



# Deep-Dixon: Deep-Learning frameworks for fusion of MR T1 images for fat and water extraction

Snehal V. Laddha<sup>1</sup> · Rohini S. Ochawar<sup>2</sup> · Krushna Gandhi<sup>3</sup> · Yu-Dong Zhang<sup>4</sup>

Received: 20 February 2024 / Revised: 14 August 2024 / Accepted: 9 September 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

## Abstract

Medical image fusion plays a crucial role in understanding the necessity of medical procedures and it also assists radiologists in decision-making for surgical operations. Dixon has mathematically described a fat suppression technique that differentiates between fat and water signals by utilizing in-phase and out-of-phase MR imaging. The fusion of MR T1 images can be performed by adding or subtracting in-phase and out-phase images, respectively. The dataset used in this study was collected from the CHAOS grand challenge, comprising DICOM data sets from two different MRI sequences (T1 in-phase and out-phase). Our methodology involved training of deep learning models; VGG 19 and RESNET18 to extract features from this dataset to implement the Dixon technique, effectively separating the water and fat components. Using VGG19 and ResNet18 models, we were able to accomplish the image fusion accuracy for water-only images with EN as high as 5.70, 4.72, MI as 2.26, 2.21; SSIM as 0.97, 0.81; Qabf as 0.73, 0.72; Nabf as low as 0.18, 0.19 using VGG19 and ResNet18 models respectively. For fat-only images we have achieved EN as 4.17, 4.06; MI as 0.80, 0.77; SSIM as 0.45, 0.39; Qabf as 0.53, 0.48; Nabf as low as 0.22, 0.27. The experimental findings demonstrated the superior performance of our proposed method in terms of the enhanced accuracy and visual quality of water-only and fat-only images using several quantitative assessment parameters over other models experimented by various researchers. Our models are the stand-alone models for the implementation of the Dixon methodology using deep learning techniques. This model has experienced an improvement of 0.62 in EN, and 0.29 in Qabf compared to existing fusion models for different image modalities. Also, it can better assist radiologists in identifying tissues and blood vessels of abdominal organs that are rich in protein and understanding the fat content in lesions.

---

✉ Snehal V. Laddha  
laddhasv2@rknec.edu

<sup>1</sup> Department of Electronics Engineering, School of Electrical and Electronics Engineering, Ramdeobaba University, Nagpur 440013, Maharashtra, India

<sup>2</sup> Department of Electronics and Communication Engineering, School of Electrical and Electronics Engineering, Ramdeobaba University, Nagpur 440013, Maharashtra, India

<sup>3</sup> M. D. Radiology, Gandhi Sonography Clinic, Nagpur, India

<sup>4</sup> Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

**Keywords** Deep learning · Image fusion · MRI · Medical images · Liver disease

## 1 Introduction

MRI is commonly used to detect and diagnose diseases and to check the effectiveness of treatment. Diagnosis scenarios often require viewing the human body in multiple depths and with a variety of dimensions that are not seen by a single modality. By aggregating the data from various modalities, a new image can be formed that provides the expert with additional information about organs, tissues, and blood vessels and overcomes the limitations of a single modality [1]. A typical MR image is composed of T1-weighted and T2-weighted regions. T1 images depict the body's fat tissue, while T2 images emphasize the body's water and fat content. Medical image fusion is crucial for integrating information from various imaging modalities to offer a more detailed understanding of the internal structures, enhancing the accuracy of diagnosis and the effectiveness of the treatment. This integration is particularly significant in complex clinical decision-making and surgical procedures where detailed visualization of anatomy is required [2].

Fat signal suppression is an extremely useful diagnostic technique because water and fat are different in terms of their signal information. The Dixon technique explains a methodology that generates fused data such as water-only and fat-only images by simply adding and subtracting in-phase and out-phase MR T1 images from each other respectively [3]. This technique may also be referred to as inversion recovery and is a hybrid chemical shift-based technique that suppresses fat [4]. Also, it aids in the characterization of tissues, particularly adrenal gland tumors, bone marrow infiltration, and fatty liver. By generating clear water-only and fat-only images, the Dixon technique improves the visualization of different tissue types, making it easier to identify abnormalities and plan appropriate interventions.

With the fusion of medical images, the diagnosis becomes easier and the decision-making process is enhanced while reducing storage costs [5]. Numerous investigations have been carried out regarding the importance of the fusion of medical images in the past decade for image-guided procedures [6]. The improved image quality and details obtained through fusion techniques support radiologists and surgeons in accurately locating and assessing the abnormalities, thereby improving the patient's case studies.

AI-based multimodal fusion of medical images has been studied extensively by many researchers and work is ongoing in this field to design a robust, self-trained model that allows the extraction of more detailed information from radiological images to assist in disease diagnosis. These AI-driven approaches not only automate the fusion process but also enhance the precision and reliability of the fused images, making them indispensable tools in modern medical diagnostics and therapeutic planning.

However, existing fusion models have limitations such as inadequate image quality, limited robustness, and insufficient extraction of diagnostic details. There is a critical need for deep learning-based approaches to address these limitations by leveraging advanced algorithms to improve image fusion accuracy and quality [5–7]. With the advent in the field of artificial intelligence, the coalescing of in-phase and out-phase images of MR T1 modality to generate water-only and fat-only images can be done with the help of Machine Learning (ML) or Deep Learning (DL) based models to obtain a fused image that contains large amounts of information by combining two different modalities [5–7].

This paper's major contributions are summarized as follows:

- Implementation of the Dixon technique using VGG19 and ResNet18 models to generate water-only and fat-only images from MR T1 in-phase and out-phase images.
- Demonstration of improvement in image fusion accuracy and quality metrics compared to existing models.
- Comprehensive evaluation of the proposed method using various quantitative parameters such as SSIM, MI, Entropy, Qabf and Nabf, showing the enhanced visual quality of fused images.
- Identification of the potential of deep learning models to assist radiologists in better visualizing and characterizing tissues, aiding in more accurate diagnosis.

The paper is organized as follows. Section 2 discusses the related work in the field of image fusion and the research gap. In Section 3, Materials and methods related to the experimentation work are discussed. Section 4 emphasizes the proposed work. Section 5 presents the performance and provides an analysis of experimentation on different fusion techniques. Section 6 draws a conclusion.

## 2 Related work

The fusion of two or more modalities in medical imaging is important as multimodal images can give a clear understanding of organ structure, vessels entering and exiting organs, and abnormalities in organs to enable more accurate diagnosis, treatment, and surgery by combining minute details from different modalities. Recent advances in deep learning technologies led to significant developments in image classification, target recognition, and image fusion [8, 9].

Huang et al. experimented with CNN to fuse source image pixels to generate a weight map based on pixel activity information [7]. An operator that utilizes weighted fusion and multiple spatial frequency bands combines the source images. Comparative experimental results showed that, when compared to the source, the fusion method provided good visual effects and successfully conserved the exact structure information of the source images. This paper describes the results of fusion experiments on CT-MRI, MRI T1-T2, MRI-PET, and MRI-SPECT for brain image datasets. The Qabf and MI parameters for the matrix fusion were 0.44 and 1.09 respectively.

Deep learning (DL) architectures have improved the early identification of Alzheimer's disease (AD) and moderate cognitive impairment (MCI) using EEG. A study by Fouladi et al. utilized modified convolutional neural networks (CNN) and convolutional autoencoders (Conv-AE) to classify EEG data into AD, MCI, and healthy control (HC) groups, achieving accuracies of 92% and 89%, respectively. Their method surpasses traditional approaches, highlighting the effectiveness of DL in medical data analysis, which informs our use of CNNs for MR image fusion [9].

Kong et al. reported CNN's superiority for image fusion as an advantage due to its in-depth feature extraction characteristics on the brain MR image dataset [10]. Barachini et al. evaluated the efficacy of combining 18 F-DOPA PET and MRI for detecting liver metastases in neuroendocrine tumor patients [11]. In 11 patients, PET-MRI fusion significantly improved sensitivity in detecting liver metastases compared to MRI alone. The results highlight the enhanced diagnostic potential of multimodal imaging for optimized patient evaluation.

Retrospectively, Parsai et al. compared fused FDG PET-CT and MRI (PET-MRI) with FDG PET-CT and MRI alone for characterizing indeterminate focal liver lesions in patients with known malignancies. Fused PET-MRI significantly improved sensitivity (91.9%), specificity (97.4%), accuracy (94.7%), PPV (97.1%), and NPV (92.5%) in identifying malignant lesions, compared to PET-CT (55.6%, 83.3%, 66.7%, 83.3%, and 55.6%, respectively) and MRI alone (67.6%, 92.1%, 80%, 89.3%, and 74.5%, respectively), demonstrating superior diagnostic performance [12].

Yin et al. introduced a technique for fusing multimodal medical images utilizing non-subsampled shearlet transforms (NSST) [13]. In this method, NSST is initially applied to decompose the source images. The high-frequency components are then merged using a parameter-adaptive pulse-coupled neural network (PA-PCNN), while a strategy is devised to retain energy and extract details from the low-frequency components. Subsequently, the fused image is reconstructed by performing an inverse NSST on the combined high- and low-frequency components. This approach was evaluated by integrating medical images across four different modalities, including CT and MR, MR-T1 and MR-T2, MR and PET, as well as MR and single-photon emission CT, using over 80 combined source images from the Whole Brain Atlas database curated by Harvard Medical School. The findings indicated that this method outperforms nine other well-known medical image fusion techniques in both visual quality and objective assessments. For instance, in the fusion of MRI T1 and T2 images, the method achieved an entropy (EN) of 3.07, mutual information (MI) of 1.08, and a Qabf value of 0.42.

In 2021, Jiang et al. introduced a technique for medical image fusion that leverages Transfer Learning alongside L-BFGS (Limited-memory Broyden-Fletcher-Goldfarb-Shanno) optimization, specifically applied to CT and MRI brain image datasets from Wenzhou Medical University [14]. In this approach, Transfer Learning was utilized to extract essential features from the medical images, with the VGG16 model employed for this task. The fusion process was further enhanced by applying L-BFGS optimization, which helped maximize feature extraction, resulting in a mutual information (MI) value of 2.47.

Lahoud et al. demonstrated a method to combine features from multiple sources into a single image, by proposing a real-time method based on pre-trained neural networks VGG-19 [15]. In this study, deep feature maps derived from convolutional neural networks were used to merge the images. Multi-modal image fusion was driven by comparing these feature maps to create fusion weights. Experimentation was performed on the fusion of various image modalities including CT-MRI, MRI T1-T2, MRI-PET, MRI-SPECT, and CT-MRI-PET images collected from the Whole Brain Atlas. On MRI T1, T2 24 images, the fusion performance was evaluated on various parameters, and the EN was evaluated as 4.68, MI was evaluated as 9.35, SSIM evaluated as 0.87, Qabf was evaluated as 0.71, and Nabf scored as 0.004.

Zhang et al. proposed a technique for multi-modal brain image fusion guided by local extreme maps. This method involves iterative smoothing of source images, extraction of bright and dark features at multiple scales, and fusion using elementwise-maximum and minimum operations, producing detailed brain images suitable for both qualitative and quantitative analysis, with potential clinical applications [16]. Additionally, J. Ma et al. explored the increasing adoption of the Dixon method in musculoskeletal MRI practices [17]. Particularly with spin-echo sequences, broadening its utility in routine practice. It reviews the impact of fat-only images derived from Dixon sequences on interpreting musculoskeletal MRI, with a focus on bone marrow imaging, elucidating key principles influenced by fat content.

Electromagnetic tracking-based fusion imaging was investigated by Lee et al. [18] using real-time ultrasound and CT-MR images fused for hepatic interventions like biopsy and RFA. It was demonstrated that synchronized fused images enhanced lesion visibility, while three-dimensional Ultrasound (US) fusion aids in radiofrequency ablation (RFA). While contrast-enhanced US complements fusion imaging for identifying small hepatic lesions when fusion imaging alone is insufficient.

The retrospective study showcased by Schwarze et al. examines fusion imaging in evaluating hepatic and renal lesions, showcasing its technical success in 92 patients without adverse effects [19]. Fusion imaging effectively clarified initially indeterminate hepatic lesions (100%) and suspicious renal lesions (97%), demonstrating potential for comprehensive lesion assessment with high accuracy and safety.

Many researchers have performed experiments on the fusion of CT-MRI, MRI-PET, and MRI-SPECT image modalities. Research on multi-modality image fusion is still ongoing, according to the literature survey [7–19]. The Dixon method reveals that combining MRI T1 in-phase and out-phase images is crucial for effective fat quantification and suppression. Implementation of the Dixon method using deep learning methods has not been reported yet so far which gives scope for further experimentation. Deep learning (DL) has proven to be the most promising method for obtaining better multimodal fusion of biomedical images. In this work, DL-based methods have been experimented for the implementation of Dixon's theory.

## 3 Materials and methods

### 3.1 Dixon Methodology

The water and fat content in the human body contributes to the formation of MR images. Water molecules spin at a different rate than fat molecules, and the Dixon technique takes advantage of this difference. This method, proposed by Dixon, exploits the alternating in-phase and opposed-phase signals of water and fat molecules over time. By acquiring images at specific echo times when the water and fat protons are in-phase and out-of-phase, pure water and lipid images can be generated. Accurate separation of these components is essential for diagnosing and monitoring various conditions, including liver disease, tumors, and metabolic disorders. The water-only images help in assessing liver lesions and fibrosis by highlighting tissue characteristics, while fat-only images can identify hepatic steatosis. Also, differentiating between water and fat content aids in the precise characterization of tumors, improving treatment planning and monitoring. However, fat-only images are valuable for diagnosing and managing conditions like fatty liver disease and lipomas.

The Dixon method, is an MRI sequence utilizing chemical shift imaging to provide consistent fat suppression. This technique involves acquiring in-phase and out-of-phase images, which are then reconstructed to produce water-only (WO) and fat-only (FO) images. The Dixon method has been gaining popularity due to its several advantages over other fat suppression techniques. It provides more uniform suppression of fat signals and is less affected by artifacts. Moreover, the Dixon method can be combined with various sequence types (e.g., spin echo, gradient echo) and weightings (e.g., T1, T2, and proton density), offering significant versatility in imaging. A notable benefit is that it can generate images with and without fat suppression from a single acquisition, allowing for the quantification of fat content, not just its presence [20–22].

In MRI T1 (in-phase-out-phase) sequence, the difference between fat and water protons is used to suppress fat. Signals are acquired twice: once when the protons in the fat and water are out of phase (the excitement protons are going back to their initial places), and once when they are in phase. Water and fat are separated after in-phase and out-of-phase images have been post-processed. The in-phase image can then be added to the out-of-phase image to create a water image, whereas the out-of-phase image can be subtracted from the in-phase image to produce a fat image [20, 21]. When using 1.5 Tesla devices, the protons in water and fat flow in opposite directions, with an in-phase duration of 4.6 milliseconds and an out-phase time of 2.3 milliseconds. To suppress fat signals, these frequencies have to be subtracted from Time of Echo (TE) values.

Simultaneous acquisition of both in-phase and opposed-phase images enables two distinct mathematical combinations of these images.

$$\text{Fat Only} = \text{In} - \text{Phase} - \text{Opposed} - \text{Phase} = (\text{Water} + \text{Fat}) - (\text{Water} - \text{Fat}) \quad (1)$$

$$\text{Water Only} = \text{In} - \text{Phase} + \text{Opposed} - \text{Phase} = (\text{Water} + \text{Fat}) + (\text{Water} - \text{Fat}) \quad (2)$$

Analyzing the fat content of lesions is made possible by using this sequence. Out-phase images are characterized by black borders around the organs because of the abrupt shift in water and fat content at the organ borders, which neutralizes the received signal [22]. Dixon method's primary benefit is that it offers numerical data on the proportions of fat and water in the same order that is utilized for qualitative assessment [20]. As a result of the Dixon technique, the fat-water fraction can be measured with a greater region of interest and at a higher spatial resolution than it can be measured with spectroscopy [21]. Generally, this concept may be used to suppress or quantify fat in a variety of MR pulse sequences. For extensive field-of-view imaging applications involving the liver and extremities, Dixon's approach is therefore well suited. Using this method, fat-suppressed and non-fat-suppressed images can be obtained from a single acquisition. It has been applied extensively with sequences that are sensitive to fluid [21].

Thus, Dixon technique is used to suppress fat since it generates fat-suppressed images (WO images) with superior signal-to-noise ratios than short-tau inversion recovery (STIR) and is more resistant to field inhomogeneities than chemical shift selective (CHESS) approaches. The Dixon method utilizes the chemical shift between water and fat to achieve uniform fat suppression even in the presence of magnetic field distortions. Compared to STIR, which often suffers from a lower signal-to-noise ratio (SNR) due to its inversion recovery process that reduces overall signal intensity, Dixon typically delivers a higher SNR. This is because the Dixon method preserves the full signal from both water and fat, enhancing image quality. Due to the versatility of the Dixon method, it can be applied to various imaging scenarios, including complex anatomical regions and large field-of-view imaging, and can be integrated with different MRI sequences and weightings. This feature allows the Dixon method to be applied across a broad spectrum of clinical contexts. Additionally, the Dixon method facilitates both qualitative fat suppression and quantitative fat analysis, providing comprehensive diagnostic information and improving its overall utility [20–22, 28]. When integrated with T1-weighted sequences, the Dixon method improves the visualization of fat-containing structures and facilitates the identification of lesions with fat components. In T2-weighted sequences, it enhances the contrast between fat and water, aiding in the detection of edema or inflammation. However, the Dixon method is prone to motion artifacts, which can be managed through breath-hold imaging and advanced

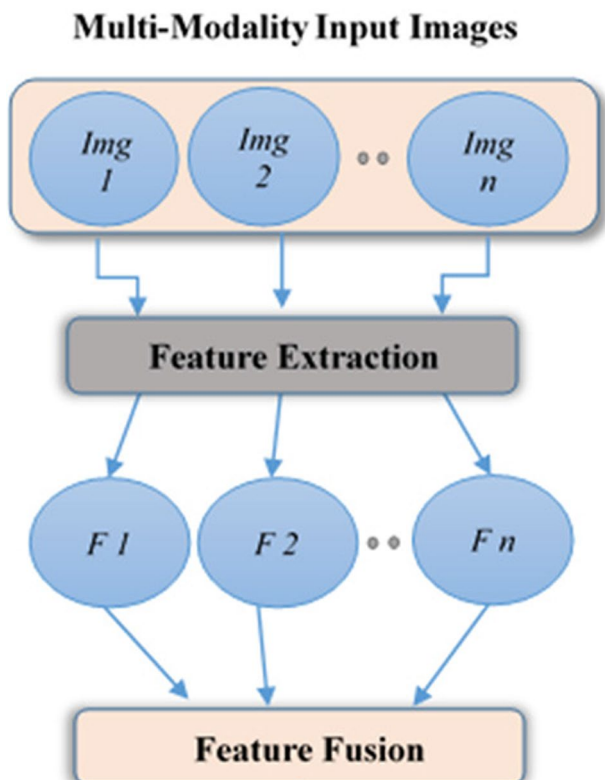
motion correction techniques. Its effectiveness can also be influenced by the strength of the magnetic field: higher field strengths generally enhance the signal-to-noise ratio but may increase susceptibility to artifacts. Addressing these issues requires careful optimization of echo times and the use of advanced reconstruction algorithms [23].

### 3.2 Fusion

Selection of the fusion algorithm plays a significant role as the resulting image from the fusion of distinct modalities must possess additional information as compared to the original image, avoiding noise, misalignment, and needless artifacts. The image fusion abstraction levels are classified into three categories: pixel level, feature level, and decision level. Based on the literature review done on various fusion models for different applications and considering the merits, demerits, and challenges of each fusion model; we have selected a feature-level fusion model for our experimentation for the implementation of the Dixon technique using a deep-learning approach.

The feature-level fusion algorithm involves the extraction of features from images. Fig. 1 represents the feature-level fusion model for multimodal images. Extracting regions at the feature level from images provides more information and a better understanding of the content than pixels. In addition, a literature review [23] indicated that the feature level offered an advantage in compressing the information and processing it in real-time.

**Fig. 1** Feature level fusion method using multimodal images





## 4 Proposed work

Our study proposes a fusion approach for the implementation of the Dixon technique. According to the Dixon methodology, when T1 in-phase and T1 out-phase images are added together, the resultant images are referred to as water-only and when they are subtracted, they are called fat-only. Our model for the generation of MRI water-only and fat-only images is shown in Fig. 2. Initially, Two different source images, T1 in-phase and T1 out-phase are pre-processed by converting them from DICOM format to lossless PNG and resizing them to a consistent resolution of  $256 \times 256$  pixels. These standardized images are then fed into deep learning models independently. Representative features from each image are extracted, fusion weights are generated according to their activation levels with weights computed at layer  $l = 1$ , and then Gaussian smoothing is performed on the weight maps to remove artifacts around the edges of both modalities, which can arise from abrupt transitions between different image regions. By smoothing the weight maps, the technique also helps correct minor misregistrations between the images, thereby improving feature alignment. Additionally, Gaussian smoothing enhances fusion quality by minimizing noise and sharp discontinuities, resulting in clearer and more accurate final images. This leads to improved diagnostic outcomes as the final fused images are more reliable and provide better clarity for assessing and interpreting medical conditions.

A convolutional neural network VGG19 has 19 layers and is considerably deeper than many other CNN architectures. The selection of VGG19 for our research work is driven by its significant depth, allowing for the extraction of a comprehensive set of hierarchical features from input images, which is essential for precise image fusion. VGG19's established performance across various datasets highlights its reliability and robustness, making it well-suited for complex tasks [24, 29]. Its ability to utilize pre-trained models through transfer learning is particularly beneficial, enabling effective training even with smaller medical datasets. VGG19's extensive feature representation and versatility validate its application, ensuring an accurate and effective fusion of MR images in the proposed method.

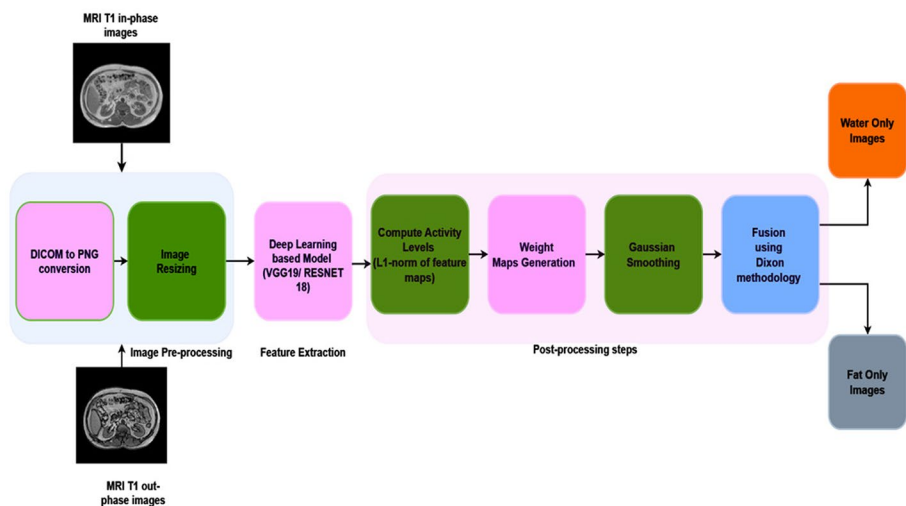


Fig. 2 Water-only and fat-only image generation using MRI T1 in-phase and out-phase images



In our approach, VGG19 is employed to extract features from MR T1 in-phase and T1 out-of-phase images. The images are first preprocessed before being input into the VGG19 network. The network’s early layers (Conv1\_1 and Conv1\_2) capture low-level features such as edges and textures, followed by max-pooling to reduce dimensionality. Intermediate layers (Conv2\_1, Conv2\_2, Conv3\_1, Conv3\_2) extract mid-level features like object parts, and further max-pooling consolidates these features. Deeper layers (Conv4\_1, Conv4\_2, Conv5\_1, Conv5\_2, Conv5\_3) focus on high-level features, including global patterns and shapes. After feature extraction, images are processed further for post-processing steps as shown in Fig. 2.

Given  $\mathcal{N}$  pre-registered source images represented as  $I_n | n \in: \{1, 2, \dots, N\}$  and a pre-trained convolutional neural network with  $\mathcal{L}$ -layers that has  $C_l$  output channels in each layer  $l$ . The feature map of the  $n$ -th image at the  $l$ -th layer of the network (post-ReLU activation) is denoted as  $f_n^{c,l}$  for the  $c$ -th feature map. This feature map is computed as:

$$f_n^l = \max(0, F_l(I_n)) \tag{3}$$

Here,  $F_l(\cdot)$  represents the processing applied by the network layers onto the input image up to layer  $l$ . The ReLU operation is represented by the function  $\max(0, \cdot)$  function. For each feature map,  $\hat{f}_n^l$  denotes the  $l_1$ -norm calculated across the  $C_l$  channels of the feature maps at layer  $l$  as given by:

$$\hat{f}_n^l = \sum_{c=0}^{C_l} \|f_n^{c,l}\|_1 \tag{4}$$

This quantifies the level of activity related to the input image at layer  $l$ . We have performed feature extraction from images, using weights computed at layer  $l = 1$ . For each image  $n$ , feature maps are extracted for all  $\mathcal{L}$  layers, resulting in the set  $\{\hat{f}_n^l | l \in \mathcal{L}\}$ . These feature maps are then utilized to generate  $n$  weight maps for each layer  $l$ , reflecting the contribution of each image to every pixel. In our approach, the softmax function is used to produce these weight maps as follows:

$$W_n^l = \frac{e^{\hat{f}_n^l}}{\sum_{j=1}^N e^{\hat{f}_j^l}} \tag{5}$$

Where,  $e^{(\cdot)}$  is the exponentiation with base  $e$ .

To eliminate artifacts around the margins of both modalities and adjust for minor mis-registration, the weight maps are smoothen using a Gaussian method. The smoothing is applied with a standard deviation ( $\sigma$ ) defined as:

$$\sigma = 0.01 \sqrt{w^2 + h^2} \tag{6}$$

In this context,  $w$  and  $h$  represent the width and height of the weight maps, respectively.

For the generation of water-only and fat-only images using the Dixon methodology, pixel-wise addition and subtraction operations are performed on T1 in-phase and out-phase images, which are processed independently following Gaussian smoothing. The addition operation, referred to as a fusion of two modality images, produces water-only images by combining the in-phase and out-phase images pixel by pixel. This fusion operation integrates information from both image modalities, enhancing the resultant image’s diagnostic value. Conversely, the pixel-wise subtraction of these images yields fat-only images, further contributing to comprehensive diagnostic imaging.

$$I_{\text{water-only}} = I_{\text{in-phase}} + I_{\text{out-phase}} \quad (7)$$

$$I_{\text{Fat-only}} = I_{\text{in-phase}} - I_{\text{out-phase}} \quad (8)$$

The deep learning model VGG19 was trained using the Mean Squared Error loss function and the Adam optimizer, with a learning rate of 0.001 and a batch size of 16, over 50 epochs with early stopping. The training was performed on Google Colab using a Tesla K80 GPU. Each epoch took approximately 45 s to 1 min, leading to a total training time of about 50 min for all 50 epochs. This setup was used to leverage the parallel processing capabilities of GPU, which significantly accelerates the training process and enhances computational efficiency compared to CPU-based training, which would typically require several hours for the same task. Also, the hyperparameter tuning was conducted via grid search, and model performance was evaluated using various evaluation metrics mentioned in the next sections as Entropy (EN), Structural similarity index (SSIM), Mutual information (MI), Edge-Based Similarity Measure (Qabf), Nabf. Thus in our experimentation, the fat-only and water-only images are obtained using the VGG19 model for feature extraction and applying the Dixon methodology to achieve fusion on multi-modal input images. The fusion is evaluated using various fusion evaluation metrics given in Section 4.1. The Dixon method implementation is also carried out with RESNET 18 architecture. Finally, fusion experimentation is quantified for both models, and results are presented in Section 5.2.

## 4.1 Fusion evaluation Metric

Medical image fusion of MRI images is evaluated with a few statistical parameters as mentioned further:

### 4.1.1 Entropy (EN)

The amount of information that is available in the source and fused images separately is measured by entropy [25]. The rich information content is indicated by the high entropy value.

$$EN = - \sum_{i=0}^{l-1} p(i) \log_2 p(i) \quad (9)$$

Where,  $p(i)$  indicates the probability of pixels gray level with the range  $[0, \dots, l-1]$ .

### 4.1.2 Structural similarity index (SSIM)

The Structural similarity index (SSIM) is a perceptual metric used to evaluate the image quality degradation. It quantifies the extent to which the salient information is preserved in the fused image, with values falling between  $[-1, 1]$ . Higher values reflect greater similarity between the original and fused images [25].

$$SSIM_{(F,I)} = \frac{((2\mu_F \mu_I + c1)X(2\sigma_{FI} + C2))}{((\mu_F^2 + \mu_I^2 + C1)X(\sigma_F^2 + \sigma_I^2 + C2))} \quad (10)$$

F represents the fused image, while I denotes the input image, and  $\mu F$  and  $\mu I$  refer to the average intensity values of F and I, respectively. Variance is represented by  $\sigma F$  and  $\sigma I$  for images F and I, while  $\sigma FI$  indicates their covariance, and C1 and C2 are the constants.

### 4.1.3 Mutual information (MI)

Mutual information (MI) measures the contribution of the input images to the fused images’s information content. As the amount of detail and texture in the fused image increases, the MI value also increases [26]. The calculation for MI is defined in terms of two input images (XA, XB) and a fused image (XF) as follows:

$$MI = I(X_A;X_F) + I(X_B;X_F) \tag{11}$$

$$I(X_R;X_F) = \sum_{u=1}^L \sum_{v=1}^L h_{R,F}(u, v) \log_2 \frac{h_{R,F}(u, v)}{h_R(u)h_F(v)} \tag{12}$$

Let R represent the reference image and F a fused image,  $h_{R, F}(u, v)$  refers to the joint gray level histogram of  $X_R$  and  $X_F$ , while  $h_R(u)$  and  $h_F(v)$  denotes the normalized gray-level histograms of  $X_R$  and  $X_F$ , respectively.

### 4.1.4 Edge-based similarity measure ( $Q_{abf}$ )

Edge-Based Similarity Measure is defined as:

$$Q^{AB/F} = \frac{\sum_{i=1}^M \sum_{j=1}^N [Q^{AF}(i, j)w^x(i, j) + Q^{BF}(i, j)w^y(i, j)]}{\sum_{i=1}^M \sum_{j=1}^N [w^x(i, j) + w^y(i, j)]} \tag{13}$$

The input and fused pictures are denoted by A, B, and F, respectively. QAF and QBF have identical definitions, which are as follows:

$$Q^{AF}(i, j) = Q_g^{AF}(i, j)Q_\alpha^{AF}(i, j), \tag{14}$$

$$Q^{BF}(i, j) = Q_g^{BF}(i, j)Q_\alpha^{BF}(i, j), \tag{15}$$

Where the edge strength and orientation preservation values for images A and B, respectively, at location (i, j), are denoted by  $Q_g^{*F}$  and  $Q_\alpha^{*F}$ . For optimal fusion, Qabf’s dynamic range, which is [0, 1], should be as near to one as feasible.

### 4.1.5 Nabf

It gauges the level of noise or artifacts that were introduced to the fused image that wasn’t in the original pictures. A lower value denotes less noise and artifacts in the fused image.

## 5 Experiments

The fusion of images at various depths of the network can be computed to understand the effect of multiple layers in a deep neural network on the output fusion. In this work using the same methodology, MRI T1 in-phase and out-phase image fusion is performed using VGG19 and ResNet18 to generate water-only images for fat suppression analysis and fat-only images for fat quantification analysis, particularly useful for distinguishing between different tissue types and identifying organ boundaries.

### 5.1 Database

We collected the dataset of MRI T1 in-phase and out-phase images from the CHAOS grand challenge [27]. The dataset comprises 120 DICOM files from two different MRI sequences: 40 files from T1 in-phase imaging and 40 files from T1 out-phase imaging. These sequences are often used to scan the abdomen utilizing various radiofrequency pulse and gradient combinations. Data were collected using a 1.5T Philips MRI scanner, generating 12-bit DICOM images. These images were subsequently converted to PNG format without loss of quality. All images, regardless of their original dimensions, were standardized to a resolution of  $256 \times 256$  pixels. T1 in-phase gives blood and tissue information while the T1 out-phase highlights the borders of organs. Because the distribution of fat and water undergoes abrupt changes at the boundaries of organs, canceling out the collected signal, the organ borders look black in out-of-phase images. Around 1300 MRI T1 (In and out phase) images were used for fusion, specifically focusing on abdominal imaging. The dataset includes critical anatomical structures such as the liver, kidneys, spleen, and other abdominal organs.

### 5.2 Results

Figure 3 shows the results of VGG 19 Model experimentation on abdominal MR images. Input images indicate T1 in-phase and out-phase images for various test cases. Correspondingly output images indicate water-only and fat-only content as shown in the same row.

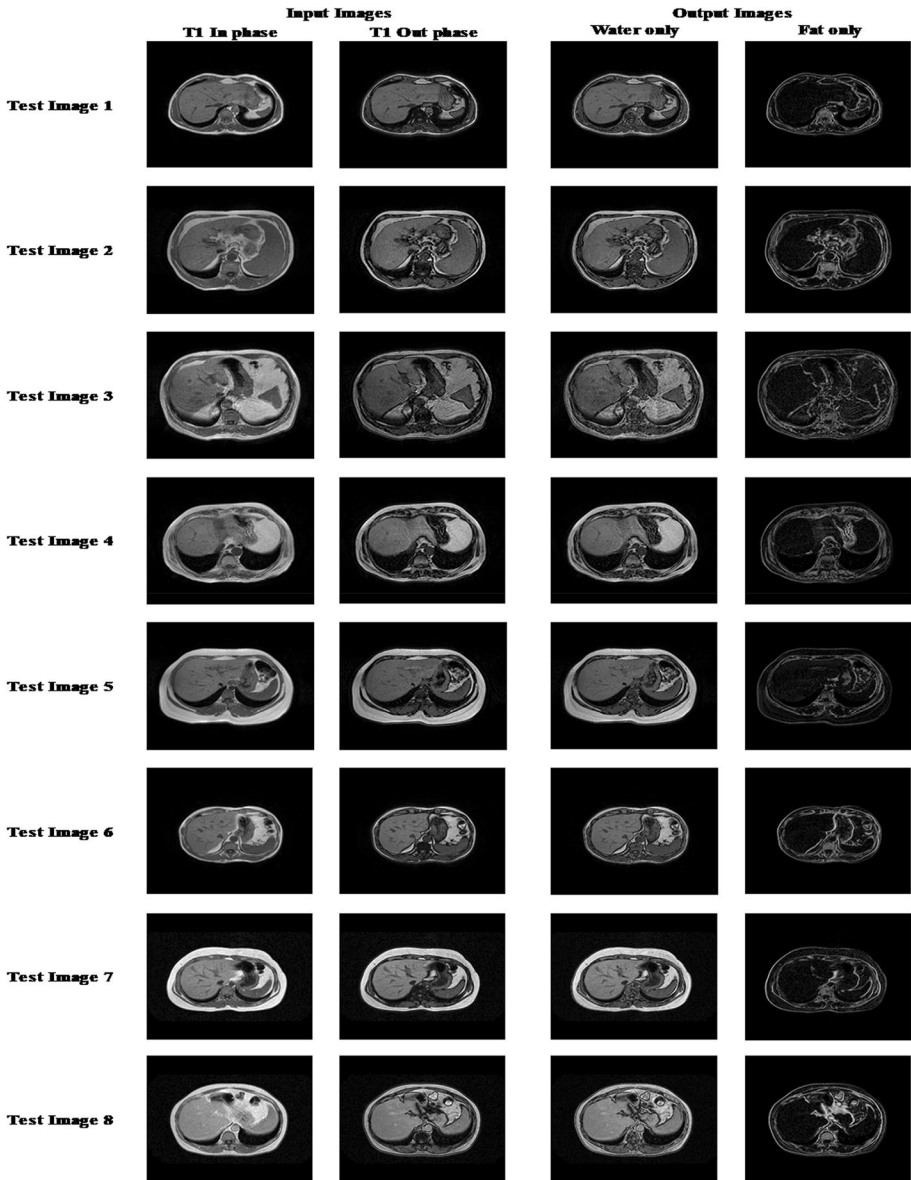
We have compared the results of our experimentation with existing work proposed by various researchers who have done multimodal image fusion on CT, MRI, and PET images and it has been proved that our model outperformed in all the evaluation metrics.

Figure 4 illustrates the fusion of T1 in-phase and out-phase images using ResNet 18 Model using abdominal MR images. The resultant water-only and fat-only images are displayed in the respective row.

A comparative analysis of various evaluation metrics for the resultant water-only images using the VGG19 and ResNet18 model for sample images is shown in Table 1. The average score of evaluation parameters for the water-only image fusion dataset on the complete dataset is shown in Table 2. Fig. 5. Illustrates the performance parameters for MRI T1 water-only images.

Table 3 depicts a comparative analysis of various evaluation metrics for the resultant fat-only images using VGG19 and ResNet18 models for sample images. Table. 4 shows the average score of evaluation parameters for fat-only image fusion. Fig. 6. Illustrates the performance parameters for MRI T1 fat-only images.

From our experimentations, we have achieved the image fusion accuracy for water-only images with EN as 5.70, 4.72, MI as 2.26, 2.21; SSIM as 0.97, 0.81; Qabf as 0.73,



**Fig. 3** Implementation of VGG 19 Model using abdominal MR images (T1in-phase and out-phase) for water-only and fat-only image generation

0.72; Nabf as low as 0.18, 0.19 using VGG19 and ResNet18 models respectively. For fat-only images we have achieved EN as 4.17, 4.06; MI as 0.80, 0.77; SSIM as 0.45, 0.39; Qabf as 0.53, 0.48; Nabf as low as 0.22, 0.27 using VGG19 and ResNet18 models respectively. Dixon technique implementation with this new approach using MRI multi-modal fusion achieved better results using VGG19 model than ResNet18.

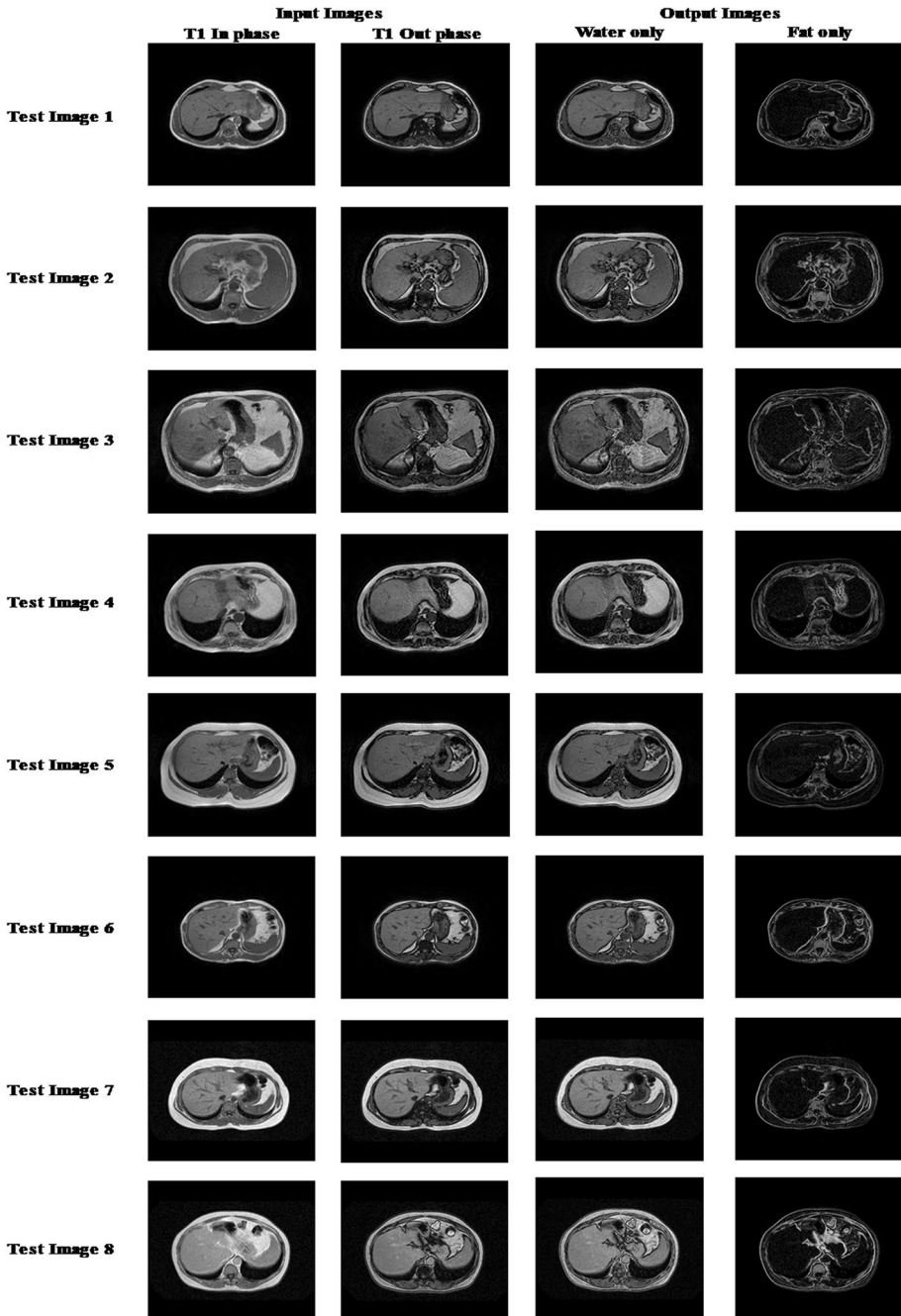


Fig. 4 Implementation of ResNet18 Model on abdominal MR images (T1 in-phase and out-phase) for water-only and fat-only image generation

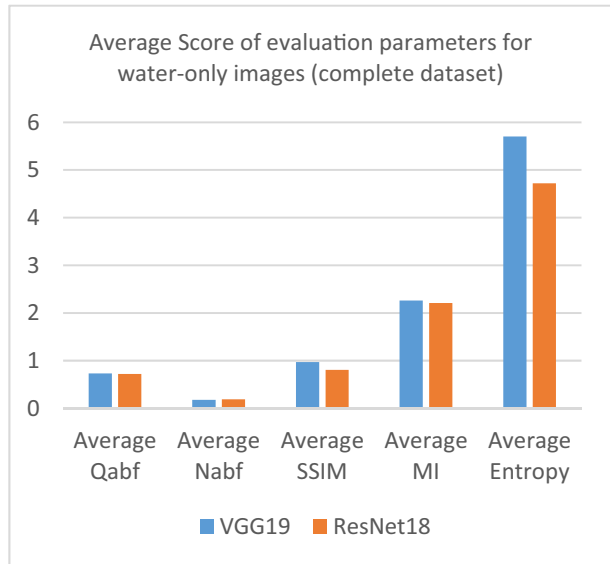
**Table 1** Comparative analysis of various evaluation parameters for water-only images with sample data using VGG19 and RESNET 18

Test Image numbers		Indices												
		EN			MI			SSIM			Nabf			Qabf
		VGG 19	ResNet 18	VGG 19	ResNet 18	VGG 19	ResNet 18	VGG 19	ResNet 18	VGG 19	ResNet 18	VGG 19	ResNet 18	
Test 1		5.07	5.96	2.39	2.37	0.92	0.78	0.18	0.13	0.73	0.73	0.73	0.73	
Test 2		5.08	5.93	2.33	2.45	0.96	0.8	0.13	0.08	0.74	0.74	0.74	0.74	
Test 3		4.33	5.05	2.15	2.2	0.95	0.84	0.27	0.18	0.75	0.75	0.75	0.75	
Test 4		4.62	4.99	2.22	2.22	0.93	0.84	0.17	0.08	0.68	0.68	0.68	0.68	
Test 5		5	5.96	2.39	2.37	0.88	0.78	0.18	0.13	0.73	0.73	0.73	0.73	
Test 6		5.08	5.93	2.3	2.16	0.92	0.74	0.13	0.13	0.66	0.66	0.66	0.66	
Test 7		4.99	5.41	2.4	2.16	0.85	0.81	0.21	0.23	0.72	0.72	0.72	0.7	
Test 8		4.85	5.13	2.29	2.06	0.86	0.81	0.23	0.18	0.68	0.68	0.68	0.68	
Test 9		4.96	5.3	2.4	2.14	0.92	0.8	0.2	0.18	0.7	0.7	0.7	0.7	
Test 10		4.12	4.8	2.15	2.24	0.9	0.86	0.3	0.3	0.76	0.76	0.76	0.76	
Test 11		5.01	5.82	2.31	2.33	0.91	0.81	0.08	0.06	0.72	0.72	0.72	0.72	
Test 12		4.97	5.38	2.38	2.07	0.82	0.79	0.18	0.2	0.69	0.69	0.69	0.69	
Test 13		4.92	5.25	2.36	2.12	0.9	0.81	0.21	0.26	0.73	0.73	0.73	0.73	
Test 14		4.23	4.78	2.11	2.15	0.92	0.85	0.25	0.19	0.77	0.77	0.77	0.76	
Test 15		4.62	5.07	2.24	2.2	0.85	0.81	0.26	0.3	0.73	0.73	0.73	0.73	
Test 16		4.99	5.39	2.39	2.16	0.86	0.81	0.21	0.23	0.7	0.7	0.7	0.7	
Test 17		4.21	4.74	2.27	2.26	0.88	0.88	0.13	0.28	0.78	0.78	0.78	0.78	
Test 18		4.04	4.69	2.26	2.22	0.91	0.89	0.13	0.17	0.77	0.77	0.77	0.77	
Test 19		5.01	5.81	2.28	2.34	0.92	0.78	0.11	0.13	0.74	0.74	0.74	0.74	
Test 20		4.16	4.69	1.75	2.13	0.84	0.84	0.27	0.23	0.75	0.75	0.75	0.75	



**Table 2** Average score of evaluation parameters for water-only images (complete dataset)

Parameters	VGG19	ResNet18
Average Qabf	<b>0.73</b>	0.72
Average Nabf	<b>0.18</b>	0.19
Average SSIM	<b>0.97</b>	0.81
Average MI	<b>2.26</b>	2.21
Average Entropy	<b>5.70</b>	4.72

**Fig. 5** An illustration of the performance parameters for MRI T1 water-only images

### 5.3 Benchmarking

Although our methodology is a standalone technique for Dixon implementation using DL based approach, we have done a comparative analysis of the quantitative metric values of the image fusion methods with previously published research on the fusion of different modalities. The metric values of the various methods on the fusion of each modality are listed in Table 5.

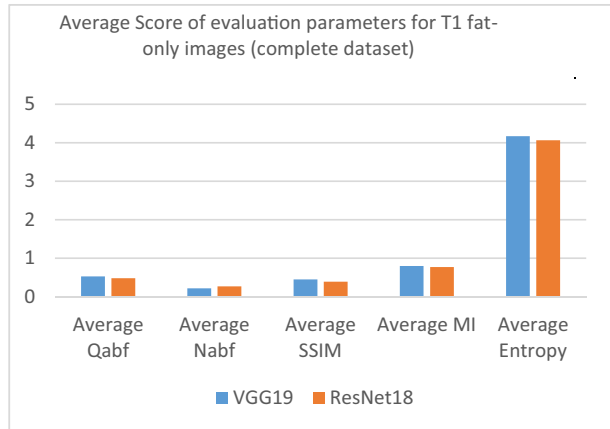
With VGG19 model, we have achieved better results than previously published work in terms of EN, SSIM, and Qabf over a dataset of more than 1300 images of MRI T1 in-phase and out-phase modality for water-only image generation. The results of the experimentations were validated by radiologist Dr. Krushna Gandhi, who confirmed the accuracy and reliability of our method. Specifically, she confirmed that our method produced clear and accurate images with high contrast and minimal artifacts.

**Table 3** Comparative analysis of various evaluation parameters for fat-only images with sample data using VGG19 and RESNET 18

Test Image numbers	Indices														
	EN			MI			SSIM			Nabf			Qabf		
	VGG 19	ResNet 18	ResNet 18	VGG 19	ResNet 18	ResNet 18	VGG19	ResNet18	ResNet18	VGG19	ResNet18	ResNet18	VGG19	ResNet18	ResNet18
Test 1	4.54	4.82	0.89	0.94	0.89	0.31	0.42	0.42	0.21	0.12	0.12	0.72	0.44	0.44	
Test 2	4.51	4.32	0.76	0.82	0.76	0.32	0.42	0.42	0.25	0.19	0.19	0.71	0.46	0.46	
Test 3	3.87	3.41	0.8	0.82	0.8	0.44	0.51	0.51	0.24	0.23	0.23	0.49	0.48	0.48	
Test 4	4	3.9	0.69	0.72	0.69	0.41	0.44	0.44	0.37	0.32	0.32	0.49	0.49	0.49	
Test 5	4.54	4.54	0.89	0.94	0.89	0.31	0.42	0.42	0.2	0.12	0.12	0.68	0.44	0.44	
Test 6	4.64	4.32	0.96	1	0.96	0.33	0.41	0.41	0.19	0.22	0.22	0.53	0.49	0.49	
Test 7	4.39	4.21	0.78	0.8	0.78	0.33	0.39	0.39	0.39	0.26	0.26	0.51	0.52	0.52	
Test 8	4.29	4.18	0.74	0.75	0.74	0.38	0.41	0.41	0.36	0.21	0.21	0.58	0.49	0.49	
Test 9	4.35	4.35	0.74	0.75	0.74	0.35	0.39	0.39	0.41	0.28	0.28	0.51	0.5	0.5	
Test 10	3.66	3.66	0.63	0.68	0.63	0.44	0.5	0.5	0.14	0.25	0.25	0.59	0.59	0.59	
Test 11	4.41	4.13	0.73	0.77	0.73	0.34	0.43	0.43	0.2	0.15	0.15	0.64	0.44	0.44	
Test 12	4.39	4.39	0.77	0.78	0.77	0.34	0.39	0.39	0.42	0.28	0.28	0.5	0.49	0.49	
Test 13	4.3	4.2	0.86	0.75	0.86	0.36	0.73	0.73	0.41	0.25	0.25	0.52	0.49	0.49	
Test 14	3.75	3.75	0.68	0.71	0.68	0.61	0.49	0.49	0.21	0.22	0.22	0.48	0.48	0.48	
Test 15	4.22	4.22	0.86	0.88	0.86	0.39	0.43	0.43	0.34	0.29	0.29	0.52	0.51	0.51	
Test 16	4.39	4.19	0.78	0.8	0.78	0.33	0.39	0.39	0.39	0.26	0.26	0.31	0.52	0.52	
Test 17	3.62	3.58	0.67	0.7	0.67	0.45	0.5	0.5	0.15	0.14	0.14	0.43	0.42	0.42	
Test 18	3.36	3.12	0.66	0.7	0.66	0.49	0.55	0.55	0.15	0.16	0.16	0.43	0.43	0.43	
Test 19	4.54	4.31	0.93	0.97	0.93	0.33	0.42	0.42	0.18	0.3	0.3	0.6	0.46	0.46	
Test 20	3.75	3.68	0.75	0.79	0.75	0.61	0.51	0.51	0.22	0.23	0.23	0.52	0.5	0.5	

**Table 4** Averagescore of evaluation parameters for T1 fat-only images (complete dataset)

Parameters	VGG19	ResNet18
Average Qabf	<b>0.53</b>	0.48
Average Nabf	<b>0.22</b>	0.27
Average SSIM	<b>0.45</b>	0.39
Average MI	<b>0.80</b>	0.77
Average Entropy	<b>4.17</b>	4.06

**Fig. 6** An illustration of the performance parameters for MRI T1 fat-only images

## 6 Conclusion

Dixon method implementation using multimodal image fusion, creates a water-only image that is helpful for fat suppression. We have introduced a completely new approach for generating water-only and fat-only MR images using a DL-based fusion technique that quantifies Dixon's method. Moreover, with our approach, it is possible to visualize the water content and fat from images by radiologists using DL-based methods with water-only and fat-only images.

Our experiments show that enhancement occurs when the fusion is performed at the feature level, retaining unique information about organs from the source images. The unique information is objectively defined as the feature maps extracted by DL-based models. On the other hand, the subjective information of the fused image is preserved with high-quality texture details. It has been demonstrated through comprehensive test findings indicating that the suggested approach clearly outperforms other multimodal fusion methods. Our deep learning-based models, VGG19 and ResNet18, achieved high fusion accuracy for water-only and fat-only images, with entropy (EN) values of 5.70 and 4.72 for water-only, and 4.17 and 4.06 for fat-only images, respectively. The visual quality of the fused images has improved as demonstrated by high structural similarity index (SSIM) values of 0.97 and 0.81, and favorable Qabf scores of 0.73 and 0.72. Our models outperformed existing fusion techniques, providing detailed information on tissues and blood vessels, which enhances the radiologist's ability to identify protein-rich tissues and understand fat content in lesions.

**Table 5** Comparative analysis of fusion evaluation metrics for benchmarking

Modality	Methodology	EN	SSIM	Qabf
MRI T1-T2 [13]	Non subsampled shearlet transforms (NSST)	3.07	-	0.42
CT-MRI [16]	Local extreme map	4.27	0.95	0.71
PET-MRI [16]	Local extreme map	4.89	0.94	0.73
SPECT-MRI [16]	Local extreme map	4.97	0.95	0.68
CT-MRI [17]	Conditional Generative Adversarial Network featuring two discriminators	4.99	0.65	0.59
PET-MRI [17]	Conditional Generative Adversarial Network featuring two discriminators	4.98	0.38	0.32
SPECT-MRI [17]	Conditional Generative Adversarial Network featuring two discriminators	5.66	0.54	0.28
MRI T1 inphase-Outphase	Transfer learning with ResNet18 (Proposed model for fat-only images)	4.06	0.39	0.48
<b>MRI T1 inphase-Outphase</b>	<b>Transfer learning with VGG-19 (Proposed model for fat-only images)</b>	<b>4.17</b>	<b>0.45</b>	<b>0.53</b>
MRI T1 inphase-Outphase	Transfer learning with ResNet18 (Proposed model for water-only images)	4.72	0.81	0.71
<b>MRI T1 inphase-Outphase</b>	<b>Transfer learning with VGG-19 (Proposed model for water-only images)</b>	<b>5.70</b>	<b>0.97</b>	<b>0.73</b>

Although promising, our deep learning models for fusing MRI T1 in-phase and out-phase images exhibit limitations related to image quality dependency and computational complexity, especially in clinical settings where resources may be limited. These limitations highlight the need for further research to strengthen the robustness of deep learning models against variations in image quality and to optimize their computational efficiency for practical medical applications. Future research will expand these models to integrate with other modalities like CT and PET, while incorporating clinical data to enhance diagnostic accuracy. To further improve diagnostic accuracy, the feasibility of integrating additional information such as clinical data or patient-specific features into the fusion process is worth investigating. Incorporating factors like patient history, lab results, and genetic information could provide a more comprehensive understanding of the patient's condition and enhance the decision-making process. By combining these data sources with multimodal imaging, we could develop a more holistic approach to medical diagnosis and treatment planning.

**Acknowledgements** We would like to express our sincere gratitude to Dr. Madhusudan Sarda, Consultant Radiologist at Nagpur Scan & Research Centre, for generously providing the dataset that was critical to this study. His support played a crucial role in the success of our research. His radiological expertise and valuable feedback significantly contributed to the development and refinement of our work. We are deeply appreciative of his continual guidance and encouragement throughout the project.

**Authors' contribution** S.V. Laddha: Conceptualization, Methodology, Investigation, Formal analysis, Data curation, Writing- Original draft preparation. R. S. Ochawar: Supervision, Writing- Reviewing and Editing, project administration. Dr. Krushna Gandhi: Validation and dataset support. Yu-Dong Zhang: Reviewing and Editing.

**Funding** None.

**Data availability** The dataset used in the experiment is derived from CHAOS - Combined (CT-MR) Healthy Abdominal Organ Segmentation Challenge Data (Version v1.03).

## Declarations

**Competing interests** The authors certify that there is NO conflict of interest about this work.

## References

1. Wang K, Zheng M, Wei H, Qi G, Li Y (2020) Multi-Modality Medical Image Fusion Using Convolutional Neural Network and Contrast Pyramid. *Sensors* 20(8):2169
2. Hermessi H, Mourali O, and Ezzeddine Zagrouba (2021) Multimodal Medical Image Fusion Review: theoretical background and recent advances. *Sig Process* 183:108036
3. Dixon WT (1984) Simple proton spectroscopic imaging. *Radiology* 153(1):189–194
4. Pezeshk P, Alian A, Chhabra A (2017) Role of chemical shift and Dixon based techniques in musculoskeletal MR imaging. *Eur J Radiol* 94:93–100
5. Chen H, Liang LJM, Liu F, Yang S, Biao Hou (2019) Fast Unsupervised Deep Fusion Network for Change Detection of Multitemporal SAR Images. *Neurocomputing* 332:56–70
6. Kaur J, and Chander Shekhar (2020) Multimodal Medical Image Fusion using deep learning. *Adv Comput Techniques Biomedical Image Anal* : 35–56
7. Huang B, Yang F, Yin M, Mo X, Zhong C (2020) A Review of Multimodal Medical Image Fusion Techniques. *Computational and mathematical methods in medicine* 144:8279342. <https://doi.org/10.1155/2020/8279342>
8. Peter JD, Fernandes SL, Thomaz CE, Viriri S (eds) (2019) *Computer Aided Intervention and Diagnostics in Clinical and Medical images*. Springer International Publishing

9. Fouladi S, Safaei AA, Mammone N et al (2022) Efficient deep neural networks for classification of Alzheimer's disease and mild cognitive impairment from Scalp EEG recordings. *Cogn Comput* 14:1247–1268. <https://doi.org/10.1007/s12559-022-10033-3>
10. Kong Y, Deng Y, Dai Q (2015) Discriminative clustering and feature selection for Brain MRI Segmentation. *IEEE Signal Process Lett* 22(5):573–577
11. Barachini O, Bernt R, Mirzaei S et al (2020) The impact of 18F-FDOPA-PET/MRI image fusion in detecting liver metastasis in patients with neuroendocrine tumors of the gastrointestinal tract. *BMC Med Imaging* 20:22. <https://doi.org/10.1186/s12880-020-00424-z>
12. Parsai A, Miquel ME, Jan H, Kastler A, Szyszko T, Zerizer I (2019) Improving liver lesion characterisation using retrospective fusion of FDG PET/CT and MRI. *Clin Imaging* 55:23–28. <https://doi.org/10.1016/j.clinimag.2019.01.018>
13. Yin M, Liu X, Liu Y, Chen X (2018) Medical image Fusion with parameter-adaptive pulse coupled-neural network in Nonsubsampled Shearlet transform domain. *IEEE Trans Instrum Meas* PP:1–16
14. Jiang J, Feng X, Hu Z, Hu X, Liu F, Huang H (2021) Medical image fusion using transfer learning and L-BFGS optimization algorithm. *Int J Imaging Syst Technol* 31:2003–2013. <https://doi.org/10.1002/ima.22560>
15. Lahoud F (2019) and Sabine Süsstrunk. Zero-Learning Fast Medical Image Fusion. In 2019 22th International Conference on Information Fusion (FUSION), 1–8
16. Zhang Y, Xiang W, Zhang S, Shen J, Wei R, Bai X, Zhang L, Zhang Q (2022) Local extreme map guided multi-modal brain image fusion. *Front Neurosci* 16:1055451. <https://doi.org/10.3389/fnins.2022.1055451> PMID: 36389249; PMCID: PMC9650390
17. Ma J, Xu H, Jiang J, Mei X, Zhang X-P (2020) DDcGAN: a dual-discriminator conditional generative Adversarial Network for Multi-resolution Image Fusion. *IEEE Trans Image Process* 29:4980–4995. <https://doi.org/10.1109/TIP.2020.2977573>
18. Lee MW (2014) Fusion imaging of real-time ultrasonography with CT or MRI for hepatic intervention. *Ultrasonography* 33(4):227–239. <https://doi.org/10.14366/usg.14021> Epub 2014 Jun 10. PMID: 25036756; PMCID: PMC4176112
19. Schwarze V, Rübenthaler J, Marschner C, Fabritius MP, Rueckel J, Fink N, Pühr-Westerheide D, Gresser E, Froelich MF, Schnitzer ML, Große Hokamp N, Afat S, Staehler M, Geyer T, Clevert DA (2020) Advanced Fusion imaging and contrast-enhanced imaging (CT/MRI-CEUS) in Oncology. *Cancers (Basel)* 12(10):2821. <https://doi.org/10.3390/cancers12102821> PMID: 33007933; PMCID: PMC7600560
20. Ma J (2008) Dixon techniques for water and fat imaging. *J Magn Reson Imaging* 28:543–558. <https://doi.org/10.1002/jmri.21492>
21. Omoumi P (2022) The Dixon method in musculoskeletal MRI: from fat-sensitive to fat-specific imaging. *Skeletal Radiol* 51:1365–1369. <https://doi.org/10.1007/s00256-021-03950-1>
22. Kavur, A. E., Gezer, N. S., Barış, M., Aslan, S., Conze, P. H., Groza, V., Pham, D.D., Chatterjee, S., Ernst, P., Özkan, S., Baydar, B., Lachinov, D., Han, S., Pauli, J., Isensee, F., Perkonigg, M., Sathish, R., Rajan, R., Sheet, D., Dovletov, G., ... Selver, M. A. (2021). CHAOS Challenge - combined (CT-MR) healthy abdominal organ segmentation. *Medical image analysis*, 69, 101950. <https://doi.org/10.1016/j.media.2020.101950>
23. Wu D, Yang A, Zhu L, Zhang C (2014) Survey of Multi-sensor Image Fusion. In: Ma S et al (eds) *Life System modeling and Simulation*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 358–367
24. Haque R, Hassan MM, Bairagi AK et al (2024) NeuroNet19: an explainable deep neural network model for the classification of brain tumors using magnetic resonance imaging data. *Sci Rep* 14:1524. <https://doi.org/10.1038/s41598-024-51867-1>
25. Jagalingam P, and Arkal Vittal Hegde (2015) A review of Quality Metrics for Fused Image. *Aquat Procedia* 4:133–142. <https://www.sciencedirect.com/science/article/pii/S2214241X15000206>
26. Singh S, Gupta D, Anand RS, Kumar V (2015) Nonsubsampled Shearlet based CT and MR Medical Image Fusion using biologically inspired spiking neural network. *Biomed Signal Process Control* 18:91–101. <https://www.sciencedirect.com/science/article/pii/S1746809414001943>
27. Kavur AE, Selver MA, Dicle O, Barış M, Gezer NS (2019) CHAOS - Combined (CT-MR) Healthy Abdominal Organ Segmentation Challenge Data (Version v1.03) [Data Set]
28. Lins C, Salmon C, Nogueira-Barbosa M (2021) Applications of the Dixon Technique in the evaluation of the Musculoskeletal System. *Radiol Bras* 54(1):33–42. <https://doi.org/10.1590/0100-3984.2019.0086> - Pubmed
29. Apostolopoulos ID, Papatathanasiou ND, Papandrianos N, Papageorgiou E, Apostolopoulos DJ (2023) Innovative attention-based Explainable Feature-Fusion VGG19 Network for Characterising Myocardial Perfusion Imaging SPECT Polar maps in patients with suspected coronary artery disease. *Appl Sci* 13(15):8839. <https://doi.org/10.3390/app13158839>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.