66.6% improvement in edge strength of fused image than existing methods considered for comparison.

Keywords: Multi-Level Local Extrema, Medical Image Fusion, Image Statistics, Qualitative Evaluation

1. INTRODUCTION

With the advancement of computer science and sensor technology, medical imaging has become an essential tool in a variety of clinical applications such as surgical navigation, medical diagnosis, treatment planning, and it is an essential tool for the doctors to accurately diagnose a particular disease[1]. Medical images are commonly produced by various imaging mechanisms that focus on specific organ or a tissue information , such as computed tomography (CT), magnetic resonance imaging (MRI), and X-ray. While CT images are being used to precisely localize dense structures such as bones and implants, MRI images could provide sufficient soft tissue information with more functional information [2]. Image fusion's primary goal is to create a comprehensive image that contains the distinctive features of multimodal medical images in it, allowing physicians to make more accurate diagnosis [3]. Multiscale transform (MST) methods for image fusion have gained considerable interest in recent years. Among the traditional MST tools are Laplacian Pyramid (LAP) [4], Discrete Wavelet Transform (DWT) [5], and contourlet transform [6]. Da Cunha et al. [7] suggested Non-Subsampled Contourlet Transform (NSCT) in order to achieve higher selectivity and regularity than Contourlet Transform and to remove pseudo-Gibbs manifestations along the edges to some extent. In comparison to other decomposition techniques, NSCT requires more computation. Non-Subsampled Shearlet transform (NSST) was suggested by Zhang et al. [8] to reduce the computational complexity of NSCT; NSST inherits the key characteristics of Shearlet and wavelet, such as anisotropy and processing speed, and possesses the shift-invariance of Non-Subsampled processes.

As a result, NSST has a substantial advantage when it comes to acquiring more information for image fusion. Edge-preserving filtering is frequently employed as an image decomposition technique in image fusion because of its superior feature of edge preservation [9-10]. Edge-preserving filtering decomposition, in contrast to MST decomposition, fully utilizes the spatial information of the image and can produce a wider variety of decomposition outputs. Bilateral filtering (BF) [11], Gaussian curvature filtering (GCF) [12], and Guided image filtering (GF) [13] are examples of common edge preserving algorithms. A crucial step in image fusion is learning how to create a thorough weight map in addition to choosing a good decomposition technique. Traditional transform domain fusion approaches used simple rules like average or choose-max to construct the weight maps. This type of fusion rule ignores the relationship among pixels, which reduces the contrast of the fused

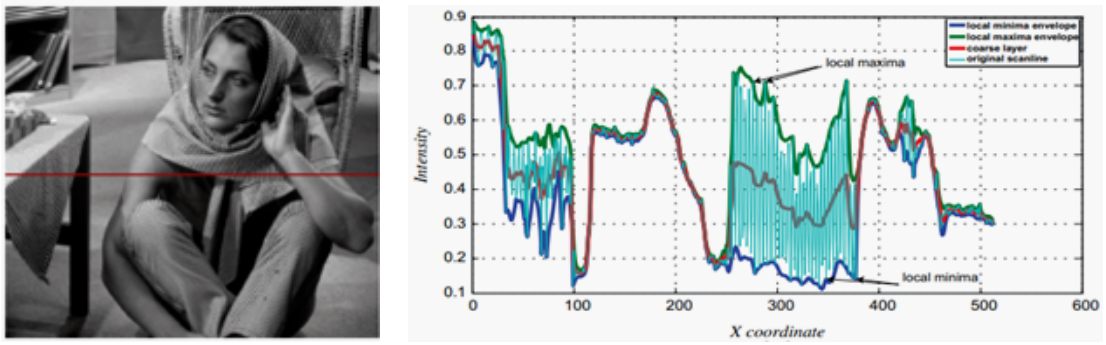
Manuscript received on April 11, 2023; revised on June 9, 2023; accepted on January 8, 2024. This paper was recommended by Associate Editor Yodchanan Wongsawat.

[1]The authors are with Usha Rama College of Engineering and Technology, Andhra Pradesh, India.

[2]The authors are with Gudlavalleru Engineering College, Andhra Pradesh, India.

[†]Corresponding author: sree.02476@gmail.com

Digital Object Identifier: 10.37936/ecti-eec.2024221.252945



(a) *(b)*

***Fig. 1**: Illustration of local maxima and minima envelopes.*

image and causes salient information to be lost to some extent. The proposed scheme generates weight maps by evaluating the strength of information in a detail image utilizing image statistics. Many neural network models for image fusion have been proposed, including pulse coupled neural networks (PCNNs) and self-organizing feature maps (SOFM). A SOFM network cluster input sampled by competitive learning, with the winner chosen as the ultimate fusion component, was proposed by Zhang et al. [14], but the total number of required output neurons had to be adjusted manually prior to each fusion. Wang et al. [15-16] performed multi-focus image fusion using dual channel PCNNs, however, the initial PCNN parameters had to be manually adjusted again, and the average brightness of the fused image was little lower than that of original source.

The following are some restrictions and problems associated with Artificial Neural Network-based fusion: (1) the ANN's longer iteration time and require manual parameterization before every fusion; and (2) a rare malfunction to converge, like with the pulse coupled neural networks based method whenever few neurons do not fire throughout the entire iteration. To conquer the complex computations in MST-based fusion, and manual parameterization and the longer iteration time in ANN based fusion, we suggest an innovative fusion algorithm at pixel-level for multi-modal image data based on the local extrema, drawing inspiration from the work of Subr et al. [18]. We decompose an image into the multi-layers with the same size as that of the original image by iteratively applying the MLE technique on the low frequency components of the image, initially starting with source image. The uniqueness of the suggested approach is in the fusion step, which creates weight maps for the pixels by using the detailed layer at each scale. This is accomplished by employing the horizontal and vertical Eigen vectors of detail layers. Pixels with high Eigen values are classified as edge pixels with important saliency features, whilst those with lower values are treated as inconsequential saliency features because Eigen values more accurately capture the ambiguity of the pixel information.

The remainder of this paper is structured as follows. The concept of coarse and detailed layers is explained in the following section. Section 3 describes a local extrema-based image fusion algorithm, and Section 4 and 5 provides our experimental data and compares the effectiveness of different methods. Finally, the paper is concluded with summarization on the obtained fusion results.

2. MECHANISM OF COARSE AND DETAIL LAYER DECOMPOSITION

An image could be divided into coarse and detail layers. As shown in Fig. 1, the coarse layer could be formed by taking the average of the local minimum and local maximum envelopes, while the detail layers could be thought of as the vibration between such envelopes [18]. In order to construct the local minima and maxima envelopes, the local minima and maxima are individually interpolated. The average of these two interpolates at each pixel makes up the coarse layer, which gives an approximate representation of the local mean around which the vibrations occur. The detail layer could be extracted by subtracting the coarse layer from the actual image. The coarse layer includes the original image's intensity distribution and easily discernible intensity variations. The detail layer contains the original's local edge information and texture pattern. The coarse layer in a medical image replicates functional and anatomical information about organs and tissues, while the detail layer conveys local edges and texture of the tissue.

The purpose of medical image fusion would be to incorporate as much diagnostically helpful data as is possible from multiple sources into a single image. In addition to texture and local boundaries of the tissue, the information about anatomical organs and functioning tissues, is always the most crucial. Information with a high diagnostic value is primarily found in areas where both the approximate and detail layers have experienced significant alterations.

The algorithm for obtaining the detailed and coarse layers from an input grayscale image F(x,y), where (x,y) represent the image's spatial coordinates and F(x,y) is the

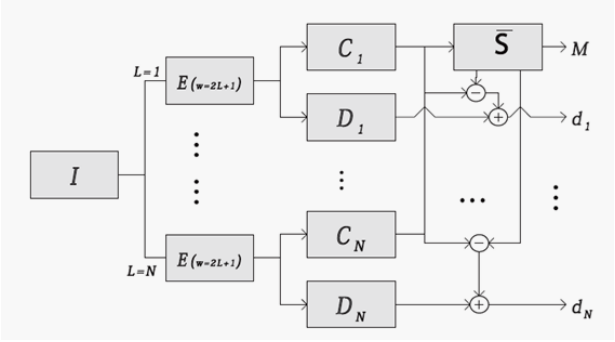


***Fig. 2**: Multiscale image decomposition using MLE.*

pixel intensity at (x,y) location is explained as follows: All pixels along four borders of F are considered extrema points, with either L_{max} or L_{min}. On the extrema pixels, Delaunay interpolation is used to create the extrema envelopes E_{max} and E_{min}. Averaging E_{max} and E_{min} yields the coarse layer M, while subtracting M from the original image F yields the detail layer D.An image is decomposed into coarse and detail layers in order to separate the image's edge and texture information. This mechanism produces quite similar results as like convoluting an image with a Gaussian mask [1] , by taking the average of the local minima and maxima envelopes. The final image cannot fully capture all of the texture information that is still available in the convoluted version of the image after it has been subtracted from it. Our decomposition scheme is capable of separating texture from the original image more effectively.

3. PROPOSED METHODOLOGY

3.1 Multiscale Image Decomposition

Let A and B are two input images obtained from different imaging modalities. The source images are decomposed into a coarse layer and multiple detail layers using the local extrema envelope method at different scales [21]. The coarse layer is an estimation of an original image, while the detail layers show high frequency details such as textures and edges at different levels of the wxw sample window. The multiscale decomposition mechanism of an input image into a coarse layer M and detail layers $d_L, L = 1, 2, 3, \ldots . N$ is illustrated in Fig. 2.

From Fig. 2, an image I is divided into coarse and detail layers at different scales L=1,2... N, which are labeled as C_L and D_L . The final coarse layer M of an image Ican be obtained from the Eq. (1) by averaging all the coarse layers at different scales. Using M, the detail layers at various levels are computed using equation (2), which are obtained by adding the difference between final coarse layer M and coarse at each scale C_L to the detail layer at that scale D_L.

$$M = \frac{1}{N}\sum_{L=1}^{N} C_L \tag{1}$$

$$d_L = D_L + (C_L - M) \tag{2}$$

In this manner, the detail layers at different levels (four levels in our work) are obtained for each source image A and B which are designated as $d_A\{L\}$ and $d_B\{L\}$ respectively.

3.2 Weight computation using image statistics

Image weights are evaluated in our scheme using statistical properties of detail coefficients (based on vertical and horizontal edge strengths). To determine the optimum weight, a square window of size *nxn* (in this work it is taken as 5) considered as a neighborhood around detail coefficients $d_A\{L\}$ for each level and is denoted by a matrix X. Determine the covariance matrix of X using Eq. (3) and unbiased estimate of X, $C_H^{u,v}(X)$ using Eq. (4) at pixel location (u, v) by treating the row as an observation and the column as a variable.

$$Covariance(X)) = \mathrm{E}[(\mathrm{X}-\mathrm{E}[\mathrm{X}])(\mathrm{X}-\mathrm{E}[\mathrm{X}])^{\mathrm{T}}] \tag{3}$$

$$C_H^{u,v}(X) = \frac{1}{n-1}\sum_{z=1}^{n}(X_z - \overline{X})(X_z - \overline{X})^T \tag{4}$$

The variance vector is given by the diagonal elements of $C_H^{u,v}(X)$. Find the Eigen values of $C_H^{u,v}(X)$, which are given by λ_H. The sum of these Eigen values yields horizontal strength of detail layer $d_A\{L\}$, which is given by Eq. (5).

$$\beta_H(u,v) = \sum_{j=1}^{n}\lambda_H^j \tag{5}$$

Using the same procedure, vertical estimate $C_V^{u,v}(X)$ is obtained at pixel location (u, v) by treating the column as an observation and the row as a variable. Now find the Eigen values of $C_V^{u,v}(X)$ and are given by λ_V. The sum of these Eigen values gives vertical strength of detail layer $d_A\{L\}$ and is given by Eq. (6).

$$\beta_V(u,v) = \sum_{j=1}^{n}\lambda_V^j \tag{6}$$

From $\beta_H(u,v)$ and $\beta_V(u,v)$, the optimal weight of pixel located at (u,v) in the detail layer $d_A\{L\}$ is obtained using the Eq. (7).

$$W(u,v) = \beta_H(u,v) + \beta_V(u,v) \tag{7}$$

3.3 Proposed image fusion algorithm

The flow diagram of proposed algorithm is presented in Fig. 3.

Algorithm

Step 1: Read two input images A and B which are to be fused.

Step 2: Apply multiscale decomposition (four levels) based on local extrema and compute coarse and detail layers of each source image using Eqs. (1) and (2) to yield M_A, $d_A\{L\}$ and M_B, $d_B\{L\}$ respectively.

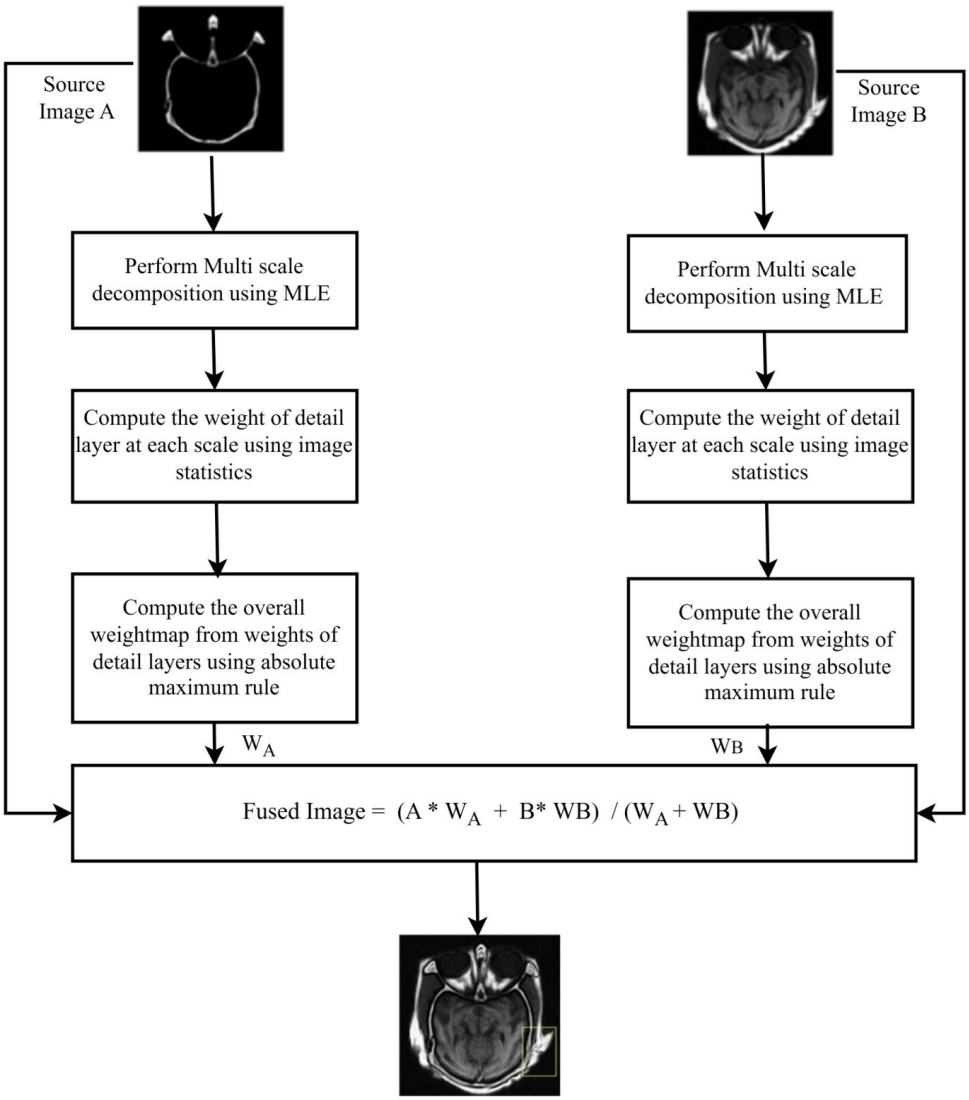


***Fig. 3**: Flow diagram of proposed algorithm.*

Step 3: Designate the detail layers of A obtained from Step 2 at four different levels as $d_A\{1\}$,$d_A\{2\}$,$d_A\{3\}$ and $d_A\{4\}$.

Step 4: Designate the detail layers of B obtained from Step 2 at four different levels as $d_B\{1\}$,$d_B\{2\}$,$d_B\{3\}$ and $d_B\{4\}$.

Step 5: Apply Eqs. (3) to (7) on Steps 3 and 4, to compute the optimal weights of each detail layer given as $W_A\{1\}$,$W_A\{2\}$,$W_A\{3\}$,and $W_A\{4\}$ for detail layers of A and $W_B\{1\}$,$W_B\{2\}$,$W_B\{3\}$ and $W_B\{4\}$ for detail layers of B.

Step 6: Find the overall weights W_A of source image A and W_B of source image B using the following pseudocode.

for $u = 0 : 255$
for $v = 0 : 255$
$W_A(u,v) = \max\left(W_A\{1\}(u,v), ..., W_A\{4\}(u,v)\right)$
$W_B(u,v) = \max\left(W_B\{1\}(u,v), ..., W_B\{4\}(u,v)\right)$
end
end

Step 7: Obtain the final fused image F of the source images A and B using the weighted average rule given in Eq. (8).

$$\text{final fused image},F = \frac{(W_A * A + W_B * B)}{(W_A + W_B)} \tag{8}$$

4. SUBJECTIVE ANALYSIS

To illustrate the efficacy of the suggested fusion method, five pairs of MRI and CT images named "Dataset-1", "Dataset-2", "Dataset-3", "Dataset-4", and "Dataset-5" of brain are chosen . They are all 256x 256 in size and cover 256 grey levels. . Dataset-1 depicts the brain of a healthy patient, Dataset-2 the brain of a patient who suffered a fatal stroke, Dataset-3 the brain images of a patient with a neoplastic tumor, Dataset-4 the sagittal plane images of the brain skull, and Dataset-5 the brain images of a patient who had a cerebellar metastasis. The Benchmark Brain Atlas, which is accessible at http://www.med.harvard.edu/aanlib/home.html, served

as the source for all the datasets. Fig. 4 depicts the fusion results for the Dataset-1 image set using various methods. Figs. 4 (a) and (b) show MRI and CT images, respectively. The fusion results produced from Sparse Representation (SR), Discrete Wavelet Transformation (DWT), Non-Subsampled Shearlet Transformation (NSST), Laplacian Pyramid (LAP), guided filter methods as well as the proposed approach are shown in Figs. 4(c)-(h). The fused results primarily preserve bony structures of the CT image along with the soft tissues of the MRI image. The approaches do, however, differ slightly in how they preserve detail and contrast.

We used a yellow rectangle to highlight the differences in comparison methods. The fusion results have slightly low intensity in the marked region, as shown in Figs. 4(c) and (d). Even though the fusion images utilizing NSST and LAP shown in Figs. 4 (e) and (f) retain more information from the CT image, but they tend to lose some details from the MRI image. With regard to visual clarity, the guided filter performs similarly to our proposed algorithm, which can completely preserve the features of the source images, but its contrast is lower than our method as shown in Figs. 4(g)-(h).

Fig. 5 depicts the second medical image set, Dataset-2. The use of DWT results in a lack of hard tissue information such as bone structures and a poor visual effect as evident from Fig. 5(d); the problem of low contrast is also present in Fig. 5(c). With respect to results of the remaining three existing methods, there were no significant differences. The proposed method outcome is inferior in terms of preserving CT details compared with that of guided filter outcome as our method in this case fail to recognize the row and column wise frequency variations of pixels in CT image.

Fig. 6 compares the proposed method to other comparison algorithms on the third medical image set Dataset-3; the fused image obtained from the suggested algorithm has good contrast and completely preserves the soft tissue information, as revealed in Fig. 6(h).

In Figs. 7 and 8 for Dataset-4 and Dataset-5, the SR, DWT, and NSST methods lose source image details, and bone structure information is insufficient. The proposed method produces results with more details, sharp edges, and increased contrast.

5. QUANTITATIVE ANALYSIS

A quantitative evaluation metric, in addition to qualitative evaluation, is an essential tool for measuring fusion performance. The quantitative evaluation metrics like 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 paper to evaluate the efficacy of different fusion methods [22-25].

1. MI depicts the relationship between two events. The mutual information of two independent random variables X and Y is given below:

$$MI(X,Y) = \sum_{x\in X}\sum_{y\in Y} p(x,y) log_2 \frac{p(x,y)}{p(x)\,p(y)} \tag{9}$$

where p(x,y) represents the joint probability distribution of X and Y and p(x) and p(y) represent the marginal probability distributions of X and Y, respectively. To denote the difference in fusion quality, the amount of MI between each of the two input images and the fused image can be calculated, and then the total mutual information measure can be defined as follows:

$$MI_F^{AB} = MI(A,F) + MI(B,F) \tag{10}$$

Eq. (10) represents the total amount of information transferred into the fused image F(i, j) from both source images A(i, j) and B(i, j). A higher value of MI depicts that the more information is extracted from the source images.

2. Standard Deviation (SD) is an indicator of the level of dispersion of an average set of image data. For fused image, standard deviation is calculated using Eq. (11).

$$SD = \sqrt{\frac{1}{MN}\sum_{i-1}^{M}\sum_{j=1}^{N}(F(i,j)-\mu)^2} \tag{11}$$

$$\mu = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} F(i,j) \tag{12}$$

where μ is called mean and F(i, j) is the fused image pixel value at the location (i,j).This metric represents the clearness of the fused output image; the greater the SD value, the better is the image quality.

3. Spatial Frequency (SF) is defined as the sum of column frequency (CF) and row frequency (RF).It is computed using Eq. (13).

$$SF = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(CF + RF) \tag{13}$$

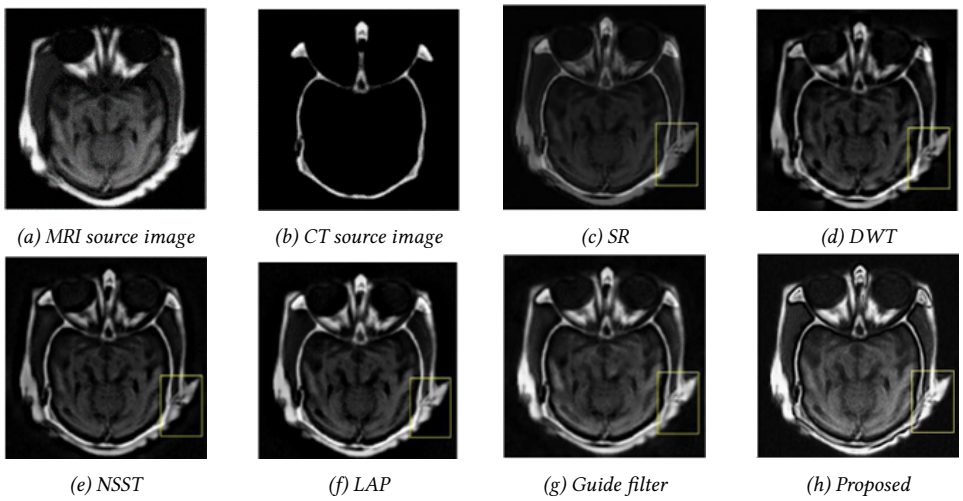
$$CF = [F(i,j) - F(i,j-1)] \tag{14}$$

$$RF = [F(i,j) - F(i-1,j)] \tag{15}$$

Here F(i,j) signifies the image's pixel value and MxN denotes the image size. The higher the score of this metric, the higher is the resolution of the fused image.

4. Image Entropy (IE) signifies the quantity of data available inside a fused image. If the grey level distribution of an image I is P = {P_0,P_1,... P_{L-1}},P_K indicates the probability of occurrence of k^{th} grey level in an image , and L indicates number of grey levels in an image (256 for grey scale image), IE could be calculated as follows:

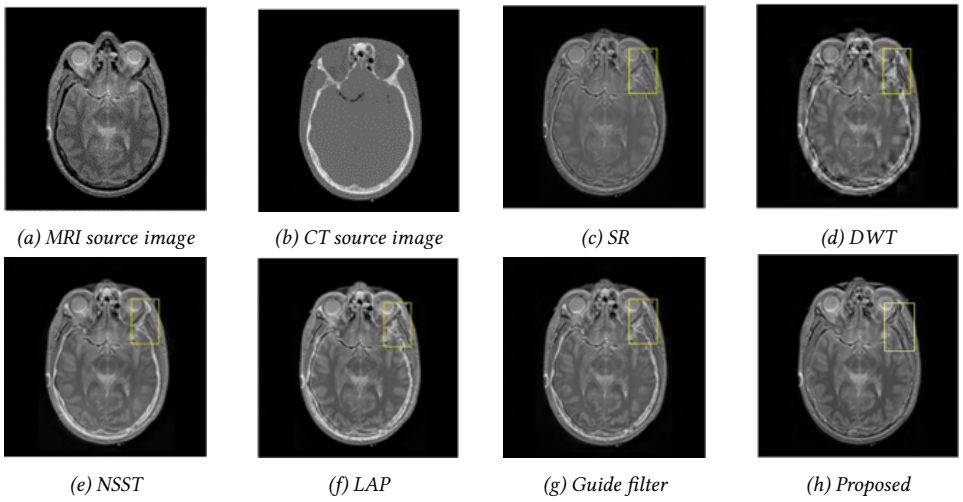
$$IE = -\sum_{i=0}^{L-1} p(i)\, log_2 p(i) \tag{16}$$



(a) MRI source image *(b) CT source image* *(c) SR* *(d) DWT*

(e) NSST *(f) LAP* *(g) Guide filter* *(h) Proposed*

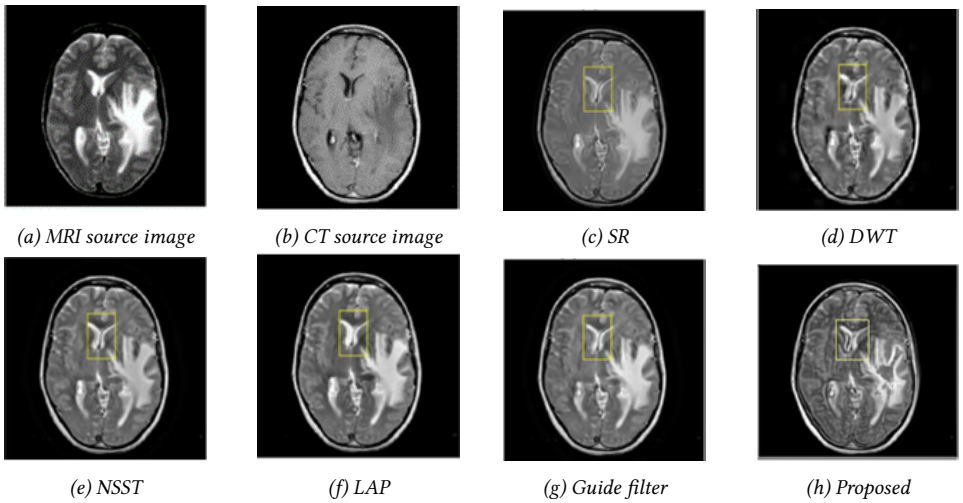
***Fig. 4**: Fused images of dataset-1 for proposed and different existing methods.*



(a) MRI source image *(b) CT source image* *(c) SR* *(d) DWT*

(e) NSST *(f) LAP* *(g) Guide filter* *(h) Proposed*

***Fig. 5**: Fused images of dataset-2 for proposed and different existing methods.*



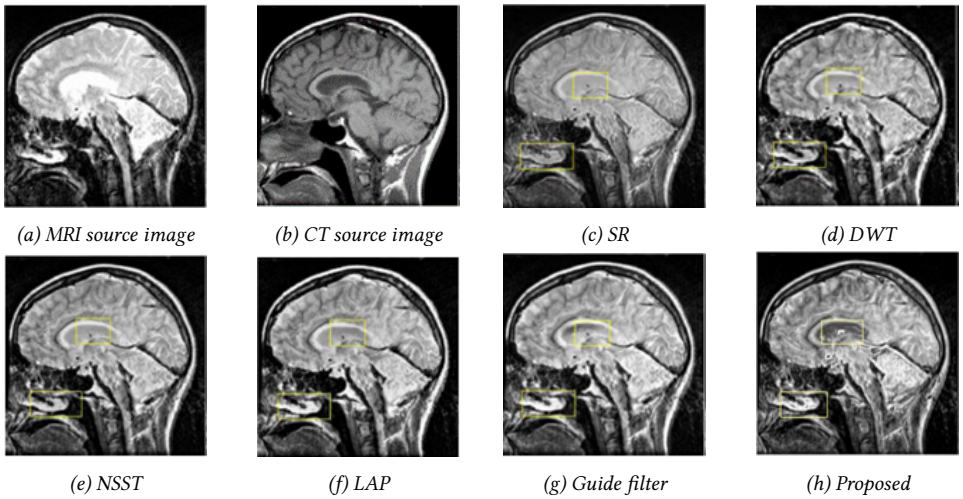
(a) MRI source image *(b) CT source image* *(c) SR* *(d) DWT*

(e) NSST *(f) LAP* *(g) Guide filter* *(h) Proposed*

***Fig. 6**: Fused images of dataset-3 for proposed and different existing methods.*

The higher the IE score, the more is the information contained in the fused image.

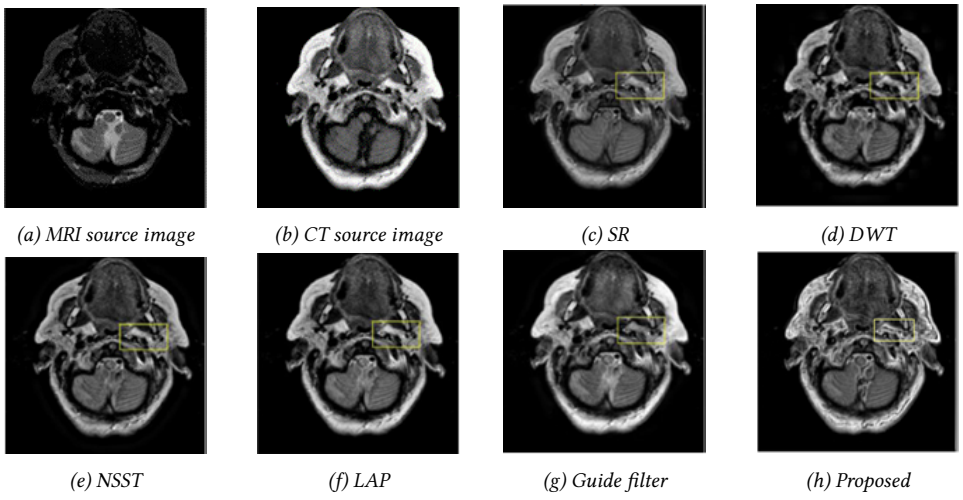
5. The Mean Structural Similarity Index Measure ($MSSIM$) is a useful measure of image similarity that



(a) MRI source image (b) CT source image (c) SR (d) DWT

(e) NSST (f) LAP (g) Guide filter (h) Proposed

***Fig. 7**: Fused images of dataset-4 for proposed and different existing methods.*



(a) MRI source image (b) CT source image (c) SR (d) DWT

(e) NSST (f) LAP (g) Guide filter (h) Proposed

***Fig. 8**: Fused images of dataset-5 for proposed and different existing methods.*

can be calculated as follows:

$$MSSIM = \frac{SSIM(A,F) + SSIM(B,F)}{2} \tag{17}$$

Here $SSIM(A,F)$ and $SSIM(B,F)$ are the correlation coefficients of image A and image B with fused image respectively.

$$SSIM = \frac{\left(2\mu_a\mu_b + C1\right)(2\sigma_{ab} + C2)}{(\mu_a{}^2 + \mu_b{}^2 + C1)(\sigma_a{}^2 + \sigma_b{}^2 + C2)} \tag{18}$$

where μ_a is the mean of a and μ_b is the mean of b , $\sigma_{ab}$is the co-variance of a and b, $\sigma_a{}^2$ is the variance of a and $\sigma_b{}^2$ is the variance of b. Two constants, C1 and C2, are employed to prevent instability that can arise from a division with a value close to zero. A possible range for SSIM values is 0 to 1, with 0 denoting low quality and 1 denoting great quality.The higher the MSSIM score, the lower is the distortion in the fused image.

6. The transition degree of the edge information of the input images into the fused image is represented by $Q_{AB/F.}$ It can be evaluated as follows:

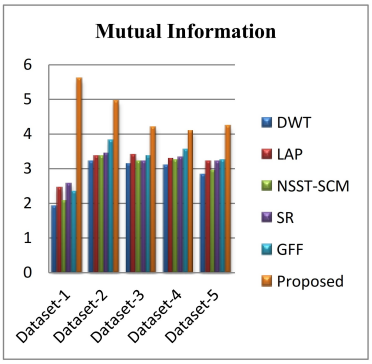
$$Q_{AB/F} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(Q_{AF}(i,j)W_A(i,j) + Q_{BF}(i,j)W_B(i,j)\right)}{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(W_A(i,j) + W_B(i,j)\right)} \tag{19}$$

where Q_{AF} signifies the edge strength and orientation preservation value between image A and fused image F. Similarly Q_{BF} signifies the edge strength and orientation preservation value between image A and fused image F. $W_A(i,j)$, $W_B(i,j)$ are the weights of Q_{AF} and Q_{BF}respectively. If the $Q_{AB/F}$ value increases and approaches unity, it indicates that the fused image has little edge information loss.

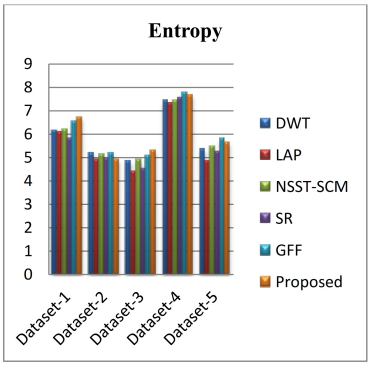
Table 1 displays the quantitative objective evaluations based on image assessment metrics. The best results are marked italic. The proposed scheme of image fusion yields better performance in terms of Image Entropy, MI, QAB/F,MSSIM, and the remaining metrics

***Table 1**: A comparison of quantitative evaluations of various methods.*

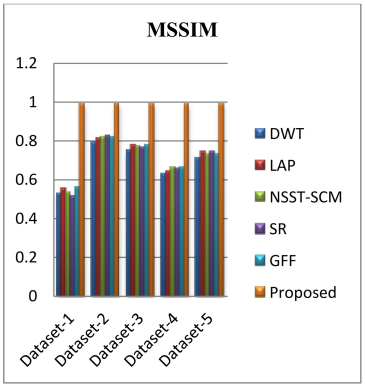
Dataset type	Method	Mutual Information (MI)	Image Entropy (H)	Spatial Frequency (SF)	Standard Deviation (SD)	MSSIM	Edge Strength (QAB/F)
Dataset1	SR	2.57	5.80	11.68	30.82	0.5122	0.5756
	DWT	1.92	6.17	17.13	44.71	0.5246	0.6073
	NSST	2.05	6.20	17.05	44.16	0.5366	0.6816
	LAP	2.43	6.07	18.10	52.89	0.5518	0.7184
	Guided filter	2.31	6.52	16.97	52.89	0.5634	0.7210
	Proposed	**5.58**	**6.72**	**20.05**	**57.35**	**0.9902**	**0.8912**
Dataset2	SR	3.42	4.94	17.76	51.40	0.8248	0.5178
	DWT	3.19	5.19	**22.15**	55.73	0.7915	0.5051
	NSST	3.34	5.12	20.95	54.56	0.8160	0.5887
	LAP	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	**4.94**	4.92	21.44	58.23	**0.9900**	**0.8441**
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
	LAP	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	**4.19**	**5.30**	25.41	**70.06**	**0.9869**	**0.8046**
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
	LAP	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	**4.07**	7.67	29.68	72.48	**0.9858**	**0.8570**
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
	LAP	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	**4.21**	5.65	**24.98**	64.99	**0.9897**	**0.8704**



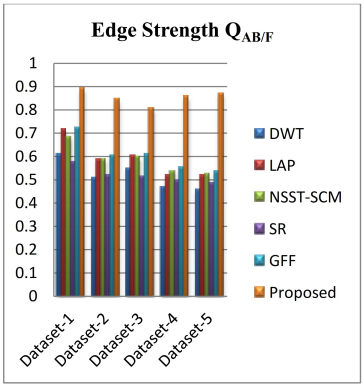
(a) Comparison bar chart of MI



(b) Comparison bar chart of IE



(c) Comparison bar chart of MSSIM



(d) Comparison bar chart of $Q_{AB/Fs}$

***Fig. 9**: Bar chart analysis of quantitative metrics of Table 1.*

differ slightly from one another, demonstrating that the proposed method is competent at preserving saliency information and edge details. Furthermore, the four important relative assessment metrics MI,IE,QAB/F and MSSIM, will be intuitively represented as bar charts, as illustrated in Fig. 9.

6. RESULT ANALYSIS

To sum up the experimental findings, the fusion results produced by the Sparse Representation (SR) and DWT methods are substandard because of poor intensity and lack of sufficient information related to bone structures. Although the visual impact of NSST, LPP, and Guided filter methods are satisfactory, the edge and texture in the yellow highlighted region have not been completely retained. The proposed method, on the other hand, achieves better and high intensity fused images by preserving saliency features, which produces bone structure and information related to the soft tissues to the greatest extent possible. Following this, we go over the quantitative evaluations of the fusion approaches in depth. Among the six considered metrics, image entropy(IE), spatial frequency(SF), and standard deviation(SD) represent the internal features of a single image and are commonly used to assess the quality of fused image. The data entropy of the fused image is represented by the IE. The SF reveals the image's clarity. The contrast of fused image is described by the SD. The spread of the image's grey level dispersion increases with larger SD, and increased contrast results in improved viewing of the fused image. However, some existing methods have more redundant features, which raises the value of these three metrics. In order to conduct more comprehensive objective analysis, three additional metrics are introduced in this work: MI, QAB/F, and $MSSIM$, which can replicate the internal relationship between the source images and fused image more accurately. The MI depicts the similarity of the corresponding image pair's intensity distributions and approximates how much data is gathered from the input images. The higher the MI value, the more details and information are obtained from CT and MRI images, as well as the more clarity and activity level within the fused images. The $MSSIM$ measures the amount of distortion in the fused image. Furthermore, the $Q_{AB/F}$ assess the quantity of edge information relocated from the input images to the fused image. This metric is critical for medical image fusion because the higher the value, the more edge details like texture and bone structure are fused, which aids in precise pathological analysis. From the quantitative evaluation of the experimental results of proposed and existing methods, the suggested technique achieves an excellent performance in terms of MI, QAB/F , and $MSSIM$ metrics, which also implies that the fused image has had sufficient dense structures, soft tissue information, salient features and edge details with less distortion. The proposed method's results are better suited for assisting doctors in the precise diagnosis of diseases.

7. CONCLUSION

A novel method of fusion for MRI and CT medical images is proposed in this paper by incorporating multiscale local extrema mechanism and image statistics in the spatial domain. MLE is used in the proposed method to decompose the input images at multiscale, while image statistics are used to calculate the strength of details in the source images, based on which the input images are fused using the weighted average rule to obtain the fused image. Extensive contrast experiments on various pairs of CT and MR images have been performed, demonstrating the supremacy of the suggested approach in both qualitative as well as objective evaluation. Furthermore, the results show that the proposed fusion scheme has potential applications in the medical field.

Compliance with Ethical Standards: This work is not communicated to any other journal.

Competing Interest: The authors have no competing interests to declare that are relevant to the content of this article.

Data availability and access: The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

REFERENCES

[1] Y. Jia, C. Rong, C. Wu and Y. Yang, "Research on the decomposition and fusion method for the infrared and visible images based on the guided image filtering and Gaussian filter," *3rd IEEE International Conference on Computer and Communications (ICCC)*, Chengdu, China, pp. 1797-1802, 2017.

[2] G. Bhatnagar, Q. M. J. Wu and Z. Liu, "Directive Contrast Based Multimodal Medical Image Fusion in NSCT Domain," in *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1014-1024, Aug. 2013.

[3] X. Liu, W. Mei, and H. Du, "Structure tensor and nonsubsampled shearlet transform based algorithm for CT and MRI image fusion," *Neurocomputing*, vol. 235, pp. 131–139, Apr. 2017.

[4] R. Vijayarajan and S. Muttan, "Discrete wavelet transform based principal component averaging fusion for medical images," *AEU - International Journal of Electronics and Communications*, vol. 69, no. 6, pp. 896–902, Jun. 2015.

[5] Toet A, "Image fusion by a ratio of low-pass pyramid," *Pattern Recogn Lett*, vol. 9, no. 4, pp. 245–253, 1989.

[6] M. N. Do and M. Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," in *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091-2106, Dec. 2005.

[7] A. L. Da Cunha, J. Zhou and M. N. Do, "The Nonsubsampled Contourlet Transform: Theory, Design, and Applications," in *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3089-3101, Oct. 2006.

[8] T. Y. Zhang, Q. Zhou, H. J. Feng, Z. H. Xu, Q. Li, and Y. T. Chen. "Fusion of infrared and visible light images based on nonsubsampled shearlet transform," In *International Symposium on Photoelectronic Detection and Imaging 2013: Infrared Imaging and Applications*, vol. 8907, pp. 366-373. SPIE, 2013.

[9] W. Tan, H. Zhou, J. Song, H. Li, Y. Yu, and J. Du, "Infrared and visible image perceptive fusion through multi-level Gaussian curvature filtering image decomposition," *Applied Optics*, vol. 58, no. 12, p. 3064, Apr. 2019.

[10] Z. Zhou, B. Wang, S. Li, and M. Dong, "Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters," *Information Fusion*, vol. 30, pp. 15–26, Jul. 2016.

[11] C. Tomasi, R. Manduchi, "Bilateral filtering for gray and color images," in *Sixth International Conference on Computer Vision, (IEEE Cat. No. 98CH36271), IEEE*, pp. 839–846, 1998.

[12] W. Tan, H. Zhou, J. Song, H. Li, Y. Yu, and J. Du, "Infrared and visible image perceptive fusion through multi-level Gaussian curvature filtering image decomposition," *Applied Optics*, vol. 58, no. 12, p. 3064, Apr. 2019.

[13] K. He, J. Sun and X. Tang, "Guided Image Filtering," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, June 2013.

[14] Z. Zhang, S.H. Sun, F. Zheng, "Image fusion based on median filters and SOFM neural networks: a three-step scheme," *Signal Processing*, vol. 81, no. 6, pp. 1325–1330, 2001.

[15] Z. Wang, Y. Ma, "Medical image fusion using m-PCNN," *Information Fusion*, vol. 9, no. 2, pp. 176–185, 2008.

[16] Z. Wang, Y. Ma, J. Gu, "Multi-focus image fusion using PCNN," *Pattern Recognition*, vol. 43, no. 6, pp. 2003–2016, 2010.

[17] N. Aishwarya and C. Bennila Thangammal, "A novel multimodal medical image fusion using sparse representation and modified spatial frequency," *International Journal of Imaging Systems and Technology*, vol. 28, no. 3, pp. 175–185, Feb. 2018.

[18] K. Subr, C. Soler, and F. Durand, "Edge-preserving multiscale image decomposition based on local extrema," *ACM Transactions on Graphics*, vol. 28, no. 5, pp. 1–9, Dec. 2009.

[19] M. Hossny, S. Nahavandi, and D. Creighton, "Comments on B Information measure for performance of image fusion," *Electron Lett*, vol. 44, no. 18, pp. 1066–1067, 2008.

[20] X. Jin, R. Nie, D. Zhou, Q. Wang, and K. He, "Multifocus Color Image Fusion Based on NSST and PCNN," *Journal of Sensors*, vol. 2016, pp. 1–12, 2016.

[21] Z. Xu, "Medical image fusion using multi-level local extrema," *Information Fusion*, vol. 19, pp. 38–48, Sep. 2014.

[22] Zhou 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, April 2004.

[23] Vladimir P, Costas X , "Evaluation of image fusion performance with visible differences," *8th European Conference on Computer Vision, ECCV 2004, Lecture Notes in Computer Science*, 3023, pp. 380–391, 2004.

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

M. V. Srikanth is a Research Scholar, pursuing Ph. D in the field of Image Processing from Jawaharlal Nehru Technological University, Kakinada,Andhra Pradesh, India. He has been working as an Academician in the capacity of an Assistant Professor in Electronics and Communication department of Usha Rama College of Engineering and Technology, a reputed Autonomous Engineering College in Andhra Pradesh, India. Apart from regular Academia, the zeal to explore new technologies and a penchant towards pursuing research in the field of Image Processing has made him to do the research work in a governmental organization in Andhra Pradesh. His keen interests are inclined more towards Signal Processing and Embedded Systems.

A. Suneel Kumar is working as Associate professor in the Department of Electronics & Communication Engineering, in Usharama College of Engineering and Technology; an Autonomous NAAC A accredited College in Andhra Pradesh, India. He received Ph. D for his work in Microstrip Patch Antenna in 2022 from Andhra University, Visakhapatnam, India. His areas of interest include Microstrip patch Antenna, Flexible and Wearable Electronics, Signal Processing, Electromagnetic fields and VLSI system designs. He has published research papers on screen printed MSPA using silver nano ink. He has attended conferences and INUP hands on workshops conducted by IIT Bombay, and IIT Hyderabad. He has keen interests in microwave communication and VLSI system designs.

B. Nagasirisha received B.Tech. (ECE) degree from JNTU, Hyderabad, India in 2005, M.Tech. (Digital systems and computer electronics) from JNTUK, ananthapuramu in 2012, pursuing Ph. D in the field of Speech processing from Jawaharlal Nehru Technological University, Kakinada,Andhra Pradesh. She has more than 15 years of experience in teaching and 7 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.

T. Venkata Lakshmi received B.Tech. (ECE) degree from ANU University, Guntur, Andhra Pradesh, India in 2002, M.Tech. (Instrumentation and control systems) from JNTUK, Kakinada in 2008, pursuing Ph. D in the field of Low Power VLSI from Jawaharlal Nehru Technological University, Kakinada,Andhra Pradesh. She has more than 20 years of experience in teaching and 7 years in R&D. Her current research interests include Low Power VLSI Design, FinFETs, Signals & Systems, Communications, Digital signal processing, RADAR and Telemetry. She is working as Associate Professor of Electronics & Communication Engineering in Gudlavalleru Engineering College, an Autonomous NBA accredited College in Andhra Pradesh.