Songklanakarin J. Sci. Technol.
46 (4), 347–354, Jul. – Aug. 2024

http://www.sjst.psu.ac.th

Review Article

Investigation of deep learning image fusion techniques for medical imaging

Angayarkanni Veeraputhiran[1*], Muthukumar Balamurugan[2], and Varun P. Gopi[2]

[1] *Department of Computing Technologies, Faculty of Engineering and Technology, SRM Institute of Science and Technology, Chengalpattu, Tamil Nadu, 603203 India*

[2] *Department of Electronics and Communication Engineering, National Institute of Technology, Tiruchirappalli, Tamil Nadu, 620015 India*

Received: 7 July 2023; Revised: 17 March 2024; Accepted: 9 May 2024

Abstract

Image fusion is becoming increasingly important as a result of recent advancements in image processing applications because of the enormous number of different capture devices available. By fusing multiple images into one, the overall image quality is improved and key characteristics are preserved efficiently. This improvement in image quality and integrity is achieved by the use of multi-temporal, multi-view, and multi-sensor information during the process of image combination. For instance, a single-mode medical image has very little information, whereas a fused image contains more important image data and provides a reliable diagnosis of the condition. Experiments on the fusion of MRI and PET medical images are being conducted to evaluate performance, algorithm stability, and other parameters. In this article, the benefits and drawbacks of different degrees of state-of-the-art picture fusion approaches are explored, and also some of the standard quality metrics that are used for medical image fusion are computed.

Keywords: fusion, medical imaging, multi-view, multi-temporal and multi-sensor

1. Introduction

The application of medical imaging has become important in clinical diagnosis. Images that use many modalities can provide information from a variety of perspectives, which can help medical professionals verify a diagnosis and determine the most appropriate treatment. Single-photon emission computed tomography (SPECT) images show the soft-tissue structures of the body, while magnetic resonance imaging (MRI) images reveal the bone structure and details on high-density tissue; positron emission tomography (PET) images reveal the metabolic action of the cells and tissues; and Computer Tomography (CT) scans reveal the bone structure and high-density tissue information. The primary goal of the fusion is to improve the perceived quality, contrast, and perceived experience of the fusion. Some of the requirements of fusion are: (a) it must retain all of the details from the source images; (b) it should not add any new details like effects; and (c) it must avoid problems like noise and misregistration (Meher, Agrawal, Panda, & Abraham, 2019). The main objective of this paper is to provide an overview of recent changes in the field of medical image fusion research and to look at how the field might grow in the future. Medical images captured using various modalities each have advantages and limitations which lead to complicated diagnosis processes and diagnostic errors. To improve the utilization rate of medical images, image fusion algorithms can be divided into two categories: transform domain and spatial domain. Algorithms based on the transform domain include Multi-scale Transform (MST) and feature space methods such as Independent Component Analysis (ICA) and Sparse Representation (SR) (Wang, Li, Zhang, & Zhang 2018). The most important spatial domain-based algorithms can be divided into block-based, region-based, and pixel-based fusion algorithms (Ancuti, Ancuti, De Vleeschouwer, & Bovik 2016).

*Corresponding author
Email address: angayarv@srmist.edu.in

Region-based algorithms break down the input images into regions and merge them to generate the fused image, whereas block-based algorithms segment the images into blocks. Pixel-based algorithms generate a fused decision graph, while deep learning promotes the progress of image fusion due to its powerful feature extraction and data expression capabilities (Huang & Jing, 2007). Deep learning methods have been proposed for image fusion, providing theoretical support and practical experience. They have advantages over traditional image fusion algorithms such as stronger feature extraction and expression, more flexible network architecture, and end-to-end fusion process. These advantages include a stronger ability for feature extraction and expression, a more flexible network architecture, and an end-to-end fusion process (Du, Li, Lu, & Xiao, 2016). This paper is divided into the following sections: Related Works, Medical Image fusion methods using deep learning, Results and Discussion, and Conclusion.

2. Related Works

Many researchers have suggested a diverse assortment of tried-and-true methods for the process of fusing medical images. Using various decomposition approaches, these techniques can be divided into the following groups: (i) methods that are based on pyramids (Ancuti, Ancuti, De Vleeschouwer, & Bovik, 2016; Huang & Jing, 2007; Wang, Li, Zhang, & Zhang, 2018) and (ii) wavelet-based techniques (Du, Li, Lu, & Xiao, 2016) including wavelet (Thévenaz & Unser, 1996), discrete wavelet (Sahu, Bhateja & Krishn, 2014), stationary wavelet (Du, Li, Xiao, & Nawaz, 2016), dual-tree discrete wavelet (James & Dasarathy, 2014), lifting wavelet (Cheng, He, & Lv, 2008), etc. (ii) In light of the recent interest in deep learning, certain deep learning based fusion algorithms, such as Enhanced medical image fusion (EMFusion) (Bhavana & Krishnappa, 2015), offer an end-to-end unsupervised network. To maintain the integrity of the one-of-a-kind information that has been added to the surface-level constraint, a deep-level constraint is applied. Original multimodal images are input into the feature in the Multiscale Residual Pyramid Attention Network (MSRPAN) (Ganasala & Prasad, 2020), and the network generates chrominance channels to decrease mosaic. The original multimodal photos serve as input for the feature extractor, which then uses those images to generate high-dimensional features (Singh, Srivastava, Prakash, & Khare, 2012).

Once the features have been extracted, they are combined with the help of the FER technique, and the resulting combined image is then provided to the reconstructor. Within this paradigm, these three methods have been compared to one another. In magnetic resonance imaging, electromagnetic signals from within the body are collected and used to recreate images of the patient (Kor & Tiwary, 2004). The technology of magnetic resonance imaging was made possible after some discoveries by two scientists and also it was followed by a way to code nuclear magnetic resonance data spatially. Using this method, it is possible to put together pictures of people. CT and MRI are two imaging methods that display the spatial distributions of physical quantities. MRI can display spectral distributions of space in four or three dimensions, while CT can display a fault in three dimensions and shows how a certain physical amount is spread out in space. Due to the greater array of MRI imaging modalities and the complexity of the imaging concepts involved, the resulting image data will be of a greater quantity. This exemplifies the primary distinction that can be made between MRI and CT scans (Xu & Ma 2021).

3. MRI and PET Imaging

Imaging modalities such as MRI-T1 and MRI-T2, CT and MRI, SPECT and PET, SPECT and CT, and PET and CT are some examples of imaging techniques that can be merged in medical image fusion. There are many different integration techniques, each with its own set of benefits and drawbacks. For instance, MRI/PET fusion images can be helpful in the identification of brain tumors, Alzheimer's disease, and liver metastases. On the other hand, positron emission is the basis for the creation of PET pictures. Along with anatomical information, this imaging method may also show the functions and metabolisms of particular tissues. From these characteristics, doctors can identify illnesses and tumors before they develop. A particular fusion approach is required to combine PET and MRI images to create a single fused picture that contains both spatial and spectral important information since PET images are colorful and have a relatively poor spatial resolution (James & Dasarathy, 2014). In this review work, three distinct methods have been compared for analyzing MRI and PET scans.

MRI is a grayscale image while PET is a color image, making fusion processing susceptible to distortion. In most of the fusion algorithms, intensity components of a PET image are decomposed by the IHS model (Haddadpour, Daneshvar, & Seyedarabi, 2017) and BEMD, Log-Gabor transform, and other algorithms are combined to process these components to preserve more of the PET image's color. (Yin *et al.*, 2018) proposed an MRI and PET image fusion algorithm based on nonsubsampled shearlet transform (NSST) and simplified pulse-coupled neural network model (S_PCNN), which transforms the PET image to the YIQ component and employs NSST to break down the MRI and the Y component of PET into low-frequency and high-frequency subbands.

3.1 MRI

The imaging principle is that the nucleus has positive electricity and the atomic nuclear energy of many elements' spins. When the magnetization vector of the spin system gradually increases from zero, the magnetization intensity reaches equilibrium. After the radio frequency pulse has stopped, the nucleus will revert to its original arrangement state in the magnetic field, releasing weak energy that will transform into a radio signal and be detected. MRI can disclose the form of the brain's soft tissues but cannot determine the brain's functional composition. With high proton densities of soft tissue, fat, the nervous system, and articular cartilage lesions, the image is very clear and does not produce any artifacts. Its high spatial resolution, absence of radiation's harmful effects on the human body, and informational advantage all play a vital part in clinical diagnostics. Since the proton density in bone is extremely low, it is not possible to obtain an accurate picture of bone using MRI (Asif, Bennamoun, & Sohel, 2017).

3.2 PET

The image obtained from PET reveals accurate details on blood flow and can precisely localize where the lesion on the patient is situated. Photons are created as a result of the collision of electrons in the flesh with positrons, which is the fundamental concept behind this theory. While PET is sensitive, it is difficult to collect exact information about the location of the brain's structural components. The purpose of PET is to count photons and build an image with information about brain function that is adequate for cancer identification. Since there is no resolution of the boundaries between soft tissue and bone, the spatial resolution is quite poor, and there will likely be spatial distortion.

4. Medical Image Fusion Methods Using Deep Learning

4.1 MRI and PET fusion

PET generates a color image, whereas MRI generates a grayscale image; the color image generated by PET is easily distorted during the fusion processing. BEMD, the Log-Gabor transform, and several additional algorithms are utilized in the color processing and preservation of the PET image. Alternatively, these methods can be combined. The intensity components of the PET image are decomposed by the IHS model, which is used in fusion methods (Lahoud & Susstrunk, 2019). The MRI and PET image fusion approach that was proposed by (Yin *et al.*, 2018) converts the PET picture into the YIQ component before utilizing NSST to split the Y component of the PET into low-frequency and high-frequency subbands based on NSST and PCNN. This is done before the MRI image is fused with the PET image. The high-frequency coefficients of the image were processed by employing a simplified version of the PCNN model, and as a consequence, the image possesses great effects, such as very little color distortion and a wealth of structural information.

For the preprocessing of MRI and PET image fusion, (Wang *et al.*, 2006) have presented a method predicated on the discrete wavelet transform. This method was successful in overcoming the decline in image quality as well as the unreadability of the raw files, and it achieved an accuracy of 90–95% in the fusion process. Shearlet transformation and discrete cosine transform were combined to create a brand-new fusion method that was suggested by (Chaitanya *et al.*, 2016). In order to maximize filter coefficients, adaptive filters were first applied to the merging of MRI-PET images by Saboori and Birjandtalab (2019). Utilizing both spatial and spectral difference parameters, they achieved this. There are additional MRI/PET fusion methods (Bhavana & Krishnappa, 2016; Shabanzade & Ghassemian, 2017; Shahdoosti & Mehrabi, 2018. Alzheimer's disease is frequently diagnosed using MRI/PET imaging, which frequently requires the fusion of MRI and PET pictures. MRI/PET is a vital part of tumor identification. In clinical oncology, PET/MRI may eventually become a powerful multimodal technique.

4.2 EMFusion

An unsupervised improved fusion model is utilized in this EMFusion effort in order to address a problem involving the fusing of multiple types of medical images. For the purpose of merging medical images, it is recommended to make use of an end-to-end autonomous network. It adds surface-level constraints as well as deep-level restrictions to the information to keep it secure. The activity level of the source images is determined at the surface level, by doing a rigorous analysis of the salience and abundance of the images. When dealing with multi-modal images, the many different representations and measuring approaches are more suitable. For objectively quantifying uniqueness, neural networks are utilized in the deep-level constraint. It does not explicitly keep the chrominance information in functional images; rather, it improves the preservation of unique details using a fusion network. This is in place of overtly maintaining the chrominance information. The mosaic was reduced by leveraging the texture information obtained from MRI scans, which resulted in an improvement in chrominance (Xu & Ma 2021). The EMFusion framework is shown in Figure 1. The main framework is split into two phases, namely Training phase and Testing phase as explained below.

4.2.1 Training phase

1. Initialize the Transnet, encoder-decoder network, and FusionNet

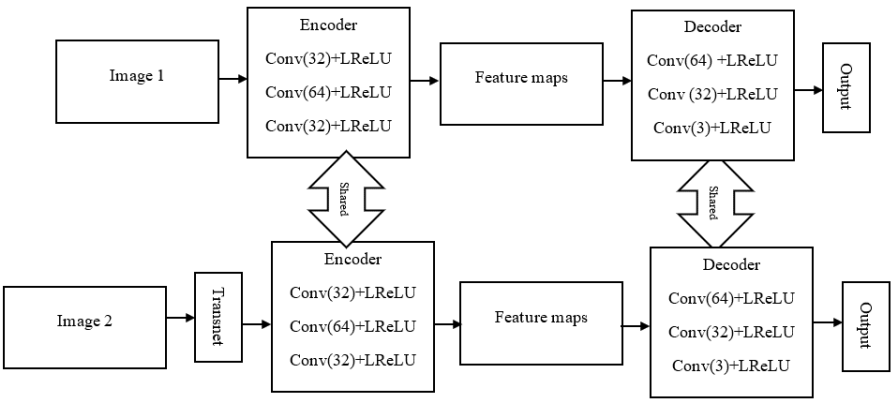


Figure 1. EMFusion framework

2. Update the parameters of Transnet

3. Fix Transnet, and update the parameters of the encoder–decoder network

4. The n channels are obtained from the output of the encoder as the unique channels

5. Fix Transnet and the encoder–decoder network, and update the parameters of FusionNet. In this step, the encoder is used as a feature extractor to provide a deep-level constraint for FusionNet.

4.2.2 Testing phase

To generate the fused image, source images should be fed into FusionNet. In this phase, encoder–decoder network and Transnet are not required.

4.3 MSRPAN fusion

A novel multiscale residual pyramid attention network-based fusion technique is analyzed for use with medical images. For the feature extractor to collect high-dimensional features, it must initially be fed the original multimodal images. Next, the FER technique is used to the retrieved features in order to fuse them, and the resulting fused picture is then provided to the reconstructor. When the training process is complete, the original photos are input, and the fused image is instantaneously received without any more parameter changes or setting adjustments being made. The numerous head-to-head comparisons of the studies reveal that the recommended approach exceeds the usual algorithm in the majority of the metrics and in terms of visual quality. It possesses both the properties of residual attention networks and pyramid attention networks. The design of the MSRPAN network is flexible in that its blocks can be altered to accommodate a variety of additional fusion operations (Fu, Li, Du & Huang 2021). Figure 2 depicts the MSRPAN algorithm's design. The three components are the feature extractor, fuser, and reconstructor. The high-dimensional features in the source images are extracted using the feature extractor, and the feature fuser is used to merge these extracted characteristics. The reconstructor then incorporates the fused features into the fused image. The feature reconstructor's job is to convert the fused features back into fusion images. Three MSRPAN blocks make up the feature extractor, while three convolution layers make up the reconstructor.

The Algorithm is given by

Input: Image 1 (OI_1) and Image 2 (OI_2)

Output: Fused image (FI)

Step 1: Input the training data into the MSRPAN.

Step 2: Once the training epochs reach 500, the training process is finished and the fusion model is obtained.

Step 3: Feed OI_1 and OI_2 to the fusion model and the features Fe1 and Fe2 are obtained respectively. Then use the feature fuser to fuse Fe1 and Fe2 and the fused feature (FF) is obtained.

Step 4: Input FF into the feature reconstructor of the fusion model and the fused image (FI) is obtained directly.

4.4 ZL Image fusion

The primary building blocks of the zero-learning image fusion technique rely on pre-trained neural networks and saliency analysis. The process involved in zero-learning image fusion is shown in Figure 3. In contrast to other neural network-based approaches, this one may be applied to a wide variety of fusion applications and does not require any training on the image modalities to be completed beforehand. Based on pre-trained CNN an innovative and speedy method for the process of picture fusion is proposed. At first, the source images are separated into a base layer and a detail layer. A pre-trained CNN with feature maps is used to combine the detail layers and visual saliency is employed to combine the base levels. The final weights map accurately reflects the source images and is ensured by a guided filter that makes the necessary adjustments. Once both the basic and detail levels have been fused, the final unified image can be created (Lahoud & Süsstrunk 2019). The block diagram of the ZL image fusion method is shown in Figure 3. The image is split into detail and base fusion layers using the two-scale decomposition technique. The detail layers are built on CNN intermediate feature maps, and the base layers are fused based on a comparison of saliency measures. The weight maps are smoothed using the guided filter, which also ensures that they are consistent with the original pictures. The merged base and detail layers are linked to create the ultimate fusion in the final stage.

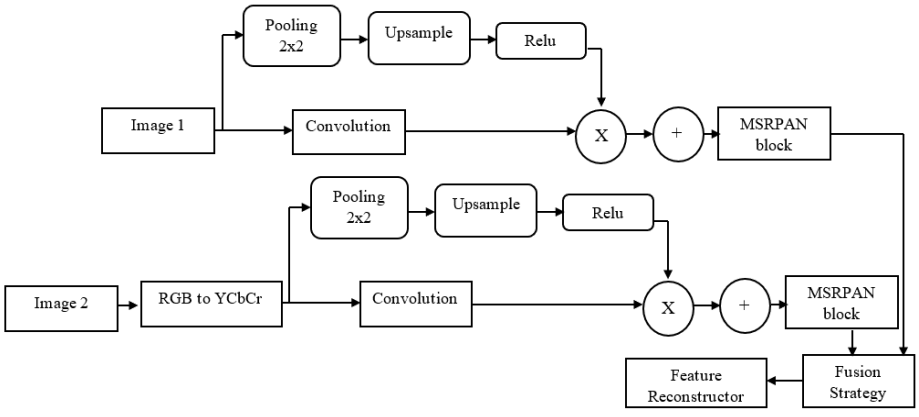


Figure 2. MSRPAN fusion framework

5. Results and Discussion

The MRI-PET fusion images using three fusion approaches are shown in Figure 4. The data sets are taken from (https://www.med.harvard.edu/aanlib/). Figures 4a and 4b show the sample MRI and PET images. The fusion of MRI and PET images combines structural and functional information, providing clinicians with a comprehensive view of the patient's anatomy and physiology. This integrated approach enhances diagnostic accuracy, improves treatment planning, and facilitates monitoring of disease progression and treatment response.

Figure 4a displays an MRI slice that was captured using various radiation techniques. The PET slices in Figure 4b were constructed after exposure to various photon radiations. This integrated approach enhances diagnostic accuracy, improves treatment planning, and facilitates monitoring of disease progression and treatment response. Figures 5a, 5b and 5c depict the fused images of EMFusion, MSRPAN, and ZL methods respectively. To examine the precise pixel values on the tumor area and obtain in-depth information about the tissue in the brain, the brain slices are gathered with various radiation excitations on the brain tissue.

The exact locations of the brain tissues are shown in Figure 5a of EMFusion model, which is obtained when combining MRI and PET images. The effectiveness of fusion techniques can be assessed quantitatively using localization measures. In medical imaging applications, techniques that improve feature localization such as identifying tumours are typically seen as more efficient. For fusion techniques such as EMfusion, localization is critical to their overall effectiveness and dependability. This is especially true in medical imaging scenarios where precise spatial feature representation is critical. The MSRPAN fusion of MRI and PET results is shown in Figure 5b. In comparison with EMFusion, MSRPAN displays better metrics, as summarized in Table 2. The image of ZL's fusion is shown in Figure 5c. Better results are being produced by ZL model fusion than by EMFusion, but it is not superior to MSRPAN. A comprehensive examination of three fusion methods allowed us to show that MSRPAN provides better metrics.

The metric's name and its abbreviation with its equations are shown in Table 1. The results are summarized in Table 2 for image fusion metrics such as Mutual Information (MI), Tsallis Entropy (TE), Nonlinear Correlation Information Entropy (NCIE), Gradient (G), Multi-Scale Analysis (M), Spatial Frequency (SF), Phase Congrency (P), Peilla-Select only one (S), Chen's algorithm (C), Yang's algorithm (Y), Chen-Varshney algorithm (CV), and Chen-Blum algorithm (BC), for the three methods compared in this work (Liu, Blasch 2011). We can create more reliable and efficient image fusion algorithms for a range of applications, such as computer vision, remote sensing, medical imaging, and image analysis, by combining these techniques.

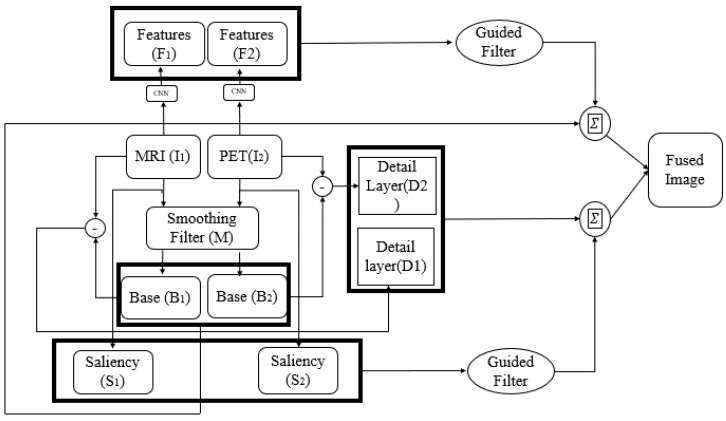
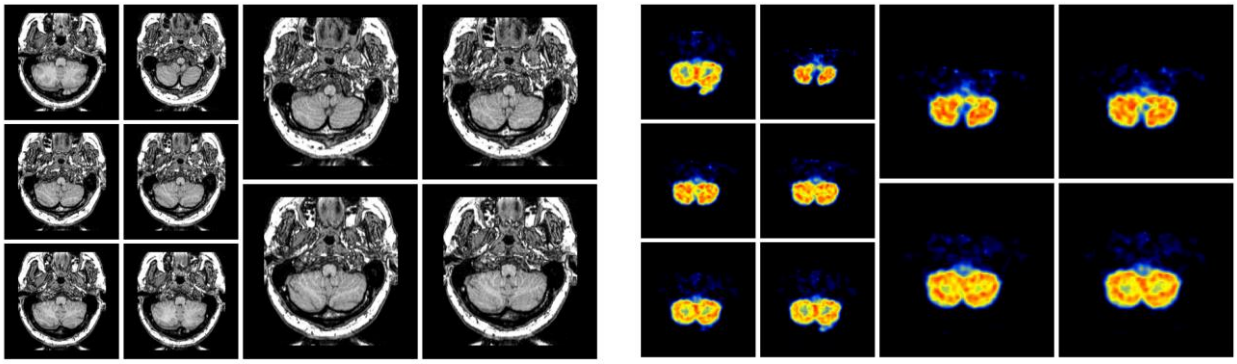
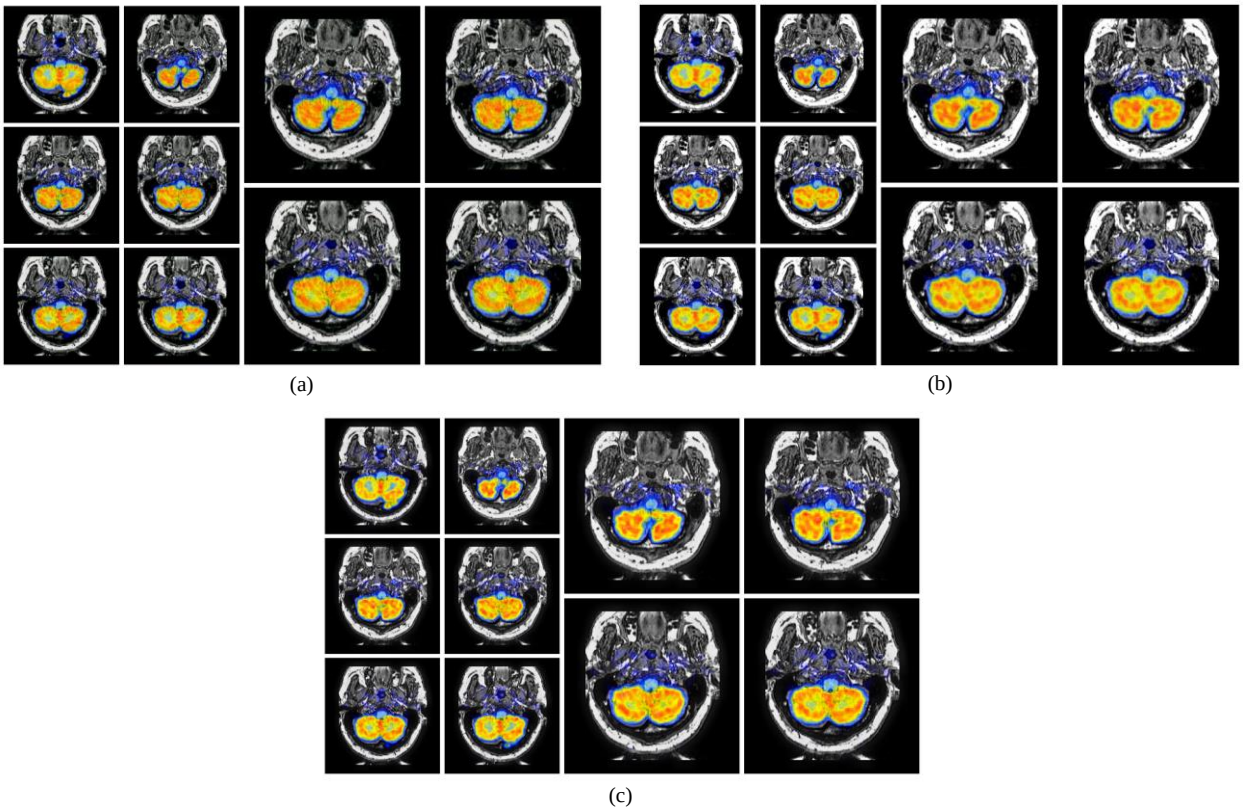


Figure 3. ZL image fusion framework



(a) (b)

Figure 4. (a) MRI samples, (b) PET samples



(a) (b)

(c)

Figure 5. (a) EMFusion, (b) MSRPAN fusion, (c) ZL fusion

Table 1. Parameters and their defining equations

Metric name	Metric equation
MI (Mutual information)	$I(X;Y) = H(X) - H(X/Y)$
TE (Tsallis Entropy)	$S_q = -k\frac{1-\sum_{i=1}^{W} P_i^q}{1-q}$
NCIE (Nonlinear correlation information entropy)	$NCC(U,V) = H'(U) + H'(V) - H'(U,V)$
G (Gradient)	$Q_{NCIE} = 1 + \sum_{i=1}^{3}\frac{\lambda_i}{3} log_b \frac{\lambda_i}{3}$
M (Multi-scale analysis)	$Q_M = \prod_{s=1}^{N}(Q_s^{AB/F})^{\alpha_s}$
P (Phase congruency)	$PC(t) = \max_{\phi}\frac{\sum_n A_n \cos(\phi_n(t) - \bar{\phi}}{\sum_n A_n}$

From the table, it is observed that the MSRPAN method shows higher values compared to the other methods. MSRPAN Fusion is a medical imaging tool that offers enhanced accuracy, improved visualization, optimized feature fusion, and efficient workflow. Its superior performance in metrics like PSNR and SSIM results in images with higher fidelity and similarity to ground truth images, reducing false positives and negatives. It also provides better contrast, sharper edges, and improved visual quality, aiding in accurate interpretation and decision-making in medical diagnoses. The significance of Table 2 is that we can create more reliable and efficient image fusion algorithms for a range of applications, such as computer vision, remote sensing, medical imaging, and image analysis, by combining these techniques. The computational resources, memory, and processing power required for each algorithm are mentioned in Table 3.

6. Conclusions

In this study, image fusion based on deep learning techniques is proposed as a method for analyzing multimodal medical images. The most essential properties from data are easily extracted by deep learning algorithms, saving developers the trouble of having to manually construct the system. In this paper, the fusion of MRI and PET images utilizing EMFusion, MSRPAN, and ZLfusion was analyzed and compared, along with a discussion of the benefits and drawbacks of each method. In addition, the values of parameters including Mutual Information (MI), Tsallis Entropy (TE), Nonlinear Correlation Information Entropy (NCIE), Gradient (G), and others, have been reviewed. When compared to EMFusion and ZLfusion, it has been discovered that the MSRPAN technique performs significantly better. Subsequent investigations in this field are anticipated to concentrate on improving quantitative analysis techniques, discovering novel imaging biomarkers, and developing imaging technology in order to augment the therapeutic usefulness of MRI-PET fusion in diverse medical contexts. The importance of EMFusion, MSRPAN Fusion, and

Table 2. MRI-PET fusion comparison of three approaches

Fusion methods	Parameters											
	MI	TE	NCIE	G	M	SF	P	S	C	Y	CV	CB
EMFusion	0.6206	0.5376	0.8074	0.7317	0.2808	-0.1014	0.7065	0.9274	0.8475	0.8635	74.99	0.5837
MSRPAN	0.6698	0.5045	0.8077	0.7607	0.3752	-0.0366	0.6726	0.9233	0.8332	0.8674	262.82	0.5895
ZL Fusion	0.5792	0.6195	0.8073	0.6409	1.0378	-0.0498	0.6903	0.8023	0.7975	0.8208	92.74	0.4846

Table 3. Computational resources, memory, and processing power required by each algorithm

Resources	EMFusion	MSRPAN	ZL
GPU	NVIDIA Geforce GTX Titan X	NVIDIA GeForce GTX 1050 Ti GPU and 8 GB of RAM	NVIDIA GeForce GTX 1050
CPU	Windows 3.4 GHz Intel Core i5	Intel Core i5− 8400 CPU with 6 core	Windows 2.8GHz Intel Core i7-7700HQ
Training parameters	PyTorch Epochs - 500 Batch size - 64	Pytorch Epochs - 300 Batch size - 32	Pytorch Epochs- 300 Batch size - 64

ZLFusion, in contrast to other image fusion algorithms, is found in their tailored methods for addressing certain requirements or obstacles in image fusion applications. To improve fusion results in their respective fields of application, they use various strategies like edge preservation, deep learning with attention mechanisms, and low-light picture augmentation, respectively. That can be pursued in future work and in extension of this research work.

References

Ancuti, C. O., Ancuti, C., De Vleeschouwer, C., & Bovik, A. C. (2016). Single-scale fusion: An effective approach to merging images. *IEEE Transactions on Image Processing*, *26*(1), 65-78. doi:10.1109/TIP.2016.2621674

Asif, U., Bennamoun, M., & Sohel, F. A. (2017). A multi-modal, discriminative and spatially invariant CNN for RGB-D object labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *40*(9), 2051-2065. doi:10.1109/TPAMI.2017.2747134

Bhavana, V., & Krishnappa, H. K. (2015). Multi-modality medical image fusion using discrete wavelet transform. *Procedia Computer Science*, *70*, 625-631. doi:10.1016/j.procs.2015.10.057.

Bhavana, V., & Krishnappa, H. K. (2016). Fusion of MRI and PET images using DWT and adaptive histogram equalization. *Proceeding of 2016 International Conference on Communication and Signal Processing* (pp. 0795-0798). doi:10.1109/ICCSP.2016.7754254

Chaitanya, C. K., Reddy, G. S., Bhavana, V., & Varma, G. S. C. (2017). PET and MRI medical image fusion using STDCT and STSVD. *Proceeding of 2017 International Conference on Computer Communication and Informatics* (pp. 1-4). doi:10.1109/ICCCI.2017.8117685

Cheng, S., He, J., & Lv, Z. (2008). Medical image of PET/CT weighted fusion based on wavelet transform. *Proceeding of 2008 2nd International Conference on Bioinformatics and Biomedical Engineering* (pp. 2523-2525). doi:10.1109/ICBBE.2008.964

Du, J., Li, W., Lu, K., & Xiao, B. (2016). An overview of multi-modal medical image fusion. *Neuro computing*, *215*, 3-20. doi:10.1016/j.neucom.2015.07.160

Du, J., Li, W., Xiao, B., & Nawaz, Q. (2016). Union Laplacian pyramid with multiple features for medical image fusion. *Neurocomputing*, *194*, 326-339. doi:10.1016/j.neucom.2016.02.047

Fu, J., Li, W., Du, J., & Huang, Y. (2021). A multiscale residual pyramid attention network for medical image fusion. *Biomedical Signal Processing and Control*, *66*,102488. doi:10.1016/j.bspc.2021.102488

Ganasala, P., & Prasad, A. D. (2020). Medical image fusion based on laws of texture energy measures in stationary wavelet transform domain. *International Journal of Imaging Systems and Technology*, *30*(3), 544-557. doi:10.1002/ima.22393

Haddadpour, M., Daneshvar, S., & Seyedarabi, H. (2017). PET and MRI image fusion based on combination of 2-D Hilbert transform and IHS method. *Biomedical Journal*, *40*(4), 219-225. doi:10.1016/j.bj.2017.05.002 Retrieved from https://www.med.harvard.edu/aanlib/

Huang, W., & Jing, Z. (2007). Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters*, *28*(4), 493-500. doi:10.1016/j.patrec.2006.09.005

James, A. P., & Dasarathy, B. V. (2014). Medical image fusion: A survey of the state of the art. *Information Fusion*, *19*, 4-19. doi:10.1016/j.inffus.2013.12.002

Kor, S., & Tiwary, U. (2004). Feature level fusion of multimodal medical images in lifting wavelet transform domain. *Proceeding of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Volume 1, pp. 1479-1482). doi:10.1109/IEMBS.2004.1403455

Lahoud, F., & Süsstrunk, S. (2019). Fast and efficient zero-learning image fusion. Retrieved from https://doi.org/10.48550/arXiv.1905.03590

Liu, Z., Blasch, E., Xue, Z., Zhao, J., Laganiere, R., & Wu, W. (2011). Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *34*(1), 94-109. doi:10.1109/TPAMI.2011.109

Meher, B., Agrawal, S., Panda, R., & Abraham, A. (2019). A survey on region based image fusion methods. *Information Fusion*, *48*, 119-132. doi:10.1016/j.inffus.2018.07.010

Saadat, S., Pickering, M. R., Perriman, D., Scarvell, J. M., & Smith, P. N. (2017). Fast and robust multi-modal image registration for 3D knee kinematics. *Proceeding of 2017 International Conference on Digital Image Computing: Techniques and Applications* (pp. 1-5). doi:10.1109/DICTA.2017.8227434

Saboori, A., & Birjandtalab, J. (2019). PET–MRI image fusion using adaptive filter based on spectral and spatial discrepancy. *Signal, Image and Video Processing*, *13*, 135-143.

Sahu, A., Bhateja, V., & Krishn, A. (2014). Medical image fusion with Laplacian pyramids. *Proceeding of 2014 International Conference on Medical Imaging, M-Health and Emerging Communication Systems* (pp. 448-453). doi:10.1109/MedCom.2014.7006050

Shabanzade, F., & Ghassemian, H. (2017). Combination of wavelet and contourlet transforms for PET and MRI image fusion. *Proceeding of 2017 Artificial Intelligence and Signal Processing Conference* (pp. 178-183). doi:10.1109/AISP.2017.8324077

Shahdoosti, H. R., & Mehrabi, A. (2018). MRI and PET image fusion using structure tensor and dual ripplet-II transform. *Multimedia Tools and Applications*, *77*, 22649-22670. doi:10.1007/s11042-017-5067-1

Singh, R., Srivastava, R., Prakash, O., & Khare, A. (2012). Multimodal medical image fusion in dual tree complex wavelet transform domain using maximum and average fusion rules. *Journal of Medical Imaging and Health Informatics*, *2*(2), 168-173. doi:10.1166/jmihi.2012.1080

Thévenaz, P., & Unser, M. (1996, September). A pyramid approach to sub-pixel image fusion based on mutual information. *Proceedings of the 3rd IEEE International Conference on Image Processing* (Volume 1, pp. 265-268). doi:10.1109/ICIP.1996.559484

Wang, A., Sun, H., & Guan, Y. (2006, April). The application of wavelet transforms to multi-modality medical image fusion. *Proceedings of 2006 IEEE International Conference on Networking, Sensing and Control* (pp. 270-274). doi:10.1109/ICNSC.2006.1673156

Wang, J., Li, X., Zhang, Y., & Zhang, X. (2018). Adaptive decomposition method for multi-modal medical image fusion. *IET Image Processing*, *12*(8), 1403-1412. doi:10.1049/iet-ipr.2017.1067

Xu, H., & Ma, J. (2021). EMFusion: An unsupervised enhanced medical image fusion network. *Information Fusion*, *76*, 177-186. doi:10.1016/j.inffus.2021.06.001

Yang, Q., Yan, P., Zhang, Y., Yu, H., Shi, Y., Mou, X., . . . Wang, G. (2018). Low-dose CT image denoising using a generative adversarial network with Wasserstein distance and perceptual loss. *IEEE Transactions on Medical Imaging*, *37*(6), 1348-1357. doi:10.1109/TMI.2018.2827462

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 Transactions on Instrumentation and Measurement*, *68*(1), 49-64. doi:10.1109/TIM.2018.2838778