Rev. Roum. Sci. Techn.– Électrotechn. et Énerg.
Vol. **70**, *2*, pp. 281–286, Bucarest, 2024

DGN-TBMF: DUAL GENERATOR NETWORK BASED ON TRI-BRAIN MODAL FUSION FOR ACCURATE BRAIN DISEASE DIAGNOSIS

THIRU GNANAM [1,*], AHILAN APPATHURAI [2], JASMINE GNANAMALAR [3], MUTHUKUMARAN NARAYANAPERUMAL[4]

Keywords: Brain disease; Image fusion; Tri-modals; Deep learning; Fusion rules; Discrete Shearlet transform.

Medical image fusion techniques are frequently used in a variety of applications. This fusion technology enables specialists to access images that incorporate anatomical and physiological data. It has been used in many clinical settings to fuse medical images of the brain for the diagnosis of brain diseases. Several methods have been proposed to fuse medical brain images, but these models need to be enhanced in terms of efficiency. This work employs a novel dual generator network-based tri-brain modal fusion (DGN-TBMF) framework to accurately predict brain diseases using tri-modality images, including MRI, CT, and PET. Initially, the gathered MRI and CT images are pre-processed using a scalable range-based adaptive bilateral (SCRAB) filter to reduce the noise artifacts. PET images are split into high and low-frequency components by the discrete Shearlet transform (DST). The proposed DGN-TBMF approach comprises two generators and a detector module. The first Generator consists of dilated convolutional layers for extracting the relevant grey matter densities and cortical thickness from MRI and CT images. Similarly, the second generator extracts the relevant voxel intensities from PET images. The image fusion is performed using four fusion rules, and these images are taken as input to the deep learning-based detector for accurately detecting brain abnormalities. According to the experimental results, the proposed DGN-TBMF performed effectively in both quantitative and qualitative analyses, yielding respective values. The accuracy achieved by the proposed DGN-TBMF network is 99.25 % for dataset-1 and 99.04 % for dataset-2.

1. INTRODUCTION

The brain is the primary component of the neurological system, which controls all the functions of the human body. It is one of the most complex organs in the human body and is covered by the skull [1]. Furthermore, it can spread to other body organs and affect human activities. Medical image fusion (MIF) aims to extract as much important data as possible from raw images to maximize the use of medical images and assist clinicians with image processing [2]. The MIF method is gaining significance in medical diagnosis and therapy. Multi-modal MIF techniques integrate additional data from multiple raw images to generate a fused image for viewing, which helps doctors make better decisions for various purposes and facilitates the early diagnosis of brain diseases [3,4].

Image fusion guided in disease prognosis has been attempted in real-world applications to assist clinicians with making judgments due to the subjectivity of human interpretation of medical images [5]. The main objective of multi-modality MIF is to combine the complementary information derived from many registered source images to produce higher-quality data [6]. PET images with molecular imaging techniques offer excellent specificity but limited resolution [7]. NSCT requires less processing than other decomposition algorithms [8,27]. The single IHS fusion technique involves replacing a PET image, a low-resolution intensity factor in IHS space, with grey-level MRI images for the detection of different brain diseases [9, 10]. Several advantages distinguish the NSST method from other sparse decomposition methods. The NSST algorithm removes all artifacts from medical images while retaining the unique soft tissue characteristics. To address the issue of the starting weight not applying to medical images, the initial weight is expressed as high-frequency and low-frequency coefficients in the multiscale domain [11, 28]. In recent years, several machine learning [12] and deep learning [13] techniques have been employed in the identification of brain diseases using MRI images [14]. This work presents a novel dual generator network-based tri-brain modal fusion (DGN-TBMF) model for accurately predicting brain disease in its early stages using tri-modality images, including MRI, CT, and PET.

The remaining work has been scheduled as follows in advance: section 2 provides an overview of the literature on brain tri-modality image fusion techniques; section 3 delineates the inclusive work of the suggested DGN-TBMF method; section 4 recounts the experimental findings and their analysis; and section 5 concludes with recommendations for future research.

2. LITERATURE SURVEY

Medical multimodality, classified by computer-based techniques, has allowed researchers to develop various methods for evaluating brain abnormalities. The development of various artificial intelligence techniques and brain disease diagnosis has become simpler in recent years.

In 2021, Li et al. [15] designed a deep learning-based multi-modal medical image to achieve multi-modal MIF based on MRI, CT, and SPECT images. It can be applied to multi-modal MIF scenarios and overcomes the limitation of processing on a single page. In 2022, Alseelawi et al. [16] suggested a low-cost multimodal MIF method based on an NSCT hybrid. A neural network was used in this method to generate a weight map based on the movement of pixels in two or more multi-modal images, such as MRI, CT, and PET scans. In 2021, Kaur et al. [17] devised a deep learning approach for multi-modal image fusion based on multi-objective differential evolution. The Inception algorithm was used to extract features from raw images. The multi-objective differential progression was utilized to select the optimal features.

In 2021, Zeyu et al. [18] designed a multimodal MIF approach based on NSCT and CNN. The NSCT was utilized to enhance fusion solutions, and a CNN was employed to extract features from various frequency subbands, creating

[1,*]Department of Electronics and Communication Engineering, Government College of Technology, Coimbatore, Tamil Nadu 641013 India. (Correspondence)
[2] Department of Electronics and Communication Engineering, PSN College of Engineering and Technology, Tirunelveli – 627152, India.
[3] Department of Electrical and Electronics Engineering, PSN College of Engineering and Technology, Tirunelveli, India.
[4] Centre for Computational Imaging and Machine Vision, Department of Electronics and Communication Engineering, Sri Eshwar College of Engineering, Coimbatore – 641 202, Tamil Nadu, India.
Emails: thirugnanam_g@gct.ac.in, akhilanappathurai@psncet.ac.in, jasmine@psncet.ac.in, : muthukumaran.n@sece.ac.in

DOI: 10.59277/RRST-EE.2025.70.2.22

decision maps and improving the quality of fused images. Then, a perceptual high-frequency CNN (PHF-CNN) is utilized to observe and choose high-frequency coefficients. The proposed approach achieves an overall accuracy of 95.62%. In 2021, Dinh et al. [19] proposed a fusion of multimodal medical images. The three-scale decomposition (TSD) extracts the base and detail layers from the input images. Second, the local energy function via the Kirsch compass operator approach was utilized to fuse detailed layers, preserving critical details in the fused image.

In 2020, Zhang et al. [20] developed an image fusion framework with a convolutional neural network (IFCNN). The image fusion techniques extract essential visual information from many input images using two convolutional layers. The informative fused image was formed by recreating the combined features via two convolutional layers. In 2020, Tan et al. [21] proposed a deep learning-based NSST method for multi-modal MIF using boundary-measured pulse-coupled neural networks (PCNN) and energy attribute (EA) fusion algorithms. The key benefit of the two fusion procedures was used at different scales.

In 2020, Wang et al. [22] devised a MIF technique based on CNN. It integrates the pixel activity information of source images using a trained Siamese convolutional network. A contrast pyramid was used to divide the source image. The dissimilar spatial frequency bands and a weighted fusion procedure merged the source images. In 2019, Xia et al. [23] proposed a multi-scale transformation and a DCNN-based image fusion method for multimodal medical images. The Gaussian and Gauss-Laplace filters divide the original image into multiple images. This fusion technique achieves 62.76% greater accuracy than CT and MR images. In 2018, Rajalingam et al. [24] proposed a Siamese CNN to create weight maps combining pixel moment data from various medical multi-modalities. Multimodal fusion strategies yield the best-fused images, the lowest processing time, and the most accurate medical image representation.

According to the literature review above, numerous researchers have reported that image fusion can improve the quality of medical images. However, existing works are unable to overcome the limitations of MRI and CT image fusion due to the inefficiency of the high-frequency component, resulting in the loss of critical information, such as edges, in the fused output images [29]. To overcome these issues, this research work incorporates multimodal images, such as MRI, CT, and PET scans, to identify brain abnormalities using deep learning networks.

3. **PROPOSED METHODOLOGY**

This section presents a novel DGN-TBMF to merge medical tri-modality images and identifying brain diseases in their early stages. Figure 1 illustrates the overall workflow of the proposed model.

3.1 DATASET DESCRIPTION

This work uses two databases, namely OASIS and MIDAS, due to their broad dataset availability. Different OASIS datasets are available, and the OASIS-3 dataset is used in this research. 1378 participants' pictures from various medical modalities, including MRI, PET, and CT scans, are gathered and stored in the OASIS dataset.

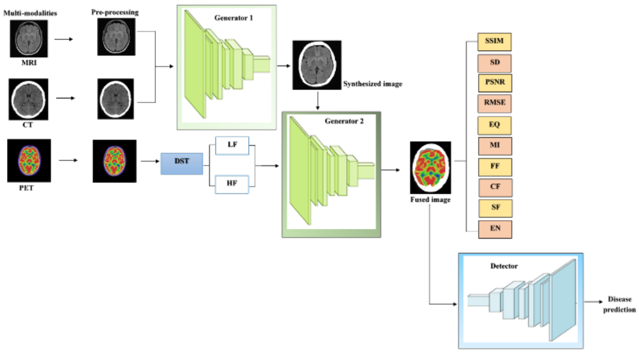

Fig. 1 – The schematic representation of the proposed model.

The MIDAS dataset, a rigid multimodality of brain scans, contains 110 images of patients' brains from various modalities under the retrospective image registration evaluation (RIRE) initiative. Table 1 illustrates the dataset description.

Table 1

Dataset description of OASIS and MIDAS datasets.

S. No	Datasets	Subjects	Multi-modality images			Total
			MRI	CT	PET	
1	OASIS dataset	1379	2842	1472	2157	7850
2	MIDAS dataset	110	1210	1320	990	3520
Total		1489	4052	2792	3147	11370

3.2 DATA PRE-PROCESSING

Pre-processing input data is crucial to reducing noise and improving subtle changes in medical images. Bilateral filters are used for noise reduction, non-linear filtering, leveling, and edge-preserving filtering. In noisy images, the bilateral filter fails to capture whole data about edge differences significantly, which is considered a major flaw. Therefore, the SCRAB filter is used to address this problem:

$$z_r(p,x,q) = \alpha\exp(-\frac{1}{2}\left(\frac{\|f(p)-f(x)-q(x)\|}{\sigma_r})^2\right) + \beta, \quad (1)$$

$$q(x) = \begin{cases} \left|f(x) - \text{mean}\left(\Omega_y\right)\right| |p-q| \leq c, \\ 0 \qquad\qquad \text{otherwise}, \end{cases} \quad (2)$$

where Ω_y describes the pixel set of $(2n+1)*(2n+1)$ pixel window here $n=2$, α and β are defined as positive parameters, Ω_y is the mean value, c denotes stable-variable and $q(x)$ is a range-based functions. The parameters utilized to control z_r are the scaling factor σ_r, the linear constant coefficient $\alpha = 2$ and $\beta = 1$. From these three, σ_r ensures the increased frequency of photometric similarity between a pixel's center (x) and its adjacent (p).

3.3 PET IMAGE DECOMPOSITION

The Shearlet transform produced the images into a low-frequency sub-band, which mostly mirrored the source images with dark qualities and layouts. The sub-band with low-frequency coefficients represents contour details, and it is liable for the local association between the input images of the adjacent coefficients:

$$L_{fre}(a,b) = \begin{cases} L_U(a,b), \ \text{if } E_{L_U}(a,b) > E_{L_V}(a,b), \\ \qquad \sum U,V(x+x',y+y'). \end{cases} \quad (3)$$

Equation (3) aids in determining the relationship between low-frequency coefficients. This will conserve the features from the input images to the maximum threshold. The variation indicates the general degree of scattering between darkly estimated pixels in an image, often known as the grey value.

$$H_{fre}(a,b) = \begin{cases} H_U(x,y), & \text{if } E_{L_U}(a,b) > E_{UV}(a,b), \\ H_m^V(x,y), & \text{otherwise.} \end{cases} \quad (4)$$

Equation (4) depicts the fusion process's high-frequency coefficients. In this case, the standard high-frequency coefficients are developed using the entire valuable feature in delivering a quality-fused image.

3.4 DGN-TBMF MODEL

A multimodal learning approach should utilize complementary data and relationships from multiple modalities to enhance learning performance. As a result, it is essential to record modality-specific data while harnessing the underlying relationships between the different modalities. A tri-modal fusion-based Dual Generator network is used to accomplish this goal, as shown in Fig. 3. The proposed DGN-TBMF network is shown in Fig. 2.

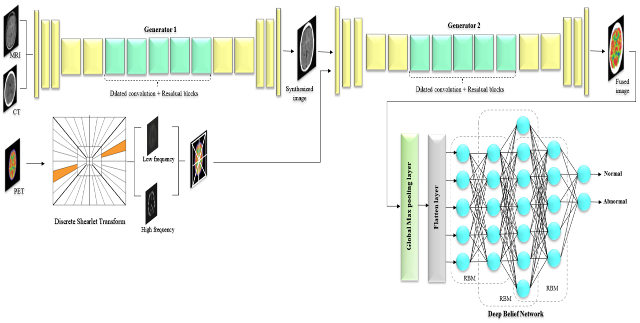

Fig. 2 – Architecture of the proposed DGN-TBMF network.

3.3.1 GENERATION PHASE

Generator-1 consists of a dilated convolutional layer for extracting the relevant grey matter densities and cortical thickness from MRI and CT images. The pre-processed MRI and CT images are merged by Generator-1 using common fusion strategies like element-wise maximization, element-wise addition, and element-wise multiplication. Generator-2 consists of a dilated convolutional layer for extracting the relevant voxel intensities from PET images. Generator networks are composed of three modules: Self-attention (SA) block, Residual feature selection (RFS) block, and tri-image fusion (TIF) block.

a) Self-attention block

The self-attention mechanism has been widely applied to solve this problem. The self-attention mechanism attains two features f_{mri} and f_{ct} are attained by 1×1 convolution. These two feature spaces form the query and key-value pairs. To estimate the attention map, both f_{mri} and f_{ct} are put into the fusion block. The resultant feature maps are up-sampled with nearest-neighbor interpolation. According to Fig. 3, the feature maps of the SA block are attained after passing through two more convolutional layers and up-sampling layers.

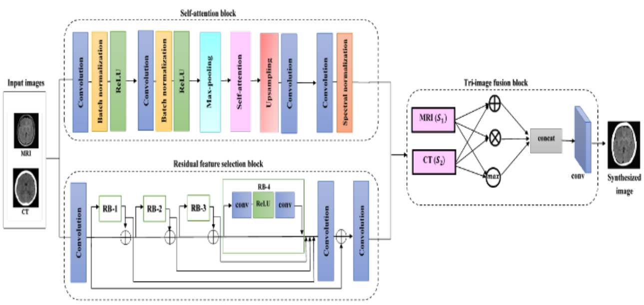

Fig. 3 – Structural representation of three blocks in Generator-1.

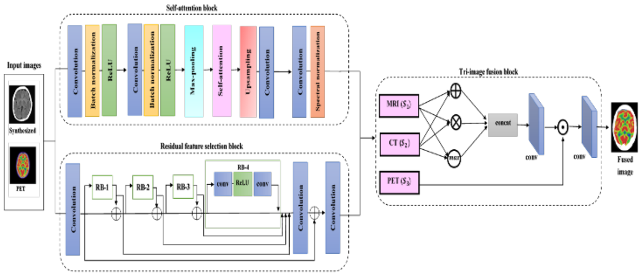

Fig. 4 – Structural representation of three blocks in Generator-2.

b) Residual feature selection block

The residual feature selection (RFS) block has been added to obtain the best local residual features. The RFS module contains four residual blocks. Before being transmitted directly to the end of the RFS module, the result of the last residual block is concatenated with the residual features of the first three blocks. A 1×1 convolution integrates these features before the element-wise addition with the identity feature. In contrast to stacking additional residual modules, the RFS framework allows non-local application of residual characteristics. In addition to conveying useful hierarchical information that preceding residual blocks contain without loss or interference, RFS modules result in a more discriminative feature representation.

c) Tri-image fusion block

The Tri-image fusion (TIF) block combines the characteristics collected by the self-attention block and the local detail block as input. Convolutional layers are present in three of the merge blocks. Spectral normalization is applied only to the top two layers rather than the third. In addition, the last layer uses a sigmoid activation function instead of a ReLU activation function to produce the final fused image. The TIF module is depicted in Fig. 4. In this case, an adaptive weight network is used to fuse feature depictions from three modalities. In addition to element-wise addition, element-wise multiplication, and element-wise maximization, the Hadamard product is another popular method for fusing multimodal data. As deed in Fig. 4, the output $S_i^{n-1} \in \mathbb{R}^{w*h*c}$ for the $(n-1)^{th}$, the ith layer modality is attained, where c denotes the number of feature channels, w and h denotes theeight and width of the feature maps, respectively. At that, the three fusion methods are employed to the inputs S_i^{n-1} $(i = 1, 2, 3)$ is used to acquire the synthesized output of MRI and CT images.

$$F_{syn} = \begin{cases} f_a = S_1^{n-1} \oplus S_2^{n-1} \\ f_s = S_1^{n-1} \otimes S_2^{n-1} \\ f_m = \max\left(S_1^{n-1}, S_2^{n-1}\right) \end{cases}, \quad (5)$$

where "$\oplus$", "$\otimes$" "max" and "$\odot$" represent element-wise addition, element-wise multiplication, element-wise maximization and hadamard product operations respectively. Then, $f_{concat} = \{f_a, f_s, f_m, f_h\}$ was combined, f_{concat} is given as an input to the first convolutional layer. The resultant of this layer is merged with the DST output f_{n-1} of the $(n-1)^{th}$ TIF block is given as an input to the second convolutional layer.

$$F_{fused} = f_{syn} \odot S_3^{n-1}. \quad (6)$$

Finally, the output F_{fused} of the n-th is obtained from the TIF module. Here, when $n = 1$, there is no previous output f_{n-1}, the output of the first convolutional layer feeds into the second convolutional layer. Remembering that the fusion strategy enables the TIF module to dynamically weight the various feature depictions from the four tri-modalities described in the fusion operations is essential.

3.3.2 DETECTION PHASE

The deep belief network (DBN) is utilized as the detection network for predicting abnormalities in the brain. DBNs are built using convolutional restricted Boltzmann machines (CRBMs). CRBM is intended to handle the problem of scaling techniques to multi-modality images.

The RBM comprises two layers: D_v and D_h are denoted as visible layers and hidden layers of DBN. Each channel of the visible layer is made-up of $ND_v * ND_h$ real-valued units. The hidden layers contain G groups, each with $ND_h * ND_h$ hidden unit. Moreover, the probabilistic max-pooling procedure decreases when the computational load allows full probabilistic inferences.

RBM associates with max-pooling features with probabilistic-based inference. The mathematical description of the probabilistic max-pooling energy function is derived as

$$E(D_h, D_v) = \frac{1}{2\sigma^2}\sum_{l=1}^{D}\sum_{i,j}(D^l_{v_{i,j}} - c_l)^2 - \sum_{g=1}^{G}\sum_{i,j}(D^g_{h_{i,j}}(\sum_{l=1}^{D}\frac{1}{\sigma}(\mathcal{A}^{g,l} * D^l_v)_{i,j} + bias_g). \quad (7)$$

Here $D_v \epsilon \mathbb{R}^{ND_v * ND_v * D}$ signifies the visible units, $D \epsilon \mathbb{R}^{ND_h * ND_h * G}$ denotes the hidden units, and σ shows the standard deviation factor about the Gaussian visible units. The visible units in the l^{th} channel share the bias c_l, whereas the hidden units in the g-th group share the $bias_g$. Moreover, the flipping of an array horizontally and vertically is defined as $\mathcal{A}$. The inputs are multiplied by the weight matrices, and the bias vectors are summed up to produce the fully connected layer results. It is derived as,

$$O(x) = AF(weight * x + bias), \quad (8)$$

where *x* designates the input of a fully connected layer, and *O(x)* signifies the resultant value of the network. The sigmoid layer interprets the values into predictions, and the classification layer categorizes the brain abnormalities accurately from the fused images.

4. **RESULTS AND DISCUSSION**

This work's experimental setup has been implemented using MATLAB 2019b, with the system requirements being an i7 processor and 8 GB of RAM. The investigations are carried out using MRI, CT, and PET images collected from publicly available datasets. Here, two datasets are used for evaluation, with each brain medical image in the dataset. In this result analysis, each dataset is separated into 75:25 for the training and testing.

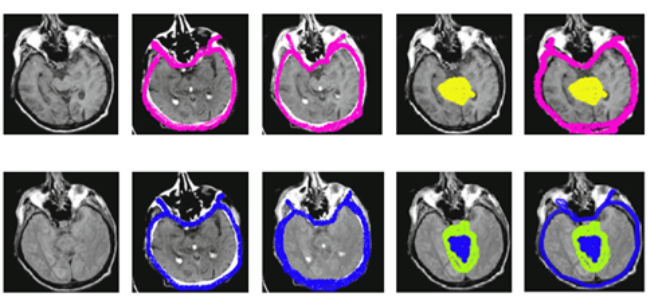

Fig. 5 – Results of the proposed DGN-TBMF model for Dataset-1.

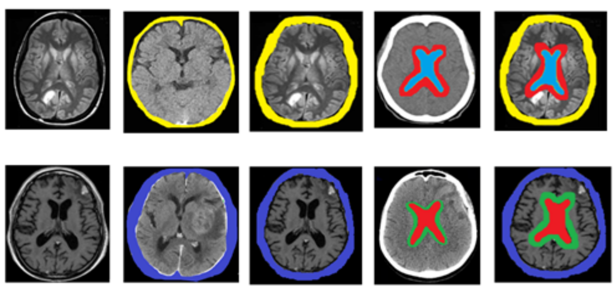

Fig. 6 – Results of the proposed DGN-TBMF model for Dataset-2.

Figure 5 displays the visualization results of the proposed DGN-TBMF model for dataset 1. Figure 6 displays the results of the proposed DGN-TBMF model for dataset 2 using the tri-modality images. The input images (column:1,2) are pre-processed using the SCRAB filter to eliminate the distortions. These pre-processed images are fused using Generator-1; this synthesized output is shown in column 3. Then, the PET images (column 4) are processed using DST to improve the image quality. These improved images are fused with the synthesized images in the Generator-2. The final tri-brain modal fused images are displayed in column 5. These fused output images incorporate different features from tri-modalities for the diagnosis of abnormalities.

4.1 PERFORMANCE ANALYSIS

This section portrays the effectiveness assessment of the proposed DGN-TBMF model, which was determined based on specificity, accuracy, recall, precision, and F1 score.

Table 1

Performance evaluation of the proposed DGN-TBMF model.

Datasets	Images	Accuracy	Specificity	Sensitivity	Recall	F1 score
Data set-1	MRI	98.99	98.02	97.96	98.10	98.92
	MRI+CT	99.09	98.25	98.12	98.79	99.01
	MRI+CT+PET	99.25	98.96	98.97	99.02	99.35
Data set-2	MRI	97.09	97.05	97.16	97.11	97.02
	MRI+CT	98.14	96.15	97.42	98.09	98.14
	MRI+CT+PET	99.04	97.16	96.27	98.12	97.35

Table 1 shows the sensitivity, specificity, precision, F1 score, and accuracy for datasets 1 and 2. The proposed network's accuracy is 99.25 % for dataset 1 and 99.04 % for dataset 2.

Table 2

Image fusion parameters for two datasets

Metrics	Dataset-1	Dataset-2
SD	98.78	94.25
EQ	0.96	0.92
MI	5.86	4.87
FF	6.54	5.88
EN	7.05	6.84
CF	0.95	0.92
SF	23.24	27.15

In Table 2, the tri-modal fusion results parameters are exposed in terms of various measurements such as standard deviation (SD), fusion factor (FF), entropy (EN), correlation factor (CF), edge quality (EQ), mutual information (MI), and spatial frequency (SF) with respective values. This work measures SF as an add-on value since it is a frequency-based measurement. Additionally, all measures are selected based on the most significant difference and data protection.

4.2 COMPARATIVE ANALYSIS

The efficiency of prior methods was evaluated to reveal that the DGN-TBMF net's finding is the most effective. The efficiency was calculated using the specific metrics for each type of brain disease. The accuracy rate attained through the proposed DGN-TBMF net is more efficient than that of state-of-the-art approaches. A comparative evaluation was made between the DGN-TBMF net and the three networks, as demonstrated in Table 2.

Table 3

Comparison of the proposed model with existing models using different performance metrics.

Table3(a): Standard deviation and edge quality.

Techniques	Dataset-1		Dataset-2	
	SD	EQ	SD	EQ
NSST [21]	90.24	0.84	80.44	0.82
ST-CNN [25]	80.35	0.54	66.21	0.83
Fuzzy-GOA [26]	94.08	0.73	56.12	0.75
DGN-TBMF (ours)	98.78	0.96	94.25	0.92

Table 3(b):Mutual information and fusion factor.

Techniques	Dataset-1		Dataset-2	
	MI	FF	MI	FF
NSST [21]	2.11	2.97	2.24	2.78
ST-CNN [25]	3.44	2.23	3.57	2.25
Fuzzy-GOA [26]	2.55	5.52	2.44	5.24
DGN-TBMF (ours)	5.86	6.54	4.87	5.88

Table 3 shows how the suggested model is evaluated in comparison to several cutting-edge models using two databases. Table 4(a) represents the SD and EQ of wavelet transforms and deep learning for image fusion. NSST [21], ST-CNN [25], and Fuzzy GOA [26] compared with the proposed network in dataset-1 while the deviations of SD of 8.54 %, 18.43 %, and 4.7 % and EQ of 0.12 %, 0.42 %, and 0.23 % respectively. The existing methods like [21, 25], and [26] compared with the proposed network in dataset-2 while the variations of SD of 13.81 %, 28.04 %, and 38.13% and EQ of 0.1 %, 0.09%, and 0.17%, respectively.

Table 3(b) shows the MI and FF of the state-of-the-art techniques. A robust DGN-TBMF model for low and high-frequency coefficients was proposed, resulting in better EQ and MI for the fused images. NSST [21], ST-CNN [25], and Fuzzy GOA [26] compared with the proposed network in dataset-1 although the deviations of MI of 3.75 %, 2.42 %, & 3.31%, and FF of 3.27 %, 4.31 %, & 0.72 % respectively. The existing methods like [21, 25], and [26] compared with the proposed model in dataset-2, whereas the variations of MI of 2.63 %, 1.30 %, & 2.43 %, and FF of 3.1 %, 3.63 %, & 0.64 %, respectively.

Table 3(c)

Entropy, correlation coefficient and spatial frequency.

Techniques	Dataset-1			Dataset-2		
	EN	CF	SF	EN	CF	SF
NSST [21]	5.18	0.46	23.34	5.31	0.34	17.73
ST-CNN [25]	5.53	0.74	7.82	6.81	0.82	8.82
Fuzzy-GOA [26]	6.47	0.85	16.42	6.23	0.87	19.08
DGN-TBMF (ours)	7.05	0.95	23.24	6.84	0.92	27.15

Table 3(c) shows the EN, CF, and SF of the state-of-the-art techniques. The tri-fused images generated by the proposed DGN-TBMF model effectively preserved the edges. This is very helpful for Human visual performance. NSST [21], ST-CNN [25], and Fuzzy GOA [26] compared with the proposed DGN-TBMF model in dataset-1 while the aberrations of EN of 1.87 %, 1.52 %, & 0.58 %; CF of 0.49 %, 0.21 %, & 0.1 %; and SF of 15.42 %, 2.9 % & 6.82 %. The existing methods like [21, 25], and [26] compared to the proposed network in dataset-2 while the deviations of EN of 1.53 %, 0.03 %, & 0.61%; CF of 0.58 %, 0.1%, & 0.05 %; and SF of 9.42 %, 18.33 %, & 8.07 %, respectively.

Table 4

Time complexity comparison of proposed and existing models.

Techniques	Time required (in minutes)	Accuracy
NSST [21]	186.4	98.54
ST-CNN [25]	191.5	98.07
Fuzzy-GOA [25]	192.6	98.21
DGN-TBMF (ours)	172.5	99.14

A comparison of different algorithms is presented in Table 4 for 190 MB images for brain tumor analysis. In terms of time required to train, the DGN-TBMF model requires the least amount of time. It requires roughly 30 minutes to train for 100 epochs. Despite its increased weight and training time, the proposed DGN-TBMF model has a higher accuracy score than the other models. The proposed DGN-TBMF model produces the most consistent output with the highest weight-to-accuracy ratio. Furthermore, it performed better than other approaches based on prediction accuracy, whereas the DGN-TBMF model is highly responsive to different brain input images.

5. **CONCLUSION AND FUTURE WORK**

This work presents a novel DGN-TBMF model for fusing medical tri-modality images to identify brain diseases in their early stages. The investigations are being conducted on a publicly accessible pre-enrolled dataset, which includes CT, MRI, and PET images. The pre-processed MRI and CT images are merged by the first generator using three popular fusion strategies, and the second generator is used to fuse the generated synthesized image and pre-processed PET images using four fusion strategies. The proposed DGN-TBMF model achieves high-performance levels based on the SD, EQ, MI, FF, EN, CF, and SF of 96.51, 0.79, 5.36, 6.21, 6.94, 0.93, and 25.19, respectively. The proposed DGN-TBMF uses a deep-learning model for image fusion and disease detection. Several aspects of this method make it faster and more effective than other methods. Through the proposed technique, multi-level decomposition fusion has significant implications for medical diagnosis. In future work, the proposed model can be applied in real-time to identify various diseases.

ACKNOWLEDGEMENT

The supervisor provided direction and constant support throughout this research, for which the author is grateful.

Received on 30 November 2023

REFERENCES

1. W. Yin, L. Li, F.X. Wu, *Deep learning for brain disorder diagnosis based on fMRI images*, Neurocomputing, **469**, pp. 332–345 (2022).
2. M. Hu, Y. Zhong, S. Xie, H. Lv, Z. Lv, *Fuzzy system based medical image processing for brain disease prediction*, Front. Neurosci., **15**, *714318* (2021).

3. X. Liang, P. Hu, L. Zhang, J. Sun, G. Yin, *MCFNet: Multi-layer concatenation fusion network for medical images fusion*, IEEE Sens. J., **19,** *16***,** pp. 7107–7119 (2019).
4. Y. Liu, X. Chen, J. Cheng, H. Peng, *A medical image fusion method based on convolutional neural networks*, 20th International Conference on Information Fusion (Fusion), pp. 1–7 (2017).
5. M. Talo, O. Yildirim, U.B. Baloglu, G. Aydin, U.R. Acharya, *Convolutional neural networks for multi-class brain disease detection using MRI images*, Comput. Med. Imaging Graphics, **78***, 101673* (2019).
6. H. Hermessi, O. Mourali, E. Zagrouba, *Convolutional neural network-based multimodal image fusion via similarity learning in the Shearlet domain*, Neural Comput. Appl., **30,** *7*, pp. 2029–2045 (2018).
7. J. Du, W. Li, K. Lu, B. Xiao, *An overview of multi-modal medical image fusion*, Neurocomputing, **215**, pp. 3–20 (2016).
8. R. Hou, D. Zhou, R. Nie, D. Liu, X. Ruan, *Brain CT and MRI medical image fusion using convolutional neural networks and a dual-channel spiking cortical model*, Med. Biol. Eng. Comput., **57,** *4***,** pp. 887–900 (2019).
9. Z. Ding, D. Zhou, R. Nie, R. Hou, Y. Liu, *Brain medical image fusion based on dual-branch CNNs in NSST domain*, BioMed Research Int., (2020).
10. R. Sundarasekar, A. Appathurai, *Efficient brain tumor detection and classification using magnetic resonance imaging*, Biomedical Physics & Engineering Express, **7,** *5***,** 055007 (2021).
11. S. Shanthi, S. Saradha, J.A. Smitha, N. Prasath, H. Anandakumar, *An efficient automatic brain tumor classification using optimized hybrid deep neural network*, International Journal of Intelligent Networks, **3**, pp.188–196 (2022).
12. H. Gupta, P. Kumar, S. Saurabh, S.K. Mishra, B. Appasani, A. Pati, C. Ravariu, A. Srinivasulu, *Category boosting machine learning algorithm for breast cancer prediction*, Revue Roumaine des Sciences Techniques – Serie Electrotechnique et Energetique **66**, *3*, pp. 201–206 (2021).
13. B. Bouchiba, I.K. Bousserhane, M.K. Fellah, A. Hazzab, *Artificial neural network sliding mode control for multi-machine web winding system*, Revue Roumaine des Sciences Techniques-Serie Electrotechnique et Energetique, **62**, *1*, pp. 109–113 (2017).
14. R. Sundarasekar, A. Appathurai, *Efficient brain tumor detection and classification using magnetic resonance imaging*. *Biomed. Phys. Eng. Express*, **7,** *5***,** 055007 (2021).
15. Y. Li, J. Zhao, Z. Lv, J. Li, *Medical image fusion method by deep learning*. Int. J. Cognit. Comput. Eng*,* **2**, 21–29 (2021).
16. N. Alseelawi, H.T. Hazim, H.T. Salim Al Rikabi, *A novel method of multimodal medical image fusion based on hybrid approach of NSCT and DTCWT*, International Journal of Online & Biomedical Engineering, **18,** *3* (2022).
17. M. Kaur, D. Singh, *Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks*, J. Ambient Intell. Hum. Comput.**12,** *2*, pp. 2483–2493 (2021).
18. Z. Wang, X. Li, H. Duan, Y. Su, X. Zhang, X. Guan, *Medical image fusion based on convolutional neural networks and non-subsampled contourlet transform*, Expert Syst. Appl., **171**, 114574 (2021).
19. 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, **67**, 102536 (2021).
20. Y. Zhang, Y. Liu, P. Sun, H. Yan, X. Zhao, L. Zhang, *IFCNN: A general image fusion framework based on convolutional neural network*. Inf. Fusion, **54**, pp. 99–118 (2020).
21. W. Tan, P. Tiwari, H.M. Pandey, C. Moreira, A.K. Jaiswal, *Multimodal medical image fusion algorithm in the era of big data*, Neural Comput. Appl., pp. 1–21 (2020).
22. K. Wang, M. Zheng, H. Wei, G. Qi, Y. Li, *Multi-modality medical image fusion using convolutional neural network and contrast pyramid*, *Sensors*, **20**, *8*, 2169 (2020)
23. K. J. Xia, H. S. Yin, J. Q. Wang, *A novel improved deep convolutional neural network model for medical image fusion*, Cluster Computing, **22,** *1*, pp.1515–1527 (2019).
24. B. Rajalingam, R. Priya, *Multimodal medical image fusion based on deep learning neural network for clinical treatment analysis*, International Journal of Chemtech Research, **11,** *06***,** pp. 160–176 (2018).
25. X. Xu, D. Shan, G. Wang, X. Jiang, *Multimodal medical image fusion using PCNN optimized by the QPSO algorithm*, Applied Soft Computing, **46***,* pp. 588–595 (2016).
26. A. K. Bhandari, K. Rahul, *A novel local contrast fusion-based fuzzy model for color image multilevel thresholding using grasshopper optimization*, Appl. Soft Comput. **81,** 105515 (2019).
27. S. Karthick, N. Muthukumaran, *Deep regression network for single-image super-resolution based on down- and upsampling with RCA blocks*, National Academy Science Letters, pp. 1-5 (2023).
28. S. Karthick, N. Muthukumaran, *Deep regression network for the single image super-resolution of multimedia text image*, IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA), pp. 394-399 (2023).
29. A. Jegatheesh, N. Kopperundevi, M. Anlin Sahaya Infant Tinu, *Brain aneurysm detection via firefly optimized spiking neural network*, International Journal of Current Bio-Medical Engineering, **01**, *01*, pp. 23–29 (2023).