MULTI-MODAL MEDICAL IMAGE FUSION BASED ON MULTI-BASED BINARY RESIDUAL FEATURE FUSION

KANDASAMY K[1]*, LATHA SHANMUGA VADIVU S[2]
[1,2]Department of Electronics and Communication Engineering,
[1]Anna university Regional Campus Coimbatore, Tamilnadu, India.
[2]Tamilnadu college of Engineering, Coimbatore, Tamilnadu, India.
e-mail: kkandasamy.03@rediffmail.com; latha@tnce.in

Abstract.

The term "medical image fusion" refers to the procedure of aligning and blending separate pictures captured using different imaging technologies. By enhancing imaging quality and decreasing unpredictability and redundancy, it aims to boost the clinical usefulness of medical pictures for the diagnosis and evaluation of medical conditions. Accuracy of clinical choices based on medical pictures has been enhanced via the use of multi model medical image fusion algorithms and tools. In this study, we show that combining pictures from chest X-rays and ultrasounds improve visibility. On the other hand, the merged picture is essential for a high-quality display. Improving the quality of the combined photos requires a classification technique. The accurate fusion and integration of medical images from multiple modalities is a critical step in the diagnostic process that leads to clinical activities and the appropriate treatment of patients, and this chapter suggests a Multi-based Binary Residual Feature Fusion (MBRFF) method for this purpose. The images are preprocessed using the Histogram equalization algorithm. For the feature extraction of medical images, the Multi-Threshold Contour Method is used. The shift variant decomposition transformation technique Discrete Wavelet Transform DWT is effective and it is used to decompose the medical image. The fused medical images are optimized by using Particle swarm optimization PSO. Finally, the simulation results reveal that the suggested mechanism outperforms the other techniques based on conventional approaches. Using the Origin Pro application, the study's outcomes are represented graphically

Keywords: *Image fusion, Multi-based Binary Residual Feature Fusion (MBRFF), Medical images, Particle swarm optimization (PSO), Discrete Wavelet Transform (DWT), origin Pro*
DOI Number:10.14704/nq.2022.20.8.NQ44507 **NeuroQuantology 2022; 20(8): 4805-4820**

1. Introduction

The scientific community has paid increasing attention to multimodal medical imaging in recent years because of its significance in medical diagnosis, computer vision, and the internet of things. One of the primary concerns in this area of study is the most effective and efficient method of combining data from several types of medical imaging sensors, such as single-photon emission tomography and magnetic resonance imaging. This image fusion approach integrates a number of methods

and areas of study, such as image processing, computer vision, and pattern recognition, to aid in the development of more accurate medical diagnoses and better medical decision-making (Dharejo et. al. (2022)). The fusion of many types of medical images is a major area of study in the field of medical image processing, and particularly in the field of image computing. Different medical imaging methods frequently offer complementing information. For more accurate analysis in several diagnostic instances, complementary data must be included. A single composite picture derived from the fusion of many multimodal medical images may be relied upon for more accurate analysis and diagnosis (Liu et al. (2022)). CXR, or chest X-ray imaging, is a non-invasive radiological test that is often used. Because radiologists can readily interpret CXR and it takes less time with fewer evaluation mistakes than other methods, it is preferred in the recent COVID-19 pandemic. As required, CXR pictures can be combined throughout the feature extraction and classification stages. Furthermore, deep convolution neural networks (DCNN) may be used with CXR as a large-scale input picture source to improve the performance of thoracic disorders of various sizes (Hanley et al. (2021)). Most images just include one or a few pieces of information. For instance, the focus angles differ while objects closer to or farther away seem indistinct in one view. If they can be rigorously separated, multiple multi-modal image types are categorized as different image types. Different image kinds reflect various information types. The outcome is an enormous tangle of information that clouds the doctor's decision. Recent studies in the field of medical image processing have recommended image fusion as an efficient method for automatically detecting information present in several pictures and combining them into a single composite image in which all aspects of interest are readily apparent (Zhao et al. (2022)). The standard workflow for Image Fusion is shown in Figure 1.

Figure 1: General procedures for Image Fusion

Extra image data has been visualized using a wide variety of medical image processing methods in recent years. The fusion of medical images is a very efficient method for merging crucial data from several images and combining many image formats into one composite image and improving diagnostic precision. By fusing medical picture technologies, as much of the original image's diagnostic information as feasible may be preserved. The original picture quality for image fusion is generally generated utilizing different energy functions using image fusion approaches based on different fusion worries. Unfortunately, even when it is possible to preserve the properties of each source picture, image fusion is often undermined because it introduces fake texture by breaking the continuity of the high and low contrast values in the domain. Medical image fusion research has been a popular area of study since the minimal information offered by one mode of medical images does not satisfy the demand for clinical analysis, which necessitates a substantial quantity of data (Wang et al. (2022)). We investigated multimodal fusion since the information provided by single-modal fusion was inadequate. To accurately fuse and integrate medical images from many modalities an

essential stage in the diagnostic process that leads to clinical actions and proper patient treatment this chapter proposes the Multi-based Binary Residual Feature Fusion (MBRFF) method. Better contrast, fusion quality, and perceived experience would be achieved by using the intended method.

This article's sections: Section II provides a literature review. Shown in section III are proposed techniques. Section IV contains the results and comments. Section V includes the proposed work's conclusion.

2. Problem Statement

For optimal treatment of patients, “the medical use of medical imaging” for the assessment and treatment of medical conditions needs to be enhanced. With the development of numerous acquisition tools and imaging techniques, it is now feasible to take images of any situation that contain a variety of details. For instance, there are several imaging techniques used in the medical industry, such as magnetic resonance and computed tomography. A method that can merge all the supplementary data from many source images into one composite image is required for quick processing. We suggest a Multi-based Binary Residual Feature technique to eliminate these drawbacks. By keeping the information from the original images and removing any extraneous data and artifacts, our method would produce the fused image.

3. LITERATURE REVIEW

In this part of the article, we will discuss some of the current works of literature. Wang et al. (2022) suggested a multi-modal color medical image fusion technique based on “geometric algebra discrete cosine transform is proposed (GA-DCT)”. The GA-DCT method incorporates the feature of GA, which reflects the multi-vector signal as a whole, to enhance the fusion image and prevent complicated encoding and decoding. First, the source pictures are separated into blocks and stated in GA multi-vector form. Second, they extend the classic DCT to GA space and propose GA-DCT. Third, they utilise GA-DCT to decompose the image to produce AC and DC coefficients. Finally, a fusion technique is used to fuse the images. Zhu et al. (2022) offered a Cross Encoder-based method for fusing many types of medical images together. In CEFusion, the fusion blocks from two different branches are sent into the cross encoder to be used for feature extraction and recycling. To train our network, which can preserve the source pictures' pixel and brightness information to the greatest degree possible, they also offered a novel estimate approach for fidelity term weighting. A weighted factor is incorporated into the loss function, and as a consequence, texture features in certain dark regions of the source picture are not adequately maintained with this approach. Palanisami et al. (2022) described that to blend multi-modality images into a single image for greater information and visual quality without artefacts and ambiguity. The raw images are divided into base and detail layers using the Gaussian filter. Detail layers are blended utilising spatial frequency to retain image clarity and edge detail. Because the foundation layer includes approximation information of a low-contrast source image, it becomes “Sugeno's intuitive fuzzy image (SIFI)”. Texture information from SIFI is used to fuse base layers. The fused output picture with improved contrast and visual effects is recreated by combining the fused base and fused detail layers. He (2022) extracts features from multi-modal medical pictures by building a neural network and analysing multi-modal feature data. The study may be used to other image fusion algorithms. The research team employs the current image fusion algorithm, multi-scale and attention mechanisms to increase the expression ability of extracted features, adaptive fusion of multi-modal medical image data, and multi-modal medicine. Image fusion is possible. Sabbava et al. (2022) describe that fuses brain CT and MRI images and classifies them using YOLO-V2. If the network is abnormal, the tumour is located. These works use CNN, YOLO-V2, and image processing. Kong et al. (2022) recommends fusing AD patients' “Magnetic Resonance Images (MRI)” with “Positron Emission Tomography (PET)” scans. They employ 3D convolutional neural

networks to test our picture fusion method in dichotomous and multi-classification tasks. 3D convolution of fused pictures is utilised to extract feature information, resulting in richer multi-modal data. Finally, a fully connected neural network classifies and predicts multi-modal features. Singh et al. (2022) introduce "MISegNet", a lightweight deep learning real-time segmentation network for medical pictures. They use DWT to extract frequency-domain characteristics from input. This technique expands neurons' receptive fields in the network. They propose a "self-attention-based global context-aware (SGCA)" module with various dilation rates to widen the field of vision and indicate the relevance of each scale to improve feature discrimination. They create a "residual shuffle attention (RSA)" technique to enhance the model's visual features and construct a novel boundary-aware loss function dubbed "Farid End Point Error (FEPE)" to accurately segment ambiguous areas for edge detection. Fang et al. (2022) propose a multi-modal medical image fusion technique using "geometric algebra sparse representation (GA-SR)". First, the multi-modal medical picture is represented as a multi-vector, and the GA-SR model is introduced to eliminate channel correlation. Second, the "orthogonal matching pursuit algorithm based on geometric algebra (GAOMP)" technique is used to produce the sparse coefficient matrix. "K-means clustering singular value decomposition algorithm based on geometric algebra (K-GASVD)" obtains the geometric algebra dictionary and updates the sparse coefficient matrix and dictionary. By combining the geometric algebra dictionary and coefficient matrix, they get the fused picture. Lian et al. (2022) presented Co Cycle Reg, a technique that formulates image registration and translation collaboratively. They integrate picture registration and translation through cycle-consistency in an end-to-end training process, so each element benefits from the other. To achieve reversibility throughout the cycle, they develop a unique dual-head registration network, consisting of one backbone to extract features and two heads to forecast deformation fields. Yang et al. (2022) propose that a new deep learning-based multi-resolution multi-modal image fusion network using "iterative back-projection (IBPNet)" is created to generate high-quality fused pictures. Our IBPNet designs up-projection and down-projection blocks to convert feature mapping between high- and low-resolution pictures. Feedback mistakes from alternation are self-corrected in reconstruction. In addition, an efficient mixed loss function is devised to adapt to multi-resolution and multi-modal picture fusion applications. Chen et al. (2022) offers an unsupervised multi-level fusion network for "multi-modal medical picture fusion (MMIF)". They build a multi-layer feature compensation module to make fused pictures richer and more complete. They also design an unique fusing approach to retain subjective definition and intuitive properties while changing fused emphasis. Functional picture fusion prevents YUV processing from distorting colours. Our network is trained using a hybrid loss. Meng et al. (2022) propose the "Unified Multi-Modal Conditional Score-based Generative Model (UMM-CSGM)" to take advantage of "Score-based Generative Model (SGM)" in modelling and stochastically tasting a target probability distribution, and to extend SGM to cross-modal conditional synthesis for varied lacking setups. UMM-CSGM uses a "multi-in multi-out Conditional Score Network (mm-CSN)" to learn cross-modal conditional distributions through conditional diffusion and reverse generation in the whole modality space. In this method, the generation process may be correctly conditioned by all available information and can accommodate all missing modalities in a single network. Zhu et al. (2022) presents an imperfect multi-modality data fusion strategy for epilepsy and its subtypes. They devised an objective function that acquires the low-rank representation of the full modality portion and recovers the incomplete modality through correlation. Alternating direction multipliers may optimise the suggested model. Our technique may obtain promising classification results in detecting epilepsy and its subtypes using incomplete diffusion tensor imaging (DTI) and functional magnetic resonance imaging (fMRI)

data. Liu et al. (2022) present a "CNN-based CT and MRI image fusion method (MMAN)" that preserves more valuable information. First, a multi-scale mixed attention block extracts features. This block can collect additional information and improve channel and spatial properties. Feature maps are fused using a "visual saliency-based method". Reconstruction blocks produce the fused image.

4. PROPOSED METHODOLOGY

The suggested method employs the histogram equalization technique for image preprocessing. Decomposition transformation using the shift variant medical images is decomposed using the efficient "Discrete Wavelet Transform (DWT)". The "Multi-based Binary Residual Feature Fusion (MBRFF)" model is used to fuse or combine the medical images into a single image. Utilizing Particle Swarm Optimization (PSO), the combined medical images are optimized. The proposed multi-based binary residual feature fusion model for medical images is shown in Figure 2.

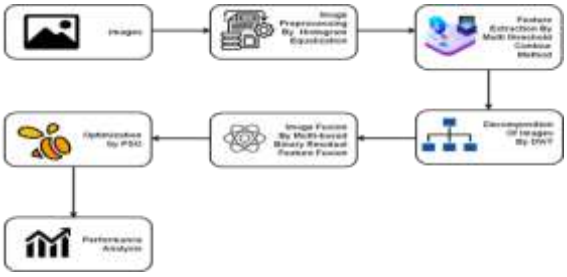

Figure 2: Proposed Multi-based Binary Residual Feature Fusion model

A. Dataset

The model makes use of chest X-rays and ultrasound images. The Houston data includes chest X-ray and ultrasound photos from Hoston University that were collected around the university of Hoston campus and the surrounding metropolitan city. Through the website, the images may be downloaded. The hyperspectral picture comprises "174 spectral bands with a wavelength range of 350 to 950 nm and measures 349 x 1905 pixels". The associated image data measures 349 by 1905 pixels. The "medical images have a spatial resolution of 2.5 m". 15 classifications were assigned to the Houston images by using modern techniques using image interpretation (Sun et al. (2021)).

B. Image Preprocessing by Histogram Equalization

One of the most well-known medical image pre-processing methods, histogram equalization (HE), is used to improve contrast in chest X-ray and ultrasound images. It effectively improves by evenly distributing the intensity values and extending the intensity range along the medical image. When its operational values are represented by a neighboring contrast picture, this approach increases the overall contrast of images. HE enables greater contrast to be obtained from smaller local differences in intensities. For a more pertinent interpretation, this approach is utilized to modify the amount of high contrast and the local details. An image's intensity spreading values can be seen as arbitrary numbers that range in value from 1 to E-1. The cumulative distribution function (CDF) associated with the random computation is another name for the estimate itself. This function establishes the likelihood that an arbitrary value will be assigned a value that is less than or equal to a given value. AHE is an algorithm that controls image intensities to enhance overall contrast. Let g be an input medical image represented as an array of numerical pixels of intensity values in the range starting with 1 and finishing with E^{-1}. Where E indicates the value of intensity probability. For this formula, q refers to the medical image's regularized histogram, and it will be as follows:

$pm = \frac{no\ of\ pixels\ with}{total\ no\ of\ pixels}$m=0,1,2,…..E-1 (1)

Let AHE image denoted as f

$f_{i,j} = flor((E-1)\sum_{m=0}^{i,j} pm)$ (2)

Consequently, the flor() changed to the closest negative integer. This is equivalent to applying the following formula to the values of the densities, i of g:

$S_a = \mathrm{flor}((E-1)\sum_{m=0}^{a} pm)$ (3)

The idea for this conversion came to mind when considering the densities for f and g as continuous arbitrary values a, b on-time ranging from 0 to E 1, with variable b being:

$B = S_a = (E-1)\int p(y)dy$ (4)

Wherever pa is the probability intensity formula of g. Where s represents the values of a's collective distribution multiplied by (E-1). Assume the variable S is differentiable and invertible for convenience of use, whereas b, indicated by the function S (A), is evenly distributed.

C. Feature Extraction using Multi-threshold Contour method

Setting threshold values using pixel differences to fully fuse images is a crucial technique in the process of processing medical image features. We may determine whether there is a connection between each characteristic by evaluating how similar the geometric sequences of various images are to one another. In contrast, because there is less connection between the two sequences, fusion is complete when the threshold value is established. It is shown that the relationship between these two characteristics is larger and cannot be divided if the distance between the two features is higher. This approach divides the edge contour elements of chest images into a white system, a black system, and a grey system. Each pixel of the chest X-ray and ultrasound image has a specific connection after contour wave conversion. The actions that follow are carried out assuming that st denotes a grey relation function of a separate sub band at the n^{-th} nodes of a medical image. First, a one-image set sequence is created by selecting several chest X-rays and ultrasound images. The second stage involves calculating the grey relation function of two neighboring sub bands. The final step involves calculating the average value of the acquired grey correlation degree and setting it to St.

Algorithm 1: Multi-threshold Contour method

The contour approach, which converts the high recurrence sub band array into one relational array of images, is used to classify the medical images into n layers.

Setting yt as one relational array of images of a sub-band in the nth layer at direction b.

By using mean processing, the starting points of several sub bands sequences in an image.

$\boldsymbol{y^*_{nb}} = \frac{y_{n,b}}{y_{n,b}(i)}, = \boldsymbol{y^*_{nb}}(1), \boldsymbol{y^*_{nb}}(2), \ldots, \boldsymbol{y^*_{nb}}(N)$ (5)

Determine the connection dimension between two neighboring sub bands.

Determine the level of sub band association in the n-th layer's neighboring direction.

$$S_{n.b} = \frac{1}{M}\sum_{a=1}^{M} S_{n.b}(a), b = 1,2,..,g \quad (6)$$

Determine the level of linkage between sub-bands in various orientations.

$$\boldsymbol{S_n = \frac{1}{g}\sum_{b=1}^{g} S_{n,a}} \quad (7)$$

The chest X-ray and ultrasound image's contour feature extraction model is then transformed into the response to a series of linear equations.

D. Decomposition of an image using Discrete Wavelet Transform

"Decomposition of medical images" is the process of extracting high pixel quality and frames from low-pixel input chest images using shared feature extractors and texture up scaling networks. Without any prior knowledge of the low-quality input image, these networks show a great capacity to create and identify a complex mapping between the low-quality input chest image and high-quality form and pattern while keeping more contextual information. Discrete Wavelet Transforms (DWT), which are effective and resilient to low-resolution images are used for a lot of decomposition tasks. A chest image displays low recurrence, smooth fluctuations with little features in between (high recurrence variation). More impulse and solicitation are present in the low recurrence portion compared to the high-frequency portion. The lower currence and higher currence components are separated using a DWT. A DWT may be obtained using a variety of filters; the Haar filter is the most basic and popular option. In our approach, the image was processed using the Haar filter. We perform the following calculations with the Haar filter to analyze the image and acquire the coefficients of each sub-band. We do calculations to determine the coefficients of each sub-band as follows,

$$LL(a,b) = \frac{q(a,b)+q(a,b+1)+q(a+1,b)+q(a+1,b+1)}{2} \quad (8)$$

$$LH(a,b) = \frac{q(a,b)+q(a,b+1)-q(a+1,b)-q(a+1,b+1)}{2} \quad (9)$$

$$HL(a,b) = \frac{q(a,b)-q(a,b+1)+q(a+1,b)-q(a+1,b+1)}{2} \quad (10)$$

$$HH(a,b) = \frac{q(a,b)-q(a,b+1)-q(a+1,b)+q(a+1,b+1)}{2} \quad (11)$$

An image is split into four sub-bands using the discrete wavelet transform: an approximation sub band (LL), and three detailed sub-bands (LH, HH, and HL), which stand for the longitudinal, orthogonal, and lateral characteristics, respectively. The initials H and L stand for high-pass and low-pass filtering when applied independently to the rows and columns. The approximation LL sub-band can be decomposed to get a high contrast representation. The applications and the kind of chest images are utilized to determine which sub-band should be used for decomposition incorporation. Without deleting any coefficients, this scanning method arranges the coefficients from low to high frequency, top to bottom, and left to right. The decomposition of our medical image is therefore achieved using the Discrete Wavelet Transform.

E. Image Fusion by Multi-based Binary Residual Feature Fusion (MBRFF)

The capacity to use medical images clinically is increased by the process of fusing several images from different imaging

*e*ISSN1303-5150 www.neuroquantology.com

techniques together. In the Multi-based Binary Residual Feature Fusion (MBRFF) design, these residue layers link the picture outcome and the probabilistic loss components of each side of the tree to two phases above on the same side.

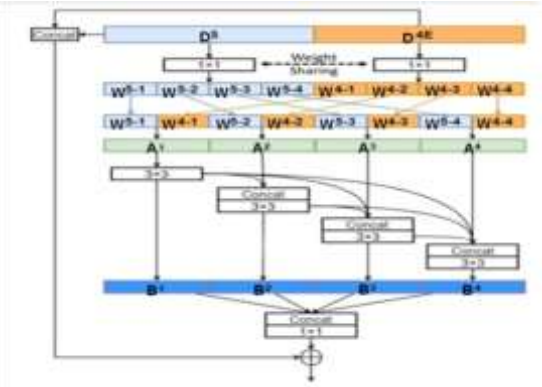

Figure 3: Architecture of proposed Multi-based Binary Residual Feature Fusion (MBRFF)

At this high level, non-quality lower parameters provide additional characteristics. As a result, in the model's tree-based structure, the residual components assist in combining the certain and improbable findings of the very low level. Figure 4.3 shows the architecture of the proposed Multi-based Binary Residual Feature Fusion (MBRFF) model. The mathematical expression for the proposed MBRFF's use of the residual feature for feature fusion and feature refinement is

$$D45F = N(D45_{concat}) + D45_{concat}; \quad (12)$$

Where N() and D45F stand-in for the convolution layers and feature concatenation, respectively, in the proposed MBRFF. To change the number of channels, two 1x1 convolution layers are first applied to D5 and D4F as shown below.

$$D4F_{norm} = E_{1x1}(D4F); \quad (13)$$

$$D5_{norm} = E_{1x1}(D5); \quad (14)$$

Two 1x1 convolution layers' weights are shared to lower because D4F and CD5 have the same number of layers, the number of training parameters. The success of such implementation is further supported by our trials, which show a 0.1 percent boost in the E1 score of the weight-sharing image when compared to a non-weight-sharing image. Multi-based representations, which are amassed across a wide variety of corresponding areas, provide advantages for more precise prediction of medical image fusion. To accomplish multi-based, we thus use numerous sub-images made up of a variety of Convolutional layers. Thus, to produce multi-based receptive fields in the proposed MBRFF, we use several sub networks made up of a variety of Convolutional layers. If the fused feature image is utilized as those sub-networks only input, it will, however, substantially increase the number of training parameters. As previously stated, this will therefore make image optimization more complex and reduce the detection accuracy. As a result, each subnet only receives a portion of the input channels in the proposed MBRFF rather than all of them.

$$S - i = concat(feat_{1+(i-2)\times(612)\backslash\backslash N},$$

$feat_{2+(i-2)\times(612)\backslash\backslash N}, \ldots, feat_{(612)\backslash\backslash N+(i-2)\times(612)\backslash\backslash N}$
(15)

After normalizing the number of channels, we use the 157 sub-band for each divide in these techniques and give M = 5 as the number of divides to fulfill our goal.

$$A_i = concat(W4 - i, W5 - i)(i = 1,2,3,4) \quad (16)$$

The intake for each split Ai is the combination of the Ai and all of the outcomes from the preceding layer, which is denoted by: where Bi is the outcome relative to Ai, and H (.) marks the convolution layer. Orders combine splits from D4F and D5, which may be stated mathematically as follows.

$$B_i = H\big(concat(A_i, B_{i-1}, B_{i-2}, \ldots., B_1)\big)(i = 1,2,3,4) \quad (17)$$

Since there isn't an output "B0" before Bi, the input is specifically Ai. In the last, the divides can be combined, and a 1x1 "convolution layer is used to change the number of channels". By employing the MBRFF model and the residual feature rule, the chest image is fused.

F. Optimization by Particle Swarm Optimization

A problem-solving method called particle swarm optimization (PSO) tries to enhance a chest image by iteratively comparing it to a set of quality images. It generates a set of potential solutions, called particles, and moves them about the search space following their position and velocity using a straightforward mathematical formula. A particle's travel is determined by its local greatest position and the search process's best-known locations, which are updated when other particles find a better image. As a simple conclusion the swarm is expected to migrate in the direction of the best possibilities.PSO starts with a collection of random particles (images).

After that, the program iteratively looks for optima. The fitness value of the particle is assessed. The particle retains the location of that value as picture best if it is the best value, it has ever obtained. Particle best refers to the position of the best fitness value that each particle has attained during any iteration. A powerful version of the PSO algorithm implies a swarm of potential images (also known as a swarm) (called particles). Some basic formulas are used to shift such particles surrounding the search area. The best places in the search area as well as the swarm's most well-known location dictate the movements of the particles. The swarm's route will be directed by any new suitable spots that are found. The procedure is repeated to get a good result; but this is not guaranteed. The optimization using a particle swarm is shown in algorithm 2.

Algorithm 2: Particle swarm optimization

For each image (particle)
Initiate the image (particle)
For each image
Do a fitness calculation
If the fitness value is higher than the historically best fitness value (best chest image), establish the current value as the new best image.
Select the image that best represents a medical image by having the highest
Fitness value among all the images.
End
For each image (particle)

Determine the image (particle) velocity.
Update the image (particle) position
Continue as long as the maximum iterations or the minimal error threshold are not met.

5. Results and Discussion

In this chapter, we will learn how to properly evaluate patients by fusing medical images using the Multi-based Binary Residual Feature Fusion (MBRFF) model. Here we evaluate the effectiveness of the suggested MBRFF model for medical image fusion. The suggested model's quality is measured using metrics including Peak Signal to Noise Ratio (PSNR), Visual Information Fidelity, Fusion factor, Fusion symmetry, Precision rate, Accuracy, F1 Score, and Recall. The findings were compared to those obtained using existing approaches. The existing methods are Stationary Wavelet Transform (SWT), Stable Graph Fourier Basis Algorithm (SGFA), Non-Subsample Contoured Transform (NSCT), and Convolutional Neural Network (CNN).

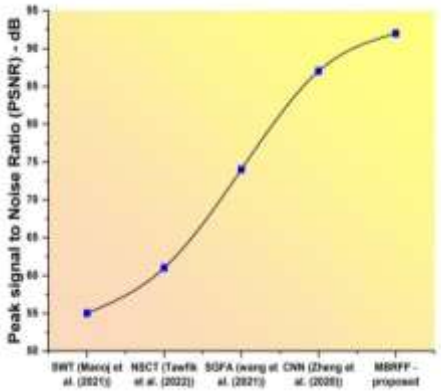

Figure 4: Peak Signal to Noise Ratio of existing and proposed methods

The reconstruction quality of loss compression in image fusion is commonly assessed using the PSNR. In this instance, the noise indicates the compression fault while the image represents the actual image. When assessing compressed images, it serves as a general measure of how well humans can rebuild an image. As a result, even when one reconstruction has a higher PSNR, it would typically signify a higher-quality reconstruction**.** Figure 4 depicts the Peak Signal to Noise Ratio of existing and proposed methods. The graph demonstrates that the recommended model has a high level of Peak Signal to Noise Ratio. The SWT received 55 percent, the NSCT received 61 percent, the SGFA received 74 percent, the CNN received 87 percent, and the recommended method received 92 percent.

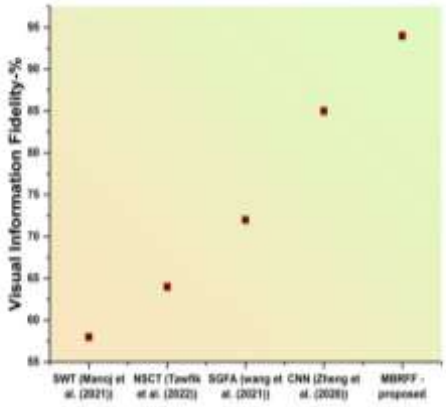

Figure 5: Visual Information Fidelity of existing and proposed methods

VIF is an essential statistic for assessing the quality of an image since it measures perceptual distortion. VIF evaluates results in the context of image fusion by determining

the amount of common information between the source and merged images. This study uses the average VIF value to assess the effectiveness of the supplied collection of approaches since VIF enables accurate distortion identification. Figure 5 depicts Visual Information Fidelity of existing and proposed methods. The graph demonstrates that the recommended model has a high level of Visual Information Fidelity. The SWT has 58 percent, the NSCT has 64 percent, the SGFA has 72 percent, the CNN has 85 percent, and the recommended method received 94 percent.

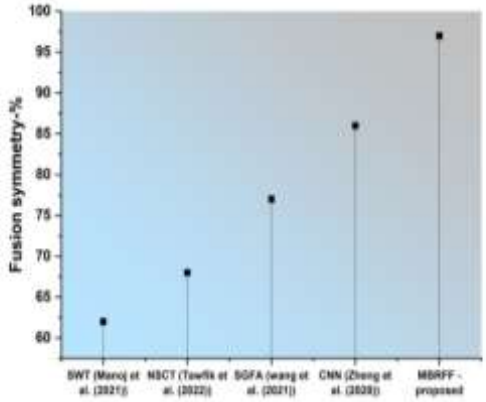

Figure 6: Fusion Symmetry of existing and proposed methods

To evaluate this metric, the Medical Image of the input and output images is employed. Higher values of obtained fusion symmetry indicate higher-quality output. Figure 6 depicts the Fusion Symmetry of existing and proposed methods. The graph shows that the suggested model has a high level of fusion symmetry. The SWT earned 62 percent, the NSCT earned 68 percent, the SGFA earned 77 percent, the CNN received 86 percent, and the recommended method earned 97 percent.

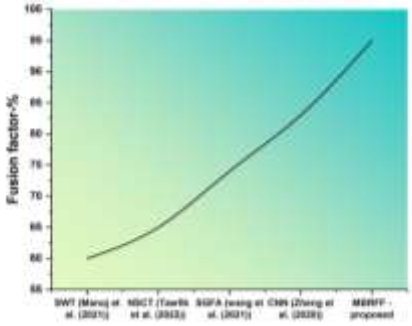

Figure 7: Fusion Factor of existing and proposed methods

By calculating how much information is present in the input and output images, or the two chest images, according to the fusion factor, the difference between the two images is assessed. The technology relies on the exchange of data between real-world medical pictures. Figure 7 depicts the Fusion Factor of existing and proposed methods. It can be seen from the graph that the suggested model has a high fusion factor. The SWT obtained 60 percent, the NSCT obtained 65 percent, the SGFA obtained 74 percent, the CNN obtained 83 percent, and the recommended method obtained 95 percent.

*e*ISSN1303-5150

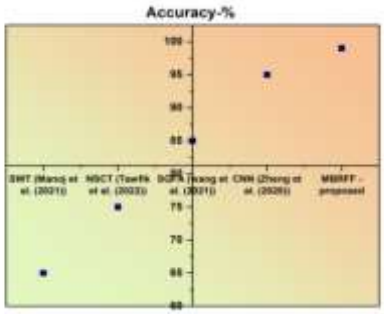

Figure 8: Accuracy of existing and proposed methods

Accuracy is one factor to consider when evaluating image fusion models. Accuracy is the proportion of predictions that our model fused properly. The correctness of every model involving the fusion of images must be assessed. A high-resolution image or fused chest image that is judged to be accurate is compared to the normal image of the chest. Figure 8 depicts the Accuracy of existing and proposed methods. The suggested model has great accuracy. The suggested approach for fusing medical images is highly accurate. The graph shows that the suggested model has high accuracy. The SWT scored 65 percent, the NSCT scored 75 percent, the SGFA scored 85 percent, the CNN scored 95 percent, and the recommended method scored 99 percent.

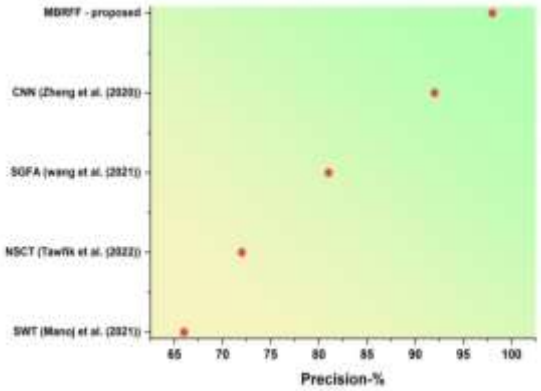

Figure 9: Precision of existing and proposed methods

Quality in a fusion model comes from its capacity to choose just the necessary chest pictures. Precision is calculated by dividing the number of correct images by the sum of the correct and incorrect ones. The number of pertinent examples out of all the examples the model found. Precision is an assessment of a model's accuracy in predicting positive labels. Figure 9 depicts the precision of existing and proposed methods. The graph demonstrates that the accuracy of the chest photos. As a result, our suggested model, MBRFF, has more precision than the current techniques. The graph depicts that the employed model has a high level of precision. The SWT received 66 percent, the NSCT received 72 percent, the SGFA received 81 percent, the CNN received 92 percent, and the recommended method received 95 percent.

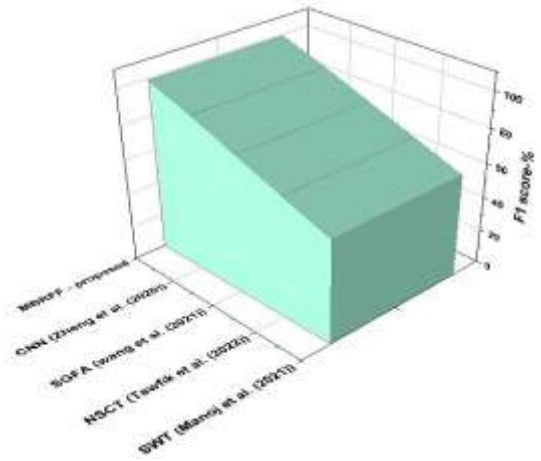

Figure 10: F1-Score of existing and proposed methods

The F-score (or F1-score) measures how well a model performs in fusing medical images. The proportion of correct predictions is found by dividing the total number of correct forecasts by the calculated accuracy of those predictions. It's possible that the most responsive model is also the one with the greatest recall. Harmonically averaging recall and accuracy yields the F1-score. Accuracy and recall are integrated into one figure. Be aware that the F1 score takes accuracy and recall into consideration, accounting for both False Positive and False Negative. Figure 10 depicts the F1-score value. Compared to other current models, the suggested MBRFF model has a higher F1 Score. The graph demonstrates the high F1-Score of the suggested model. The SWT gained 60 percent, the NSCT gained 70 percent, the SGFA gained 80 percent, the CNN gained 90 percent, and the recommended method gained 99 percent.

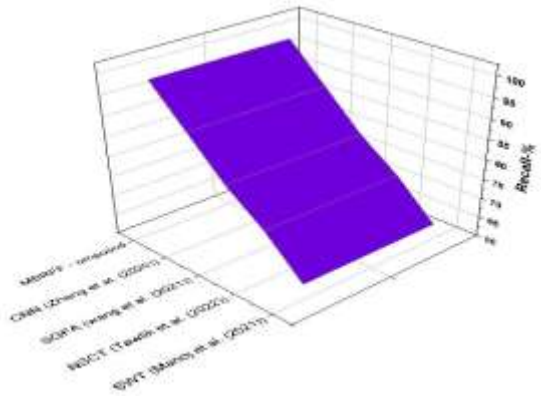

Figure 11: Recall existing and proposed methods

The recall is calculated as the percentage of all Positive images that were correctly labeled as Positive and how effectively the model can detect positive images is shown by a recall. The recall is higher the more positive images that are detected. The recall is more beneficial when missed images are more expensive than false images. Figure 11 displays the model's Recall metric value. When compared to other approaches, the MBRFF method exhibits a high recall. The graph depicts the high Recall of the proposed model. The SWT achieved 64 percent, the NSCT achieved 73 percent, the SGFA achieved 80 percent, the CNN achieved 89 percent, and the recommended method achieved 98 percent.

6. Discussion

Manoj et al. (2021) employed a "Stationary Wavelet Transform (SWT) technique for the fusion of medical images for the assessment of disease". Blurring at the pixel level has a direct impact on the image's

contrast under the SWT rule. Tawfik et al. (2022) improves the image fusion in the medical application by using the Non-Subsample Contoured Transform (NSCT) model. It uses a spatial-based method for contrasting images which will affect the visual fidelity of the medical image. Wang et al. (2021) proposed the Stable Graph Fourier basis Algorithm (SGFA) for the fusion process of the medical image in clinical assessment. It failed to combine all the information from the medical images. Zheng et al. (2020) utilizes the "Convolutional neural network" in medical assessment to detect the actual information using the fusion of medical images. It entirely loses all memory of the images' positioning and composition, and they pass the information on to a neuron that might not be able to fuse the image. As a consequence of this, we employ the Multi-based Binary Residual Feature Fusion (MBRFF) model for the fusion of the chest image to enhance multi-modality image fusion. The images are pre-processed by the Histogram equalization algorithm. The shift variant decomposition transformation technique Discrete Wavelet Transform DWT is effective and it is used to decompose the medical image. "Particle swarm optimization (PSO)" is used to optimize the medical images.

7. Conclusion

Medical image fusion is an essential part of medical treatment since accurate illness diagnosis relies on high-quality images obtained by techniques like chest X-ray and ultrasound. These medically merged images are utilized by medical professionals for accurate illness diagnosis and subsequent therapy. Two multi-modal images must be combined, which is a difficult procedure but one that produces great levels of information. The suggested approach, Multi-based Binary Residual Feature Fusion (MBRFF), efficiently completes the image fusion process and generates an output image with a high degree of perception and superior quality. This research shows that the suggested strategy is effective with respect to the Peak Signal-to-Noise Ratio, Visual Information Fidelity, Fusion Factor, Fusion Symmetry, Precision, Accuracy, Recall, and F1 Score. The overall quality beats the traditional algorithms such as Stationary Wavelet Transform (SWT), Stable Graph Fourier Basis Algorithm (SGFA), Non-Subsample Contoured Transform (NSCT), and Convolutional Neural Network (CNN).Future research on the suggested topic may focus on increasing image resolution and computing time of image decomposition. In the future, we may explore introducing optimization approaches to boost additional performance measures and improve the assessment of medical image fusion in these domains.

References

[1]. Dharejo, F.A., Zawish, M., Deeba, F., Zhou, Y., Dev, K., Khowaja, S.A. and Qureshi, N.M.F., 2022. Multimodal-Boost: Multimodal Medical Image Super-Resolution Using Multi-Attention Network With Wavelet Transform. IEEE/ACM Transactions on Computational Biology and Bioinformatics, (01), pp.1-14.

[2]. Liu, X., Yang, L., Chen, J., Yu, S. and Li, K., 2022. Region-to-boundary deep learning model with multi-scale feature fusion for medical image segmentation. Biomedical Signal Processing and Control, 71, p.103165.

[3]. Hanley, M., Brosnan, C., O'Neill, D., Ni Mhuircheartaigh, N., Logan, M., Morrin, M.M., Hurley, K., Sulaiman, I., O'Brien, E., Morgan, R. and Lee, M.J., 2021. Modified Brixia chest X-ray severity scoring system and correlation with intubation, non-invasive ventilation and death in a hospitalised COVID-19 cohort. Journal of Medical Imaging and Radiation Oncology.

[4]. Li, Y., Zhao, J., Lv, Z. and Li, J., 2021. Medical image fusion method by deep learning. International Journal of Cognitive Computing in Engineering, 2, pp.21-29.

[5]. Li, J. and Wang, Q., 2022. Multi-modal bioelectrical signal fusion analysis based on different acquisition devices and scene settings: Overview, challenges, and novel orientation. Information Fusion, 79, pp.229-247.

[6]. Wang, R., Fang, N., He, Y., Li, Y., Cao, W. and Wang, H., 2022. Multi-modal Medical Image Fusion Based on Geometric Algebra Discrete Cosine Transform. Advances in Applied Clifford Algebras, 32(2), pp.1-23.1

[7]. Zhu, Y., Wang, X., Chen, L. and Nie, R., 2022. CEFusion: Multi-Modal medical image fusion via cross encoder. IET Image Processing.

[8]. Palanisami, D., Mohan, N. and Ganeshkumar, L., 2022. A new approach of multi-modal medical image fusion using intuitionistic fuzzy set. Biomedical Signal Processing and Control, 77, p.103762.

[9]. He, B., 2022. Research on Multi-modal Medical Image Fusion Based on Multi-Scale Attention Mechanism. In Journal of Physics: Conference Series (Vol. 2173, No. 1, p. 012037). IOP Publishing.

[10].Sabbava, S.R., Boddapati, S.P., Dasari, T. and Bollam, N., 2022 CNN based multi modal medical image fusion with classification using YOLO-V2.

[11]. Kong, Z., Zhang, M., Zhu, W., Yi, Y., Wang, T. and Zhang, B., 2022. Multi-modal data Alzheimer's disease detection based on 3D convolution. Biomedical Signal Processing and Control, 75, p.103565.

[12].Singh, V.K., Kalafi, E.Y., Wang, S., Benjamin, A., Asideu, M., Kumar, V. and Samir, A.E., 2022. Prior wavelet knowledge for multi-modal medical image segmentation using a lightweight neural network with attention guided features. Expert Systems with Applications, p.118166.

[13].Li, Y., Fang, N., Wang, H. and Wang, R., 2022. Multi-Modal Medical Image Fusion With Geometric Algebra Based Sparse Representation. Frontiers in genetics, 13, p.927222.

[14].Lian, C., Li, X., Kong, L., Wang, J., Zhang, W., Huang, X. and Wang, L., 2022. CoCycleReg: Collaborative Cycle-consistency Method for Multi-modal Medical Image Registration. Neurocomputing.

[15].Liu, C., Yang, B., Zhang, X. and Pang, L., 2022. IBPNet: a multi-resolution and multi-modal image fusion network via iterative back-projection. Applied Intelligence, pp.1-17.

[16].Chen, L., Wang, X., Zhu, Y. and Nie, R., 2022. Multi-level difference information replenishment for medical image fusion. Applied Intelligence, pp.1-13.

[17].Meng, X., Gu, Y., Pan, Y., Wang, N., Xue, P., Lu, M., He, X., Zhan, Y. and Shen, D., 2022. A Novel Unified Conditional Score-based Generative Framework for Multi-modal Medical Image Completion. arXiv preprint arXiv:2207.03430.

[18].Zhu, Q., Li, H., Ye, H., Zhang, Z., Wang, R., Fan, Z. and Zhang, D., 2022. Incomplete multi-modal brain image fusion for epilepsy classification. Information Sciences, 582, pp.316-333.

[19].Liu, Y., Yan, B., Zhang, R., Liu, K., Jeon, G. and Yang, X., 2022. Multi-Scale Mixed Attention Network for CT and MRI Image Fusion. Entropy, 24(6), p.843.

[20].Ge, C., Du, Q., Sun, W., Wang, K., Li, J. and Li, Y., 2021. Deep residual network-based fusion framework for hyperspectral and LiDAR data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, pp.2458-2472.

[21].Manoj.K., Tripathi, A., Joshi, K., Sharma, A., Singh, P. and Memoria, M., 2021. A comparative review: Medical image fusion using SWT and DWT. Materials Today: Proceedings, 37, pp.3411-3416.

[22].Tawfik, N., Elnemr, H.A., Fakhr, M., Dessouky, M.I. and El-Samie, F.E.A., 2022. Multimodal Medical Image Fusion Using Stacked Auto-encoder in NSCT Domain. Journal of Digital Imaging, pp.1-18.

[23].Wang.H,Agham, N.D. and Chaskar, U.M., 2021. An advanced LAN model based on optimized feature algorithm: Towards hypertension interpretability. Biomedical Image Processing and Control, 68, p.102760.

[24].Zheng, M., Wang, K., Wei, H., Qi, G. and Li, Y., 2020. Multi-modality medical image fusion using Convolutional neural

network and contrast pyramid. Sensors, 20(8), p.2169