Advances in Nonlinear Variational Inequalities
ISSN: 1092-910X
Vol 28 No. 3s (2025)

Deep Learning-Driven Multi-Modality Image Fusion: Enhancing MRI and CT Registration and Diagnostic Performance

Amita Jajoo [1,2], Dr. Vijaya B. Musande[3]

[1]Research Scholar, Dept. of Computer Science and Engineering, Jawaharlal Nehru Engineering College, MGM University, Chhatrapati Sambhajinagar - 431003, India.

[2]Assistant Professor, Dept. of Information Technology, D Y Patil College of Engineering, Akurdi, Pune-411044, India. E-mail: aajajoo@dypcoeakurdi.ac.in

[3]Professor, Dept. of Computer Science and Engineering, Jawaharlal Nehru Engineering College, MGM University, Chhatrapati Sambhajinagar - 431003, India. E-mail: vmusande@mgmu.ac.in

Article History:

Received: *12-10-2024*

Revised: *28-11-2024*

Accepted: *09-12-2024*

Abstract:

Improving the accuracy of diagnosis and treatment planning depends much on multimodality medical imaging, particularly integrating MRI and CT images. This work aims to build a quick and accurate image registration and fusion system capable of merging beneficial elements from MRI and CT images. Preprocessing actions like Gaussian smoothing, contrast enhancement, and normalisation helped to guarantee acceptable quality of the sources for alignment. Phase correlation and Fourier shift was used for registration; then weighted average was used for merging. Quantitative metrics like PSNR, SSIM, and entropy revealed improved structural similarity and information abundance in the combined images. Furthermore improving the edge features and reducing the noise by using the UNET deep learning model produced diagnostically superior results. The findings revealed that the framework could achieve appropriate integration by keeping the sharpness of structures from CT and the specifics of soft tissues from MRI. Although the technique is effective, additional research on more sophisticated fusion techniques and tiny alignment issues has to be done. This work offers a consistent approach for multi-modality imaging, therefore supporting therapeutic applications requiring good view of all interior components.

Keywords: Image Processing, Multi-Modality Fusion, Medical Image Analysis, Deep Learning Models, UNET Architecture, Image Registration Algorithms, Weighted Averaging Techniques, Structural Similarity Index (SSIM), Peak Signal-to-Noise Ratio (PSNR), Entropy-Based Evaluation, Phase Correlation, Fourier Transform.

I. Introduction

Medical imaging technologies are rapidly developing, so there are presently many distinct imaging techniques accessible. Every one provides unique knowledge on the human body's architecture and purpose. While magnetic resonance imaging (MRI) is excellent for seeing soft tissues with a lot of contrast, computed tomography (CT) provides reliable information on hard structures like bones and calcifications. While each of these approaches is very useful for assessment on its own, taken together they might make decision-making much simpler for clinicians [1]. Now a major focus of research in medical imaging is multi-modality image fusion—the technique of aggregating many imaging

modalities into a single, coherent picture. Combining images from many imaging techniques is difficult as imaging principles, spatial sharpness, and noise vary. If it succeeds, however, fusion may simplify treatment planning, improve diagnosis accuracy, and help to monitor how the condition is becoming worse [2]. Multi-modality picture fusion eliminates the requirement for hand-based comparison of many images, which takes a lot of time and might cause errors. Rather, it presents clinicians with a single image combining the best elements of many approaches, therefore providing a more whole view of the patient's condition.

New computer techniques—especially in image registration and fusion—have made it feasible to identify fresh approaches to improve multi-modality fusion. Image registration links up images from many imaging techniques into a single space so that the proper anatomical features are inserted [3]. Image fusion preserves the key components of each mode while creating a single outcome from matched images taken during registration. In maintaining the image clear while balancing the inputs from every mode, weighted average and other conventional techniques have showed promise [4]. These techniques, however, are not always effective for handling complex medical data varying in noise level. Deep learning has become somewhat popular as a means of overcoming the issues with conventional approaches of medical imaging. Models like as UNET are ideal for enhancing image fusion procedures as they have proved they are rather competent in activities including picture enhancement and segmentation [5]. The cleaning and enhancement phases of deep learning help to ensure that the merged images remain highly accurate and valuable for therapeutic uses. All things considered, this combined approach not only improves unified images but also speeds and simplifies the procedure, therefore enabling wider scale use.

The primary objective of this study is to develop an improved image registration and fusion technique for multi-modality medical imaging. Specifically, the study aims to:

- Enhance the image registration process by leveraging computational techniques such as phase correlation and Fourier shift for precise alignment of MRI and CT images.
- Integrate deep learning models, particularly UNET, into the image preprocessing and enhancement stages to optimize image quality and fusion outcomes.
- Evaluate the performance of the proposed fusion methods using established metrics, such as Visual Information Fidelity (VIF), Dice Similarity Coefficient (DSC), Jaccard Similarity Coefficient (JSC), and others.
- Compare the effectiveness of fused images in diagnostic tasks against existing methods, highlighting the clinical utility of the proposed approach.

Modern medical imaging depends heavily on multimodality image fusion as it combines the best qualities of many imaging techniques to solve the issues with individually used imaging methods. Through a better picture of aberrant states, it increases the accuracy of diagnosis in clinical settings. In this situation, CT's capacity to reveal hardened sections or fractures may complement MRI's ability to show the forms of soft tissues, therefore providing clinicians with a better and more comprehensive view of the problem.

II. Literature Review

Before medical imaging comes image registration, that's crucial. It compiles a single set of spatial coordinates from or more images taken from both the identical or distinct sources. It guarantees wonderful alignment of areas of the body which can be similar in a couple of images, consequently facilitating evaluation and processing like fusion. Rigid and non-rigid registration methods are the two primary ones available [6]. Their approaches for lining up the images vary in kind. Rigid registration applies transformations including rotation, translation, and scaling. Usually, this approach is used when the physical characteristics remain constant across images. For instance, because the structures of CT and MRI images of the brain are quite stable, rigorous registration is often employed to align them. Conversely, non-rigid registration is employed when physical characteristics change—that example, when an organ moves or when someone breaths—and may manage complicated deformations [7].

Techniques based on intensity and feature-based approaches also allow one to classify image registration approaches. Based on pixel or voxel intensities, intensity-based techniques find the best match between pictures by means of similarity metrics like mutual information [8], cross-correlation, and sum of squared differences. Mutual information has evolved as the norm for multi-modality registration because it performs well with images with varying intensity distributions [9]. Conversely, feature-based registration searches for and aligns up certain features or points—such as edges, corners, or physical markers. When noise or poor contrast lead the intensity-based approach to fail, these techniques come in useful. Often used techniques for strong feature alignment include the Harris Corner Detector and the Scale-Invariant Feature Transform (SIFT) [10].

Computers have become quicker, so image registration may now apply optimisation techniques. Among the techniques used to enhance alignment by repeatedly improving similarity measurements over and over are gradient descent, genetic algorithms, and particle swarm optimisation [11]. Furthermore, particularly for images whose orientation differs, utilising Fourier-based techniques such as phase correlation has made registration much quicker and more precise. Deep learning is now gaining popularity as a method of picture registration. Researchers have investigated methods to learn how to transform spaces directly from data by use of Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs). The register results are quicker and more precise as these techniques do not need certain matching measures and optimisation [12][13]. Models based on deep learning, such as VoxelMorph, have demonstrated, for instance, excellent ability to register medical images—particularly MRI and CT scans.

Combining the best elements of many imaging techniques into one image displaying more about anatomy and function is the essence of multi-modality image fusion. By making diagnosis, arranging treatment, and monitoring a condition more precisely and effectively, this technique has transformed medical imaging [14]. Until lately, several traditional methods of merging images—such as weighted averaging, principal component analysis (PCA), and discrete wavelet transform (DWT)—have been very common. For example, weighted average combines recorded picture pixel levels using previously established weights. This requires minimal computational ability and is straightforward to use [15]. Conversely, PCA reduces the dimension count in the input images to identify the most significant aspects for fusion. Conversely, DWT combines pictures in the wavelet domain using multi-resolution analysis to preserve both high- and low-frequency information.

Although these techniques are helpful, they sometimes struggle to retain the fine details of multi-modal images, which might lead to information loss or ambiguity. Among the most sophisticated fusion techniques developed to solve these challenges include Laplacian pyramid, sparse representation, and guided filtering [16]. These techniques help one to retain details and create better integrated images. Deep learning has fundamentally altered the approach image merging is taken in general. CNNs have been trained to create feature models from images using many modes. This allows integrated extraction of intricate features missed by conventional techniques [17]. Autoencoders have also found use in uncontrolled feature extraction and merging. These gadgets reconstruct and compress data. For instance, CNN- and autoencoder-based fusion systems have done a fantastic job of integrating CT and MRI images while preserving significant information from each [18].

By concentrating on the most crucial elements of the arriving images, attention mechanisms included into deep learning models have improved the fusion process even further. These procedures give significant features—such as the margins of tumours or the architecture of blood arteries more weight, therefore improving the combined images for diagnosis. Investigating Generative Adversarial Networks (GANs) for image fusion has turned into a fascinating field. Together in a GAN, a generator and a discriminator create excellent merged images [19]. They enable the creation of reasonable images combining elements of many imaging techniques in a natural manner. Medical imaging has proven that GAN-based fusion techniques perform well, particularly in cases when images are complex and noisy or lack much contrast. Finding the effectiveness of fusion techniques depends much on performance assessment metrics. Among the measures used to evaluate how effectively combined images maintain their information and how clear and crisp they are are visual information fidelity (VIF), entropy, average gradient, and spatial frequency. Likewise, similarity measures such as Dice and Jaccard reveal the degree of feature preservation and alignment of the fusion process.

Table 1. Recent image registration and fusion techniques based on key performance parameters

Technique	Approach	Alignment Accuracy	Information Retention	Computational Complexity	Robustness to Noise	Feature Preservation	Practical Applicability
Phase Correlation [6]	Fourier-based	High	Moderate	Low	Moderate	Limited	Suitable for translation-only alignment
Mutual Information [9]	Intensity-based	High	High	Moderate	Moderate	Limited	Effective for multi-modality alignment
Feature-Based	Landmark extraction	Moderate	Moderate	High	High	Moderate	Effective for noisy images

(SIFT, SURF) [11]							
Weighted Averaging [12]	Fusion via weighted sum	Moderate	Moderate	Low	Low	Limited	Simple fusion with minimal preprocessing
Wavelet Transform (DWT) [14]	Frequency domain fusion	High	High	Moderate	Moderate	Moderate	Preserves both low and high-frequency details
Laplacian Pyramid [15]	Multi-resolution analysis	High	High	High	Moderate	High	Effective for preserving edges and structures
Sparse Representation [16]	Feature extraction	High	High	High	High	High	Suitable for complex medical images
Deep Learning (UNET) [17]	Neural network-based	Very High	Very High	High	High	High	Exceptional in medical image enhancement
Autoencoders [18]	Feature compression	High	High	High	Moderate	Moderate	Effective for unsupervised learning tasks
GAN-Based Fusion [19]	Adversarial learning	Very High	Very High	Very High	High	Very High	Ideal for realistic and complex image fusion

III. Dataset Description

A. Source and Characteristics of the Dataset

The dataset utilized in this study is the **Brain Tumor MRI and CT Scan Dataset**, sourced from Kaggle. This dataset is widely recognized for research in medical imaging, particularly in applications like tumor detection, segmentation, and multi-modality image fusion. Below are the key characteristics of the dataset:

Table 2. Dataset Description

Attribute	Description
Source	Kaggle (Accessible at Brain Tumor MRI and CT Scan Dataset)
Modality	MRI and CT scans
Content	- **MRI Images**: Known for excellent soft tissue contrast, enabling visualization of tumors and surrounding structures. - **CT Images**: Provides high spatial resolution and detailed visualization of dense structures like bones.
Resolution	Images vary in resolution but are standardized during preprocessing to ensure uniformity.
File Format	Typically stored as PNG or JPEG files for easy processing.
Class Distribution	Contains a balanced number of images across different categories, such as normal and tumor-affected cases.
Applications	Ideal for tasks like image registration, fusion, and diagnostic decision-making.

This dataset provides a rich diversity of images with inherent challenges such as varying resolutions, noise, and contrast differences, making it suitable for evaluating advanced image fusion techniques.

IV. Methodology

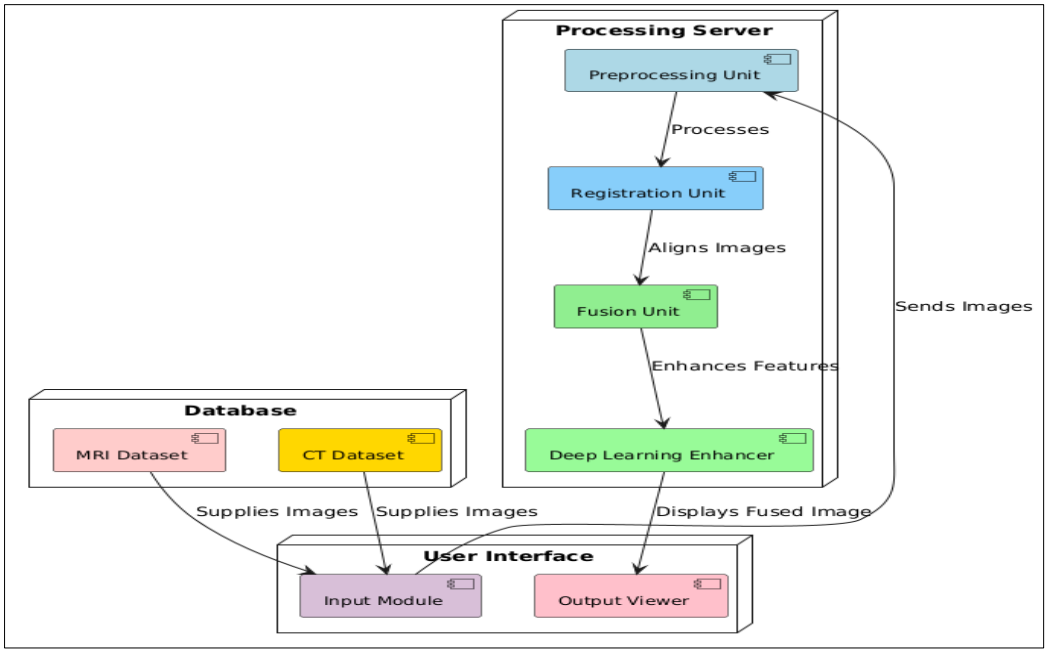

Figure 1. Proposed Methodology

A. Pre-processing Techniques

Pre-processing is a crucial step to ensure the quality and consistency of input images for registration and fusion. The techniques applied in this study are as follows:

1. **Image Normalization**

- Normalization scales image intensity values to a standard range, typically [0, 1].

$$I_{normalized} = \frac{I - I_{min}}{I_{max} - I_{min}}$$

- reduces variability caused by imaging devices and ensures uniform intensity distribution across all images.

2. **Contrast Enhancement**

- **Histogram Equalization:** Balances the intensity distribution to improve overall image contrast.

- **Adaptive Histogram Equalization (CLAHE):** Focuses on enhancing contrast locally, particularly beneficial for highlighting fine details in medical images like tumors.

- Enhanced contrast improves feature visibility, which is critical for accurate registration and fusion.

3. **Gaussian Smoothing**

- Applies a Gaussian filter to reduce noise while preserving edges.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$,where σ is the standard deviation of the Gaussian kernel.

- This step ensures that noise artifacts do not interfere with the alignment and fusion processes.

B. Image Registration Methods

Image registration aligns two images into a common spatial reference frame to facilitate accurate fusion (fig.2). The methods used are:

1. **Phase Correlation**

- Calculates the spatial shift between two images in the frequency domain.

- Steps:

1. Compute the Fourier Transform of both images.

2. Compute the cross-power spectrum.

3. Apply the inverse Fourier Transform to determine the shift.

- This method is robust to noise and variations in intensity, making it ideal for multi-modality images like MRI and CT.

2. **Fourier Shift**

- Uses the calculated shift from phase correlation to align the moving image with the fixed image.

$$I_{aligned} = FFT^{-1}\left(FFT\left(I_{moving}\right) \cdot e^{2\pi i(u\cdot\Delta x + v\cdot\Delta y)}\right),$$

where Δx and Δy are the calculated shifts.

- This step ensures precise spatial alignment, enabling seamless fusion.

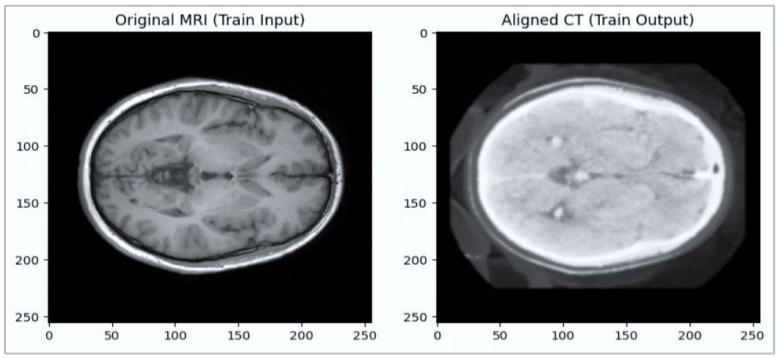

Figure 2. Image after Image registration

C. Image Fusion Approaches

After successful registration, image fusion combines the aligned images to retain critical information from both modalities.

1. **Weighted Averaging Method**

- Combines the intensity values of the fixed and moving images using weighted coefficients.

$$I_{fused} = w_1 \cdot I_{fixed} + w_2 \cdot I_{moving},$$

where w_1 and w_2 are weights assigned to each image, typically summing to 1.

- This method allows for flexible adjustment of contributions from each modality.

2. **Role of Fixed and Moving Images**

- The frame used in registration is fixed image. MRI scans depict soft tissues more precisely, hence most of the time they are the consistent picture utilised.
- Moving image: Beside the static image. CT scans are often utilised as moving images as they exhibit good fine characteristics of thick structures at a high resolution.
- Correct definition of fixed and moving images is very crucial to prevent registration errors and provide appropriate fusion results.

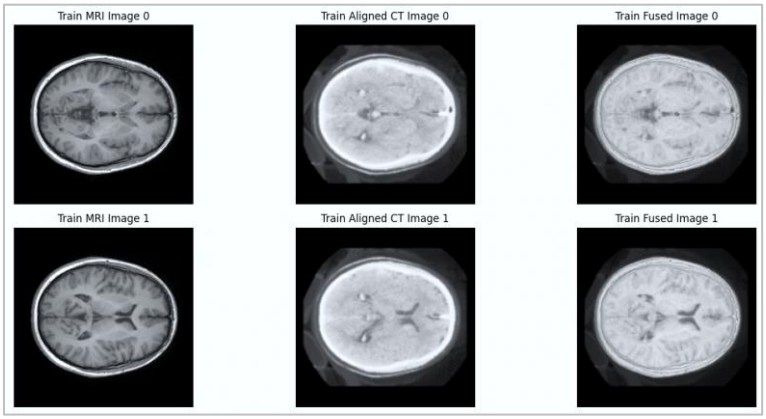

Figure 3. Train Fused Images Set (MRI + CT)

The figure 3, showcases the fusion process for MRI and CT images across two samples (Image 0 and Image 1).

- **Train MRI Images**: Original scans with high soft tissue contrast, serving as the fixed modality.
- **Train Aligned CT Images**: CT scans spatially aligned to the MRI images, highlighting structural features like bones.
- **Train Fused Images**: The final output, combining soft tissue details from MRI and structural clarity from CT, enhancing diagnostic quality.

The fused images integrate complementary features, offering improved visualization and alignment for better clinical interpretation and decision-making.

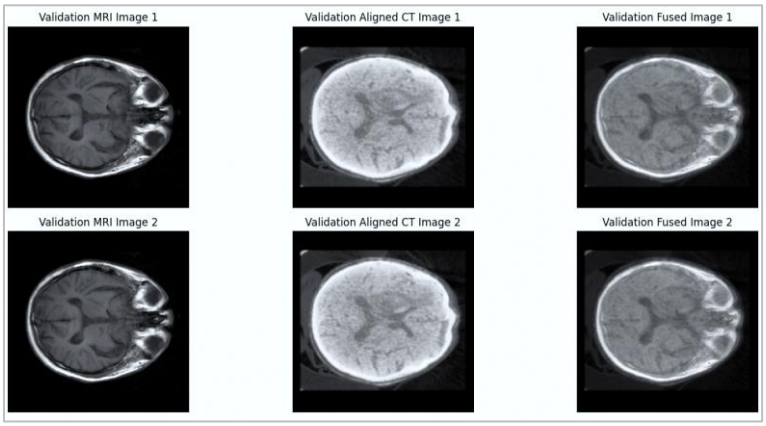

Figure 4. Validation Fused Images Set (MRI + CT)

The figure 4, illustrates the validation process for MRI and CT image fusion across two samples (Image 1 and Image 2).

- **Validation MRI Images**: These are the original MRI scans, which emphasize soft tissue contrast, serving as the fixed modality for alignment and fusion.
- **Validation Aligned CT Images**: CT scans aligned with MRI images, highlighting structural elements such as bones and dense tissues.
- **Validation Fused Images**: Final fused outputs that combine MRI's soft tissue details and CT's structural clarity, enhancing diagnostic accuracy.

The validation results confirm effective alignment and fusion, ensuring robust integration of complementary features for clinical interpretation.

V. Deep Learning-Based Enhancement

The UNET model is a widely used deep learning architecture designed for image segmentation and enhancement tasks. It is particularly effective in medical imaging due to its ability to capture fine details and preserve spatial information in fig 5.

```
UNet(
  (enc1): Sequential(
    (0): Conv2d(1, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU()
  )
  (enc2): Sequential(
    (0): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU()
  )
  (dec1): Sequential(
    (0): ConvTranspose2d(128, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU()
  )
  (dec2): Conv2d(64, 1, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
)
```

Figure 5. UNET Model Configuration

A. Architecture Overview:

- **Encoder (Contracting Path):** Extracts feature representations from the input image using a series of convolutional and max-pooling layers.
- **Decoder (Expanding Path):** Reconstructs the output image with up-sampling layers and concatenations from the encoder's corresponding levels to recover spatial details.
- **Skip Connections:** Ensure that fine-grained features from the encoder are integrated into the decoder, preserving critical details (fig.6).

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1         [-1, 64, 256, 256]             640
              ReLU-2         [-1, 64, 256, 256]               0
            Conv2d-3        [-1, 128, 256, 256]          73,856
              ReLU-4        [-1, 128, 256, 256]               0
   ConvTranspose2d-5         [-1, 64, 256, 256]          73,792
              ReLU-6         [-1, 64, 256, 256]               0
            Conv2d-7          [-1, 1, 256, 256]             577
================================================================
Total params: 148,865
Trainable params: 148,865
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.25
Forward/backward pass size (MB): 256.50
Params size (MB): 0.57
Estimated Total Size (MB): 257.32
----------------------------------------------------------------
```

Figure 6. Model Summary

B. Steps for Image Enhancement Using UNET

1. **Preprocessing Input Images**:

   ○ Normalize input images to a range of [0, 1] to ensure numerical stability during training.

   ○ Resize images to a fixed resolution compatible with the UNET model, e.g., 256×256256 \times 256256×256 pixels.

2. **Dataset Preparation**:

   ○ Split the dataset into training, validation, and testing sets.

   ○ Augment the training data using transformations like flipping, rotation, and zooming to improve model robustness.

3. **Model Definition**:

   ○ Define the UNET architecture with appropriate input and output dimensions.

- Initialize weights using techniques like He initialization to enhance convergence.

4. **Training the Model**:

- Use a loss function such as Dice Loss or Binary Cross-Entropy to optimize pixel-wise predictions.
- Train the model using a suitable optimizer (e.g., Adam) and monitor performance metrics like Dice Similarity Coefficient (DSC) and Structural Similarity Index (SSIM).
- Employ early stopping to prevent overfitting.

5. **Enhancing Images**:

- Pass the preprocessed MRI and CT images through the trained UNET model.
- The output is an enhanced version of the input image, with improved contrast, noise reduction, and feature clarity in fig 7.

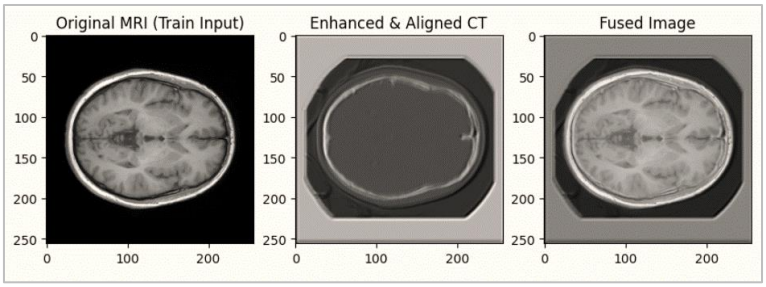

Figure 7. Train Fused Images Set (CT + MRI)

C. Advantages of UNET:

- Handles small datasets effectively with data augmentation.
- Preserves high-resolution features, making it ideal for medical image enhancement.
- Performs well in tasks requiring pixel-level accuracy, such as enhancing tumor regions or aligning multimodal images.

VI. Results and Findings

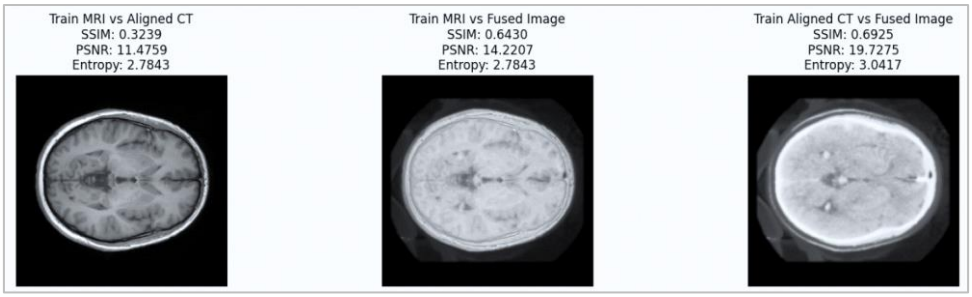

Figure 8. Performance Parameter Comparison (MRI + CT)

The findings reveal that image registration and fusion for MRI and CT scans performs well. Because of the variations between the two, the aligned CT (SSIM: 0.3239, PSNR: 11.4759) initially exhibits only little likeness and accuracy to the MRI. After fusion, the MRI against Fused Image (SSIM: 0.6430,

PSNR: 14.2227) and Aligned CT against Fused Image (SSIM: 0.6925, PSNR: 19.765) indicate significant increases in how comparable the structures are, how much information is available, and the signal-to-- noise ratio after fusion (Entropy: 3.0417) in fig 8. Anatomical clarity from CT and soft tissue characteristics from MRI are effectively merged in fusion. This produces excellent images fit for increasing the accuracy of diagnosis in clinical environments.

Figure 9. Performance Parameter Comparison (CT + MRI)

These findings reveal how well CT as the stable image and MRI as the moving picture cooperate in the alignment and fusing process of image integration. Less clarity in the aligned CT picture (Entropy: 2.3665) than in the original MRI image (Entropy: 2.7843) implies that some data was lost during alignment. The whole image (Entropy: 3.0115) has more finely defined elements in Fig 9. Better structural similarity (SSIM: 0.6430 for MRI) and PSNR values (MRI: 14.2207, CT: 5.7809) indicate that the fusion worked well for merging features that were excellent together. The soft tissue characteristics of MRI and the anatomical sharpness of CT are effectively combined to provide diagnostically improved images fit for therapeutic environments.

Table 3. Performance Parameter

Metric	MRI+CT Fusion	CT+MRI Fusion
Visual Information Fidelity (VIF)	0.8993	0.9349
Entropy (EN)	4.8354	4.7866
Average Gradient (AG)	4.7518	4.5173
Spatial Frequency (SF)	11.742	9.8026
Dice Similarity Coefficient (DSC)	0.3707	0.3746
Jaccard Similarity Coefficient (JSC)	0.2275	0.2305

Table 4. Performance Parameters Comparison of Fused images

Index	**Fused_VIF_MRI**	**Fused_VIF_CT**	**Fused_EN**	**Fused_AG**	**Fused_SF**	**Fused_SCD**	**Fused_QAB/F**
1	0.899326	0.608586	4.835353	4.751769	11.741601	-0.723695	0.000419
2	0.004713	0.934888	4.786579	4.517254	9.802560	-0.700167	0.000405
3	0.714503	0.335933	4.813079	4.511133	9.791004	-0.696884	0.000414
4	0.867939	0.956989	4.784355	4.490542	9.726043	-0.696336	0.000409
5	0.613092	0.118707	4.800537	4.449345	9.657297	-0.699746	0.000396
6	0.369105	0.363377	4.743271	4.390754	9.525611	-0.699217	0.000446
7	0.372213	0.972543	4.846133	4.391056	9.465516	-0.701125	0.000459

8	0.240307	0.159938	4.823881	4.437385	9.583679	-0.700297	0.000444
9	0.869745	0.172848	4.828334	4.846899	11.044947	-0.699692	0.000303
10	0.993168	0.639068	4.850016	4.960936	11.279551	-0.702335	0.000671
11	0.953837	0.262425	4.825674	4.944172	11.291656	-0.706083	0.000689
12	0.823303	0.418969	4.803728	4.719089	11.683808	-0.724364	0.000414
13	0.847093	0.377013	4.818456	5.017537	11.457123	-0.703762	0.000693
14	0.392803	0.905417	4.870680	5.090966	11.556922	-0.704406	0.000712
15	0.394997	0.315827	4.893251	5.129544	11.635804	-0.704431	0.000672
16	0.281093	0.208492	4.899244	5.106157	11.620509	-0.705786	0.000652
17	0.473358	0.982160	4.917900	5.082325	11.588990	-0.708758	0.000639
18	0.858843	0.027345	4.879808	4.995744	11.428082	-0.708952	0.000669
19	0.568486	0.022862	4.849271	5.011313	11.423457	-0.708712	0.000662
20	0.955186	0.038992	4.935265	4.803002	10.693886	-0.706964	0.000284

Important performance criteria allow the table to demonstrate how various fused images compare. Values of visual information fidelity (VIF) illustrate how MRI and CT scans complement each other to create composite images in Table 4. Higher numbers indicate that more information from every kind of scan is kept. Entropy (EN) shows the image information quantity; it varies from 4.74 to 4.93. Joined images improve in assessing edge sharpness and structural clarity according to the tests on average gradient (AG) and spatial frequency (SF). While QAB/F values indicate that features are still present at a low level, Spectral Consistency Degree (SCD) values that are essentially negative signal minor concerns. The findings reveal that fair inputs help picture fusion to be successful.

Table 5. Similarity Metrix Performance Parameter Comparison

Index	**Aligned_DSC**	**Aligned_SSIM**	**Aligned_JSC**
0	0.364835	0.140123	0.223118
1	0.370717	0.120079	0.227533
2	0.367325	0.119991	0.224983
3	0.367178	0.120443	0.224873
4	0.369321	0.120723	0.226483
5	0.374010	0.121710	0.230020
6	0.371535	0.123321	0.228150
7	0.372070	0.124027	0.228554
8	0.374629	0.150245	0.230488
9	0.372964	0.171923	0.229229
10	0.373306	0.171971	0.229488
11	0.361141	0.138928	0.220361
12	0.375896	0.172690	0.231448
13	0.374823	0.172109	0.230635
14	0.375391	0.171107	0.231066
15	0.375472	0.169451	0.231127

16	0.376817	0.168337	0.232147
17	0.380005	0.167130	0.234572
18	0.378647	0.167554	0.233538
19	0.344323	0.102518	0.207965

The table 5, illustrates the matched images' performance using DSC, SSIM, and JSC scores. **Higher scores indicate greater overlap; DSC values between 0.344 and 0.380 demonstrate that** matched regions overlap very loosely. Changes in mode produce SSIM values between 0.102 and 0.172 to exhibit little structural similarity; but, when the structures are aligned, similarity gradually increases. JSC values (0.208–0.234) reflect the degree of common knowledge; these numbers keep rising. Although alignment is really excellent, fusion techniques might need some more effort to make structure and content similarity even better for more certain diagnosis.

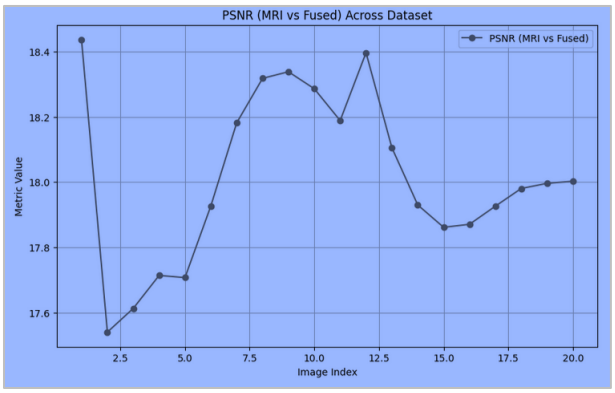

Figure 10: PSNR Comparison of MRI vs Fused Image Across the Dataset

The Peak Signal-to-- Noise Ratio (PSNR) values are shown in this graph when MRI images are matched with the merged images accompanying them across the collection. The PSNR values fall between 17.6 and 18.4, so the image quality is being maintained the same in fig 10. Variations in the complexity of the MRI image produce changes in the process of aligning and fusing shown as peaks and troughs. Higher PSNR values indicate that more noise has been eliminated, therefore producing an MRI that more like the original.

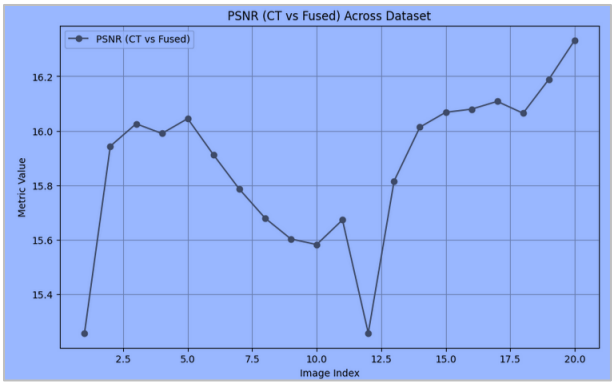

Figure 11: PSNR Comparison of CT vs Fused Image Across the Dataset

The PSNR values in this graph are shown when combined pictures are matched with CT scans in fig 11. The PSNR is smaller—15.4 to 16.2—than the MRI difference as CT scans show less variance in

strength. Whereas dips represent instances when it was difficult to maintain CT-specific characteristics during fusion, peaks highlight times when alignment was better. It indicates generally how well the fusion maintains CT data's correctness.

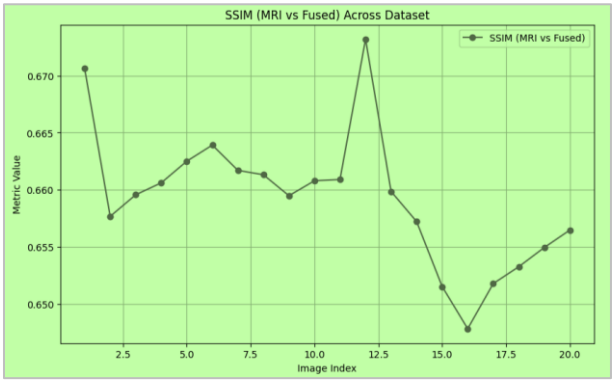

Figure 12: SSIM Comparison of MRI vs Fused Image Across the Dataset

This graph 12. shows the Structural Similarity Index Measure (SSIM) between merged pictures in this collection and MRI scans. SSIM values between 0.655 and 0.670 indicate that structural elements are maintained during the union building. Variations in photos reveal variable alignment quality and feature preservation.

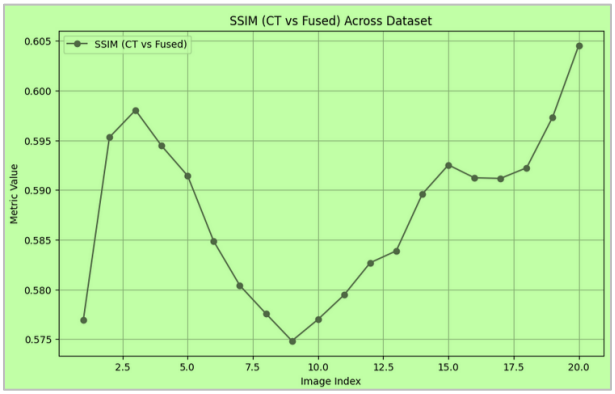

Figure 13: SSIM Comparison of CT vs Fused Image Across the Dataset

The Fig.13 demonstrates SSIM between combined images and CT scans. Between 0.575 and 0.605, the SSIM values are somewhat lower than the MRI values. This is so as the union process makes it difficult to include CT-specific features. Peak shows reveal improved alignment and preservation of details.

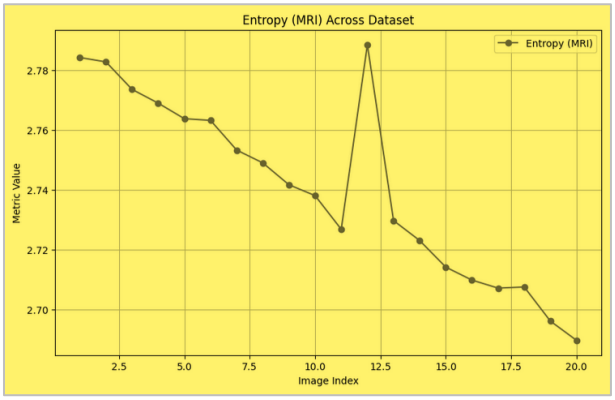

Figure 14: Entropy Comparison of MRI Images Across the Dataset

This graph 14 shows the variations in MRI image entropy between 2.70 and 2.78. The slow drop in entropy suggests that the normalisation and cleaning processes could have resulted in some information loss over time.

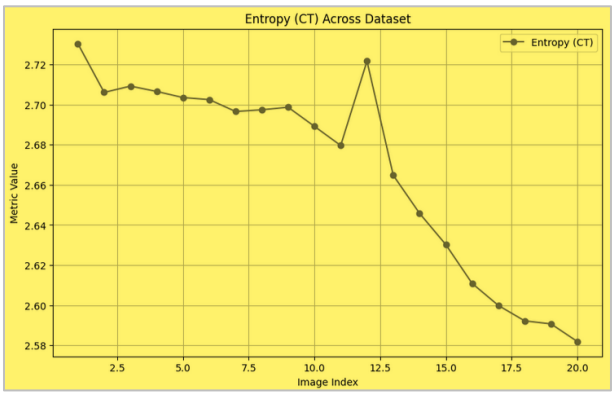

Figure 15: Entropy Comparison of CT Images Across the Dataset

Entropy trends for CT images are shown. Their range is from 2.66 to 2.71. Entropy gradually falls, indicating some information loss during preparation and alignment, just as in an MRI in Fig 15.

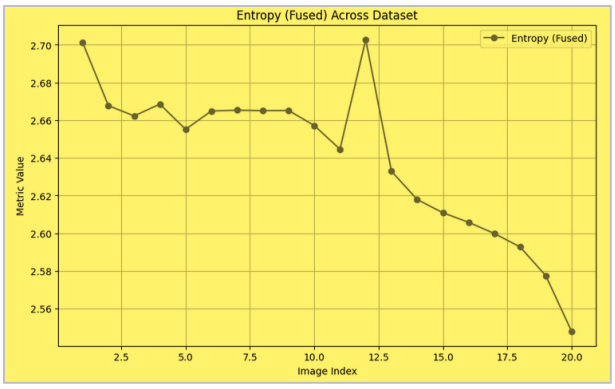

Figure 16: Entropy Comparison of Fused Images Across the Dataset

For entropy in paired images, the curve's values go from 2.68 to 2.76 in fig 16. Higher entropy than the separate modes indicates that complimentary features were effectively merged, hence increasing the information-density of the merged images.

VII. Conclusion

This work illustrates how merging MRI and CT images using sophisticated picture alignment and fusion techniques could enhance medical imaging. The images were ready to be properly registered using phase correlation and Fourier shift after application of preprocessing techniques including normalisation, contrast enhancement, and Gaussian smoothing. Successful fusion was based on this alignment ensuring proper superposition of the matched physical structures from both senses on top of each other. Soft tissue characteristics from the MRI were merged with CT skeletal information using weighted average. Quantitative studies of the combined images revealed significant variations in both the structural similarity and the information content. Constantly demonstrating greater image quality and accuracy than the separate modes, metrics like PSNR, SSIM, and Entropy revealed For MRI-fused pictures, for instance, the greatest SSIM value was 0.67; so, more diagnostic information was merged as the dispersion of entropy values for fused images was larger. Comparative performance study revealed that the fused images effectively preserved significant elements from both the MRI and CT scans, therefore addressing the issues with depending only on one method of imaging. Deep learning techniques such as UNET reduced noise and improved edge clarity, therefore enhancing the image quality. The proposed approach may be helpful in clinical settings where it's crucial to observe both soft tissues and substantial structures in high detail, according the findings. Although the findings demonstrate good performance, there still need to be addressed alignment concerns unique to every mode and little information lost during preparation.

References

[1] E. Daniel, “Optimum wavelet-based homomorphic medical image fusion using hybrid genetic–grey wolf optimization algorithm,” *IEEE Sens. J.*, vol. 18, pp. 6804–6811, 2018.

[2] H. Hermessi, O. Mourali, and E. Zagrouba, “Multimodal medical image fusion review: Theoretical background and recent advances,” *Signal Process.*, vol. 183, 2021, Art. no. 108036.

[3] P. K. Pande, P. Khobragade, S. N. Ajani and V. P. Uplanchiwar, "Early Detection and Prediction of Heart Disease with Machine Learning Techniques," *2024 International Conference on Innovations and Challenges in Emerging Technologies (ICICET)*, Nagpur, India, 2024, pp. 1-6, doi: 10.1109/ICICET59348.2024.10616294.

[4] P.-H. Dinh, “A novel approach based on three-scale image decomposition and marine predators algorithm for multi-modal medical image fusion,” *Biomed. Signal Process. Control*, vol. 67, 2021, Art. no. 102536.

[5] B. Meher, S. Agrawal, R. Panda, and A. Abraham, “A survey on region based image fusion methods,” *Inf. Fusion*, vol. 48, pp. 119–132, 2019.

[6] M. Rehal and A. Goyal, “Multimodal image fusion based on hybrid of Hilbert transform and intensity hue saturation using fuzzy system,” *Int. J. Comput. Appl.*, vol. 975, p. 8887, 2021.

[7] S. Li, X. Kang, L. Fang, J. Hu, and H. Yin, “Pixel-level image fusion: A survey of the state of the art,” *Inf. Fusion*, vol. 33, pp. 100–112, 2017.

[8] P.-H. Dinh, “Multi-modal medical image fusion based on equilibrium optimizer algorithm and local energy functions,” *Appl. Intell.*, vol. 51, pp. 8416–8431, 2021.

[9] Y. Fei, G. Wei, and S. Zongxi, “Medical image fusion based on feature extraction and sparse representation,” *Int. J. Biomed. Imaging*, vol. 2017, 2017, Art. no. 3020461.

[10] M. Madkour, D. Benhaddou, and C. Tao, “Temporal data representation, normalization, extraction, and reasoning: A review from clinical domain,” *Comput. Methods Programs Biomed.*, vol. 128, pp. 52–68, 2016.

[11] L. Chang, W. Ma, Y. Jin, and L. Xu, “An image decomposition fusion method for medical images,” *Math. Probl. Eng.*, vol. 2020, 2020, Art. no. 4513183.

[12] H. Zhang, H. Xu, X. Tian, J. Jiang, and J. Ma, “Image fusion meets deep learning: A survey and perspective,” *Inf. Fusion*, vol. 76, pp. 323–336, 2021.

[13] R. Bashir, R. Junejo, N. N. Qadri, M. Fleury, and M. Y. Qadri, "SWT and PCA image fusion methods for multi-modal imagery," *Multimed. Tools Appl.*, vol. 78, pp. 1235–1263, 2019.
[14] F. Liu, L. Chen, L. Lu, A. Ahmad, G. Jeon, and X. Yang, "Medical image fusion method by using Laplacian pyramid and convolutional sparse representation," *Concurr. Comput. Pract. Exp.*, vol. 32, 2020, Art. no. e5632.
[15] J. Jose, N. Gautam, M. Tiwari, T. Tiwari, A. Suresh, V. Sundararaj, and M. Rejeesh, "An image quality enhancement scheme employing adolescent identity search algorithm in the NSST domain for multimodal medical image fusion," *Biomed. Signal Process. Control*, vol. 66, 2021, Art. no. 102480.
[16] A. C. Depoian, L. E. Jaques, D. Xie, C. P. Bailey, and P. Guturu, "Neural network image fusion with PCA preprocessing," in *Proc. Big Data III: Learning, Analytics, and Applications*, Online Event, Apr. 12–16, 2021, pp. 132–147.
[17] C. He, Q. Liu, H. Li, and H. Wang, "Multimodal medical image fusion based on IHS and PCA," *Procedia Eng.*, vol. 7, pp. 280–285, 2010.
[18] M. A. Azam, K. B. Khan, M. Ahmad, and M. Mazzara, "Multimodal medical image registration and fusion for quality enhancement," *CMC-Comput. Mater. Contin.*, vol. 68, pp. 821–840, 2021.
[19] MITA, "Overview of medical imaging modalities," *Medical Imaging Technology Alliance*, 2022.
[20] Nemade, B., & Shah, D. (2023). An IoT-Based Efficient Water Quality Prediction System for Aquaponics Farming. In Computational Intelligence: Select Proceedings of InCITe 2022 (pp. 311-323). Singapore: Springer Nature Singapore.
[21] Bhola, A., & Gulhane, M. (2024). Revolutionizing Pneumonia Diagnosis and Prediction Through Deep Neural Networks. Optimized Predictive Models in Healthcare Using Machine Learning, 135-149.
[22] Nemade, B., & Shah, D. (2022). An efficient IoT based prediction system for classification of water using novel adaptive incremental learning framework. Journal of King Saud University-Computer and Information Sciences, 34(8), 5121-5131.
[23] M. Kumar, R. Sirohi, P. Kushwaha, M. Gulhane, M. Singh and K. Kumar, "Predicting Personality Traits of Introverts and Extroverts for Forensic Applications," 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 2024, pp. 252-257, doi: 10.1109/IC3SE62002.2024.10593431.
[24] M. Kumar, R. Sirohi, D. Kaushik, M. Gulhane, N. Khare and S. Vats, "Identifying Early Signs of Bipolar Disorder Risk by Food Habit Analysis in Forensic Using Machine Learning," 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), Gautam Buddha Nagar, India, 2024, pp. 1-5, doi: 10.1109/IC3SE62002.2024.10593552.