MULTIMODAL MEDICAL IMAGE FUSION USING HYBRID DOMAINS

A.RAJESH NAIDU, D.BHAVANA *

Abstract. In a variety of clinical applications, image fusion is critical for merging data from multiple sources into a single, more understandable outcome. The use of medical image fusion technologies to assist the physician in executing combination procedures can be advantageous. The diagnostic process includes preoperative planning, intra operative supervision, an interventional treatment. In this thesis, a technique for image fusion was suggested that used a combination model of PCA and CNN. A method of real-time image fusion that employs pre-trained neural networks to synthesize a single image from several sources in real-time. A innovative technique for merging the images is created based on deep neural network feature maps and a convolution network. Picture fusion has become increasingly popular as a result of the large variety of capturing techniques available. The proposed design is implemented using deep learning technique. The accuracy of the proposed design is around 15% higher than the existing design. The proposed fusion algorithm is verified through a simulation experiment on different multimodality images. Experimental results are evaluated by the number of well-known performance evaluation metrics

Key words: Image fusion, PCA, CNN, VGG 16, VGG19

AMS subject classifications. 68U10

1. Introduction. Medical imaging may provide information in a number of ways. For clinical diagnosis, it is vital to keep an eye on the distinguishing aspects of different medical images. Image fusion is a method used to produce a single image by integrating the features of numerous image sensors. This information may come from a single source across many time periods or from a number of sensors during a single period. Fusion is the process of joining two or more things together to form a single entity[4],[10]. To improve the use of medical images and assist clinicians in interpreting image content, medical image fusion attempts to extract as much useful information as possible from source images[8]. Image fusion is the process of fusing two or more images into a series of images that incorporates the data from the individual images. The result is an image with higher information content than any of the input images. The purpose of image fusion is to create images that are more acceptable and understandable to both humans and machines.

MRI scanners employ magnetism, high magnetic fluctuations, and radio waves to generate images of the inside of the body's tissues[5]. An MRI scan that detects structural issues before they become irreparable may help to avoid nerve damage. Neuropsychological testing is often used to identify neural injury [18]. Medical image fusion was used to merge images from several imaging modalities, including computed tomography (CT), magnetic resonance imaging (MRI), and proton emission tomography (PET) [6].

To eliminate superfluous data, source register images are joined with other important data in an image fusion process. To do quantitative analysis, entropy and other statistical approaches are applied. Liguo Zhang has developed a complete deep learning model for extracting features, merging features, and restoring pictures that does not involve the development of complex feature matching and fusion criteria [1]. According to the Associated Press, Jamesa provides a comprehensive list of possibilities as well as an explanation of the several scientific obstacles that medical image fusion faces [2].

Deep learning has been used to address some of the most difficult challenges in medical imaging. Simultaneous PET/MR imaging combines the soft-tissue contrast of MR imaging with the molecular sensitivity and specificity of PET. Obtaining a valid photon attenuation correction (AC) map, which is critical for accurate PET quantification, is still a difficulty [3] [19-22]. Because traditional MR imaging-based AC (MRAC) techniques lack direct bone measurement, a number of new methodologies have been used.[15] PET and SPECT imaging

---

*Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Guntur District, India (`bhavanaece@kluniversity.in`)

applications of AI-based algorithms range from low-level electronic signal formation and processing to high-level internal dosimetry and diagnostic/prognostic modelling. Deep learning techniques have mostly been employed to improve incident photon timing resolution and localization accuracy with the objective of enhancing overall spatial and time-of-flight (TOF) resolutions in PET. The purpose of the research being done on quantitative SPECT and PET imaging right now is to get rid of the effects of noise, artifacts, and movement [16].

A wide variety of computer vision problems, including identification, fragmentation, and very low-level recognition, have been improved through deep learning. It's called image fusion when two or more separate images are combined to create a new image. To enhance the original multi-spectral image's resolution, it takes data from several sources while still conserving spectral information. Radiology uses computed tomography, or CT, to produce detailed images of the inside of the human body without the need for any incisions [9]. It's impossible to view your bones, muscles, and other biological parts in the same way that this image does. Tumors may be identified on CT scans by their shape, size, and location.

Utilize contourlet transforms and adaptive weighted PCA in the frequency domain to create a multimodal fusion. Describe the fusion mechanism for the DTCWT and SOFM modalities. Describe the use of NSCT-based image fusion for clinical research. For the purpose of improving input photos, the hybrid fusion method is introduced. Review and elucidate the fusion algorithms overview. developed a hybrid approach for accurate treatment analysis based on a blend of curvelet and wavelet transforms. To combine structural and functional information from the source images, introduce the fusion method. A-PCNN has been developed, and NSST has been suggested for multimodality medical pictures [8]. Describe the NSCT for multimodality medical image fusion. a curvelet-based fusing algorithm was invented [10]. An algorithm for fusing multimodal medical images has been created [11]. The DFRWT technique is described here for effective image fusion [12]. Create and modify the surface of the fusion system [13]. Describe the fusion technique for medical pictures based on SWT and NSCT [14].

2. Proposed Methodology. In this paper, the proposed method is a combination of CNN and PCA. Radiological images were collected, followed by normalization, which enhances both computation efficiency and execution. The normalised images were then fed into the CNN algorithm for fusion, and the fused images were then gathered and submitted to the PCA method for dimensionality reduction. the whole process of proposed model is shown in figure 2.1. As a result, the technique is now known as a CNN + PCA combine model, which produces superior results than other techniques. The normalized image is submitted to VGG16, VGG19, SqueezeNet, and PCA to verify the accuracy of alternative algorithms for image fusion, although the results are less accurate than the combination model of CNN+PCA. For image fusion, several models of radiological imaging such as MRI, CT, PET, and SPECT were collected. Two different format photographs are used in this image fusion.

2.1. Convolutional Neural Network. Convolutional networks, also known as convolutional neural networks, are a type of neural network that uses a grid-like design to process input (LeCun, 1989). In real-world applications, convolutional networks have had a lot of success. A "convolutional neural network" is a network that uses the convolutional mathematical approach as shown in figure 2.2. It's a very specialized linear procedure. Simple neural networks with a minimum of one layer that simply use convolution instead of generic matrix multiplication are used in these networks. Convolutional networks were among the first neural networks to solve key business problems, and they continue to be at the core of commercial deep learning applications today.[14] A convolutional neural network consists of several layers which are classified as the Input layer, Convolutional layer, Max Pooling layer, flatten layer, Output layer which helps us to extract features for the fusion of images. A tiny portion of the visual field reacts to stimuli in the Receptive Field which are known as neurons. A ring of comparable fields forms around the entire field of vision.

2.2. VGG16. It is a deep Convolutional Neural (CNN) architecture with several layers. VGG means Visual Geometry Group.VGG-16 is a deep convolutional neural network with 16 layers. It has a very appealing architecture due to its uniformity. It features only 3x3 convolutions, but a lot of filters, similar to AlexNet [19]. It may be taught for two to three weeks on 4 GPUs. It is now the most popular method for extracting characteristics from images in the communityIn ImageNet, The top-5 precision of the VGG16 model was 92.7 percent. The VGG16 Architecture was invented and presented by Karen Simonyan with Andrew Zisserman of
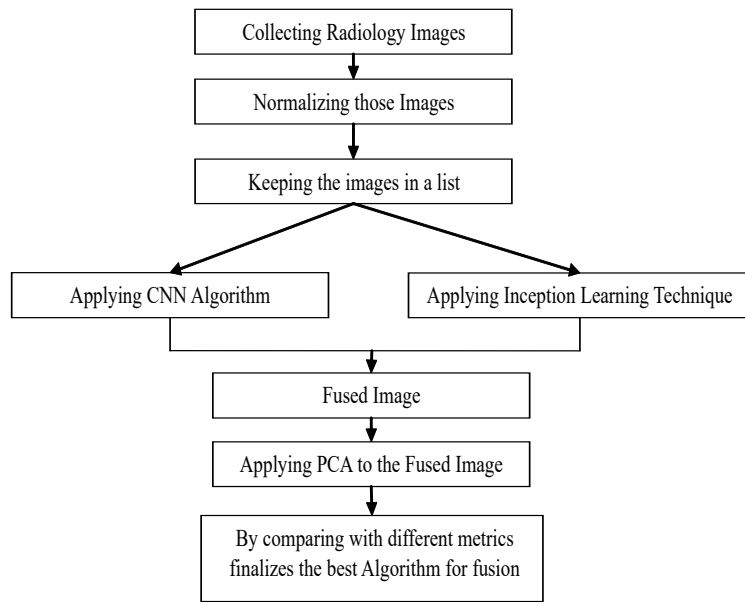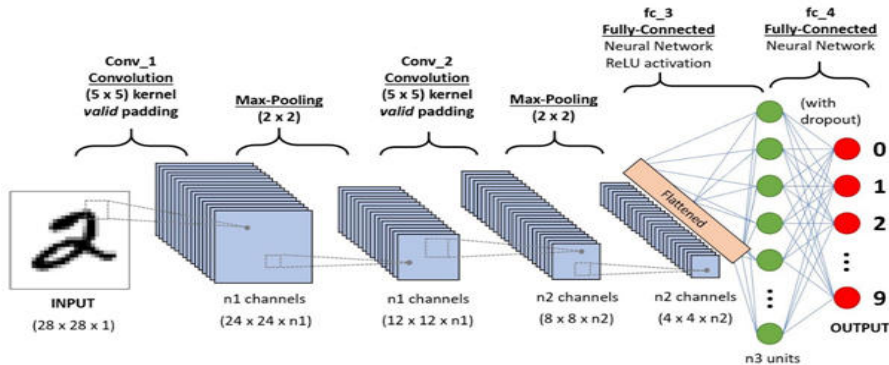
Fig. 2.1: Proposed Model

Fig. 2.2: CNN sequence to classify handwritten digits

Oxford University in their article "Very Fully Convolutional Networks for Large Object Recognition" in 2014.

2.3. VGG19. The VGG19 model is a VGG version with nineteen layers which is a combination of sixteen convolution layers, three fully linked layers, five max pool layers, and one Soft Max layer. The VGG19 (also known as VGGNet-19) is a model similar to VGG16 only the difference is it contains 19 layers [11,13]. The weight layers of the model are represented by the digits "16" and "19". In comparison to VGG16, VGG19 has three more convolution layers. VGG19 is a complex CNN having pre-trained phases and a deep understanding of how a picture is described in terms of appearance, color, and structure. VGG19 is a neural network that was learned on millions of images to solve difficult classification issues. For a wide range of images, the system has learned rich feature representations.
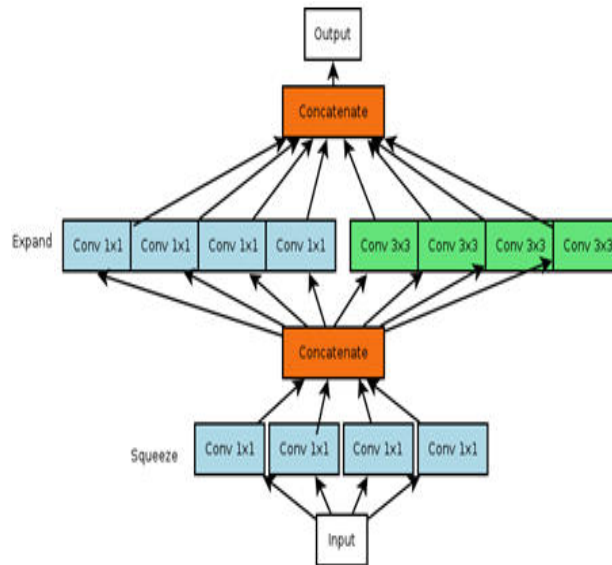
Fig. 2.3: CNN sequence to classify handwritten digits

2.4. SqueezeNet. SqueezeNet is a deep convolutional neural network with eighteen layers. SqueezeNet is a form of neural network that uses design tactics to minimise dimensionality, most notably the usage of flame components, which use 1x1 CNNs to "squeeze" variables. SqueezeNet was created with the goal of creating a compact neural network with fewer elements that could fit into memory storage and be conveyed more easily over a computer network. SqueezeNet is a small network designed to be a more compact version of AlexNet. It has more than 50 times less parameters than AlexNet yet is three times faster [12]. The SqueezeNet design is made up of the "squeeze" and "expand" layers. In a squeeze convolutional layer, only one filter is used. In an extended layer, these are routed via a mix of 1 1 and 3 3 convolution filters. SqueezeNet was trained with a learning method of 0.04, which lowers linearly over time. SqueezeNet makes the setup process easier because of its small size. This network was created in Caffe, but it has since gained popularity and has been adopted by a number of platforms.

2.5. Principal Component Analysis. Principal Component Analysis is an unsupervised learning approach for reducing dimensionality in machine learning. It is a mathematical method that uses orthogonal information to turn correlated attribute outputs into a collection of uncorrelated, linear qualities. The Principal Components are the newly amended features.

It is necessary to limit the number of variables and interpret linear combinations of data in order to understand data with a large number of variables in a meaningful way. PCA is a method for reducing a large number of potentially associated variables to a smaller number of variables known as principal components [7]. A method for condensing a large number of samples into a smaller set that preserves the majority of the data from the larger set is principal component analysis. We can utilize the principal component analysis technique to create and use a smaller set of variables called primary factors. It's much easier to research and interpret a smaller set of data.

3. Performance Evaluation. The source images are obtained using several radiological techniques from a normal brain. The fuse output is formed again by integrating its inverted output of the resultant connection with actual U and V segments, using real brain MRI and PET input images as a source for testing. For source images, the process is simple to implement; nevertheless, the fusion result is not flawless. The fusion image is unstable because the information in the fused image changes from one algorithm method to the next, while the data image has many modalities depending on the algorithm utilized. The hue of the PCA image changes when

compared to other fused images. Because we are using PCA, the generated fused image will be in grayscale mode. In this section, we will discuss the comparative analysis of different algorithm methods using various metric indexes. The comparative analysis is based on source images and fused image with help of comparing algorithm methods including the add-weighted average of them. The fused images were created utilizing several methods such as VGG-16, VGG-19, Squeeze Net, PCA, and CNN+PCA.

The following metrics can be used to assess the performance of image fusion algorithms:

Mean Square Error [17] (MSE):

$$MSE = \frac{1}{X * Y} \sum_{m=0}^{X} \sum_{n=0}^{Y} (S_1(m,n) - S_f(m,n))^2 \tag{3.1}$$

where $X * Y$ image size. $S_1(m,n)$ is the input image and the fused image.

Structural Similarity Index Measure [17 (SSIM)]

$$SSIM = \frac{(2\mu_{S_r}\mu_{S_f} + K_1)(\sigma_{S_r}\sigma_{S_f} + K_2)}{(\mu_{S_r}^2 + \mu_{S_f}^2 + K_1)(\sigma_{S_r}^2 + \sigma_{S_f}^2 + K_2)}) \tag{3.2}$$

where $\mu(S_r)^2 + \mu_(S_f)^2$ is close to 0, the constant K_1 is used to prevent instability and when $\sigma_(S_r)^2 + \sigma_(S_f)^2$ is close to 0, the constant K_2 is used to prevent instability

$\mu_{S_r)} = \frac{1}{XY} \sum_{m=0}^{X} \sum_{n=0}^{Y} S_r(m,n)$

$\mu_{S_r)} = \frac{1}{XY} \sum_{m=0}^{X} \sum_{n=0}^{Y} S_f(m,n)$

$\sigma_{S_r}^2 = \frac{1}{(XY-1)} \sum_{m=1}^{X} \sum_{n=0}^{Y} (S_r(m,n) - \mu_{(S_r)})^2$

$\sigma_{S_f}^2 = \frac{1}{(XY-1)} \sum_{m=1}^{X} \sum_{n=0}^{Y} (S_f(m,n) - \mu_{(S_f)})^2$

$\sigma_{S_r S_f} = \frac{1}{(XY-1)} \sum_{m=0}^{X} \sum_{n=0}^{Y} (S_r(m,n) - \mu_{(S_r)})(S_f(m,n) - \mu_{(S_f)})$

Visual Information Fidelity [17] (VIF): It is the parameter for evaluating image quality that is based on information aspects of natural scenes statistical (NSS).
$A(p_{i,k}, q_{i,k})$= actual image details
$A(p_{i,k}, r_{i,k})$ = unbalanced image details
$p_{i,k}, q_{i,k} and r_{i,k} = p, q$, and r in images frames in a variety of sizes
M_k= scale's amount of image pixels k
s = the variety of categories in which the image is split

Mutual Information [17] (MI) Mutual information is used to measure the quality of the fused image, it is given as:

$$MI = I(S_X; S_f) + I(S_Y; S_f) \tag{3.3}$$

where $(S_X; S_f) = \sum_{m=1}^{M} \sum_{n=1}^{M} h_{R,F}(m,n) \log_2 \frac{(h_{R,F}(m,n))}{(h_R(m)h_F(n))}$, where $h_R(m), h_F(n)$, are S_R and normalised grey level histograms, which correspond to the reference and fused images, respectively. When the MI value of a fused image is higher, it means the image has more features and texture information.

Entropy The expected value along with the random variable is measured using entropy [17]. The image's color histogram can be thought of as a probability density function. The detailed entropy of a picture is represented with $h(i)$ where $h(i)$ reflects the quantized color's % of pixel i in the overall image.

$$Entropy = \sum_{i=1}^{n} h(i) \log_2 h(i) \tag{3.4}$$

4. Results. The comparative analysis of different algorithm methods using various metric indexes. The obtained results indicates the proposed algorithm is the best algorithm method among the tested algorithms. How to accurately evaluate the efficiency of fusion methods is an essential topic in picture fusion. In recent years, several other assessment indexes have been proposed. There are two types of assessment methods now in use,
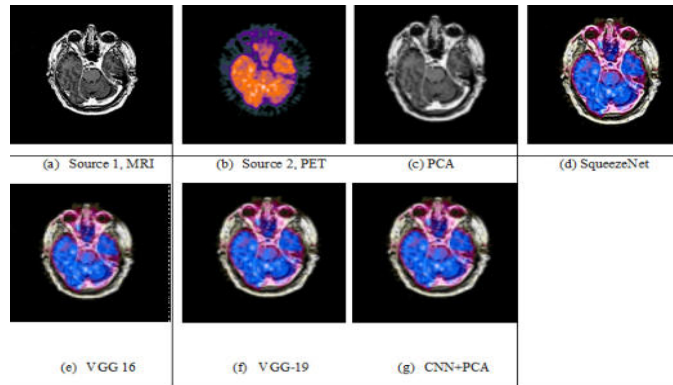
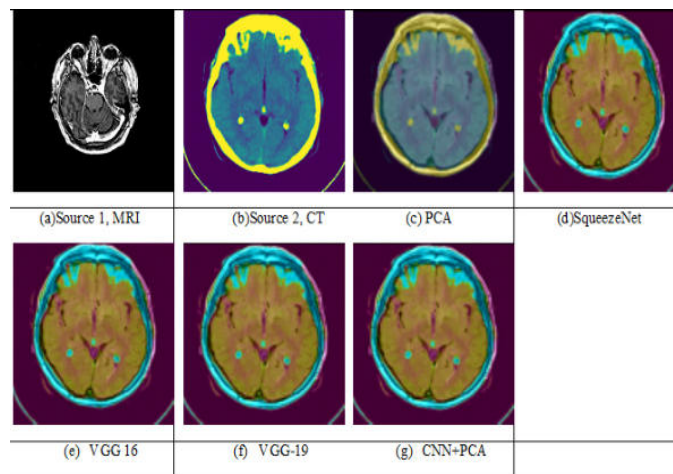Fig. 4.1: Normal brain MRI,PET medical images fusion results

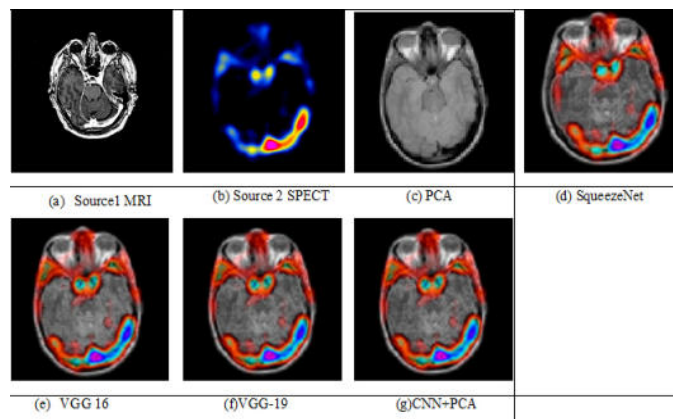Fig. 4.2: Normal brain MRI and CT medical images fusion results

Fig. 4.3: Normal brain MRI and SPECT medical images fusion results

Table 4.1: MRI-PET Metric Score [19]

	VGG-16	VGG-19	SQUEEZE NET	PCA	CNN+PCA
MSE	38.81896868	43.26268603	52.48530215	38.41147947	37.95749468
SSIM	0.634070678	0.599964187	0.576918209	0.593477189	0.631366731
VIF	0.020115451	0.026862346	0.017607955	0.022060916	0.082619463
Entropy	3.770389411	4.070330749	4.112055998	3.990689894	3.001083163
MI	0.436537877	0.435251605	0.371655424	0.467336177	1.145331026

Table 4.2: MRI-CT Metric Score

	VGG-16	VGG-19	SQUEEZE NET	PCA	CNN+PCA
MSE	52.56273155	46.16454714	53.09027707	22.221134702	20.85981113
SSIM	0.338170304	0.373294043	0.353296692	0.428951084	0.060058943
VIF	0.03505622	0.053515519	0.045902515	0.078706307	0.015024352
Entropy	6.420616688	6.279529215	6.173356001	5.989346358	4.957725674
MI	0.658167886	0.724539243	0.718558411	0.878390975	1.265098822

Table 4.3: MRI-SPECT Metric Score

	VGG-16	VGG-19	SQUEEZE NET	PCA	CNN+PCA
MSE	42.7168407	42.27611837	45.56879791	57.79010585	34.55778519
SSIM	0.486359572	0.473912872	0.470555961	0.591496644	0.48038565
VIF	0.047101677	0.044480608	0.037290363	0.107424537	0.048830219
Entropy	4.729863707	4.815029541	4.693647639	4.380235834	3.202091852
MI	0.362361713	0.324649206	0.308850748	0.586975232	1.573748002

Table 4.4: MRT1- MRT2 Metric Score

	VGG-16	VGG-19	SQUEEZE NET	PCA	CNN+PCA
MSE	51.78371531	54.17560435	30.95919525	47.35199752	26.68707336
SSIM	0.17952148	0.167316343	0.478195945	0.515009571	0.4094382486
VIF	0.031350142	0.024646476	0.062285785	0.075168443	0.0615285748
Entropy	5.695356778	5.763374805	5.290079636	5.277351219	3.9653502677
MI	0.493445079	0.459940614	0.531062598	0.626701992	1.119491017

the quality evaluation using a source image and a non-source image. Evaluation of image quality, we choose five metrics, for the accurate understanding of their outcomes Mean Square Error (MSE) be a part of the objective evaluation index that must be consulted the Structural Similarity Index (SSIM), Information Entropy (EN), Normalized Mutual Information (Qmi) and visual information fidelity (VIF) pixel, all of which are objective evaluation indicators. a few pairs of radiology medical images taken as the test data. For test data, applied filters for better understanding. The comparing algorithms are based on VGG-16, VGG-19, SQUEEZENET, Principal Component Analysis (PCA), and CNN+PCA. Among them, we randomly choose two source images with diverse modalities. The input image is transformed from RGB to HSI space at the time of testing, and various another format like grayscale, YCbCr etc.

5. Conclusion. The method extracts further information from source images using weights that are not shared. CNN is used with PCA in the proposed method for image reconstruction, and it also aids during

lowering the computational complexity and dimensionality of the original images. MSE loss is the basis for the proposed method loss function. Minimizing MSE loss raises the objective assessment index, however, this does not guarantee acceptable image fidelity. Its focus on the future discussion will be refining the loss function as well as enhancing its effectiveness images that were fused to get a bunch more durable fusion method according to CNN. The proposed model's fusion image can effectively fully preserve the useful information in the source images possible. The proposed method outperforms other comparison methods in terms of evaluating image fusion metrics. The proposed method is stable in terms of durability and may be applied to image fusion.

REFERENCES

[1] X. Liang, P. Hu, L. Zhang, J. Sun and G. Yin, "MCFNet: Multi-Layer Concatenation Fusion Network for Medical Images Fusion," in IEEE Sensors Journal, vol. 19, no. 16, pp. 7107-7119, 15 Aug.15, 2019

[2] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art", Inf. Fusion, vol. 19, pp. 4-19, Sep. 2014.

[3] Rajalingam, B., and Priya, R. (2018). Multimodal medical image fusion based on deep learning neural network for clinical treatment analysis. International Journal of ChemTech Research, 11(06), 160-176.

[4] E. Daniel, "Optimum Wavelet-Based Homomorphic Medical Image Fusion Using Hybrid Genetic–Grey Wolf Optimization Algorithm," in IEEE Sensors Journal, vol. 18, no. 16, pp. 6804-6811, 15 Aug.15, 2018, doi: 10.1109/JSEN.2018.2822712.

[5] R. Zhu, X. Li, X. Zhang, and M. Ma, "MRI and CT Medical Image Fusion Based on Synchronized-Anisotropic Diffusion Model," in IEEE Access, vol. 8, pp. 91336-91350, 2020.

[6] R. Nie, J. Cao, D. Zhou, and W. Qian, "Multi-Source Information Exchange Encoding With PCNN for Medical Image Fusion," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 3, pp. 986-1000, March 2021

[7] He, C., Liu, Q., Li, H., and Wang, H. (2010). Multimodal medical image fusion based on IHS and PCA. Procedia Engineering, 7, 280-285.

[8] Hermessi, H., Mourali, O., and Zagrouba, E. (2021). Multimodal medical image fusion review: Theoretical background and recent advances. Signal Processing, 183, 108036.

[9] Tan, W., Tiwari, P., Pandey, H. M., Moreira, C., and Jaiswal, A. K. (2020). Multimodal medical image fusion algorithm in the era of big data. Neural Computing and Applications.

[10] Bai L, Xu C, Wang C (2015) A review of fusion methods of multi-spectral image. Optik-Int J Light Electron Optics 126(24):4804–4807

[11] Thepade, S. D., and Dindorkar, M. R. (2022). Fusing deep convolutional neural network features with Thepade's SBTC for land usage identification. Engineering Science and Technology an International Journal, 27(101014), 101014.

[12] Ucar, F., and Korkmaz, D. (2020). COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. Medical Hypotheses, 140(109761), 109761. doi:10.1016/j.mehy.2020.109761

[13] Victor Ikechukwu, A., Murali, S., Deepu, R., and Shivamurthy, R. C. (2021). ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images. Global Transitions Proceedings, 2(2), 375–381. doi:10.1016/j.gltp.2021.08.027

[14] Gai, D., Shen, X., Chen, H., and Su, P. (2020). Multi-focus image fusion method based on two stage of convolutional neural network. Signal Processing, 176(107681), 107681.

[15] M. Yin, X. Liu, Y. Liu, and X. Chen, "Medical image fusion with parameter-adaptive pulse coupled neural network in nonsub sampled shearlet transform domain," IEEE Trans. Instrum. Meas., vol. 68, no. 1, pp. 4964, Jan. 2019.

[16] Venkatesan, B., and Ragupathy, U. S. (2022). Integrated fusion framework using hybrid domain and deep neural network for multimodal medical images. Multidimensional Systems and Signal Processing. doi: 10.1007/s11045-021-00813-9

[17] Naidu, V. P. S., M. Divya, and P. Mahalakshmi. "Multi-modal medical image fusion using multi-resolution discrete sine transform."Control and Data Fusion e-Journal 1, no. 2 (2017):13-26.

[18] B, R., and R, P. (2018). Multimodal medical image fusion using various hybrid fusion techniques for clinical treatment analysis. Smart Construction Research. doi: 10.18063/scr.v0.594

[19] Rajalingam, B., Priya, R., and Bhavani, R. (2019). Hybrid multimodal medical image fusion using combination of transform techniques for disease analysis. Procedia Computer Science, 152, 150–157. doi: 10.1016/j.procs.2019.05.037

[20] Yadav, Satya Prakash, and Sachin Yadav. "Image fusion using hybrid methods in multimodality medical images." Medical and Biological Engineering and Computing 58.4 (2020): 669-687.

[21] Alseelawi, Nawar, Hussein Tuama Hazim, and Haider TH Salim ALRikabi. "A Novel Method of Multimodal Medical Image Fusion Based on Hybrid Approach of NSCT and DTCWT." International Journal of Online and Biomedical Engineering 18.3 (2022).

[22] Singh, Sneha, and Radhey Shyam Anand. "Multimodal medical image fusion using hybrid layer decomposition with CNN-based feature mapping and structural clustering." IEEE Transactions on Instrumentation and Measurement 69.6 (2019): 3855-3865.