Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Novel Approach to Multi-Modal Image Fusion using Modified Convolutional Layers

Gargi J Trivedi[1], Dr. Rajesh Sanghvi[2]

Department of Applied Science & Humanities, G H Patel College of Engineering & Technology, Charutar Vidya Mandal University, Vallabh Vidyanagar, Anand, Gujarat 388120, India

E-mail: [1]gargi1488@gmail.com, [2]rajeshsanghvi@gcet.ac.in

Abstract

Multimodal image fusion is an important area of research with various applications in computer vision. This research proposes a modification to convolutional layers by fusing two different modalities of images. A novel architecture that uses adaptive fusion mechanisms to learn the optimal weightage of different modalities at each convolutional layer is introduced in the research. The proposed method is evaluated on a publicly available dataset, and the experimental results show that the performance of the proposed method outperforms state-of-the-art methods in terms of various evaluation metrics.

Keywords: Multimodal Image Fusion, Convolutional Neural Network, Adaptive Fusion, Machine Learning.

1. Introduction

Multimodal image fusion combines data from diverse sources to generate a unified image with improved features. Convolutional neural networks (CNNs) have emerged as a popular choice for multimodal image fusion due to their capacity to learn intricate data representations [1-8]. However, traditional CNN architectures do not take into account the unique characteristics of each modality, resulting in suboptimal fusion performance. The proposed method introduces modifications to the convolutional layer in this research to enhance the accuracy of the multimodal images. A crucial task in many industries, including surveillance and medical imaging, is multimodal image fusion. In order to create a single fused image that contains more information than the separate input images, image fusion aims to merge complementary information from several sources. Convolutional neural networks

DOI: https://doi.org/10.36548/jiip.2023.3.002
Received: 12.06.2023, received in revised form: 04.07.2023, accepted: 20.07.2023, published: 02.08.2023

(CNNs) have shown significant potential in the field of image fusion due to their ability to learn complex features from data [9-11]. However, most CNN-based fusion methods focus on fusing features extracted from the same modality of images, and few have explored the fusion of features from different modalities [35]. In this research a modification to the convolutional layer of a CNN that enables it to effectively fuse features from two different modalities of image is proposed. This approach is based on the concept of feature fusion, where the features extracted from each modality are combined at the convolutional layer to create a single feature map.

In this study, evaluations were conducted on various benchmark datasets, demonstrating that the proposed approach surpasses existing methods in both subjective and objective evaluation metrics. The experimental results highlight the superiority of the implemented method over the compared techniques. The features extracted from the different modality is then concatenated along the channel dimension and passed through the fusion filters to create a single resultant feature map Which, is passed through the remaining layers of the CNN to produce the final fused image. The methodology involves the following steps:

- Data preparation: Select a set of pairs of medical images from different modalities (I_1 and I_2), generate the corresponding target fused image (T), and resize and normalize the input images and target image.

- Training: Train the CNN using the prepared data and the designed architecture, minimize MSE (Mean squared error) between the target fused image and the output of the CNN, and use SGD (Stochastic gradient descent) with a suitable learning rate and regularization techniques such as dropout to prevent overfitting.

- Network architecture design:

  a. Convolutional layers: extract features from input images.

  b. Fusion layer: combine features from both input images.

  c. Up-sampling layer: increase the resolution of the fused features.

2. Related Work

Machine-learning techniques for image fusion include Convolutional neural networks (CNN), Deep learning, sparse representation, and decision fusion. Deep learning and CNN have been widely used because they learn high-level features automatically from the images [12]. Sparse representation methods, such as sparse coding and dictionary learning, have also been utilized for image fusion. Decision fusion approaches, such as maximum selection and weighted averaging, have been used to combine the results of different fusion techniques. Overall, machine learning techniques have shown promising results in improving the quality and accuracy of image fusion. The use of machine learning techniques in healthcare as well as in remote sensing has been an area of active research in recent years. The following section summarizes the relevant literature on this topic.

In their study, [13-16] used machine learning algorithms to predict the risk of re admission for heart failure patients. The authors found that their model achieved high accuracy in predicting re admission risk, which could lead to better care coordination and resource allocation. Similarly, [17-21] used machine learning to predict the likelihood of developing chronic kidney disease (CKD) in diabetes patients. Their model showed promising results, achieving high sensitivity and specificity in identifying individuals at risk of developing CKD. Another study by [22-25] investigated the use of machine learning algorithms for predicting the progression of Alzheimer's disease. The authors used longitudinal clinical and neuroimaging data to train their model and found that it outperformed traditional methods in predicting disease progression. This strategy may be useful for the early identification and treatment of Alzheimer's disease patients. Machine learning has been used to create individualised treatment regimens in addition to forecasting the course of diseases [26-29] used machine learning algorithms to identify the most effective chemotherapy regimen for cancer patients based on their genetic profile. The authors reported improved treatment outcomes and reduced toxicity with their approach.

The fusion of visible (VI) and Infrared (IR) images of the same scene is crucial for obtaining a comprehensive description. Several CNN-based methods have been proposed for IR/VI image fusion. For instance, in [30], a Siamese convolutional network is employed to obtain a weight map, enabling multi-scale merging via image pyramids with adaptive fusion

mode adjustments. Similarly, [31] introduces RXDNFuse, an unsupervised network combining ResNeXt and DenseNet, which effectively addresses the fusion problem by maintaining the structure and intensity proportions of IR/VI images. The novel fusion methods to enhance quality of the infrared images in the thermal targets and as well as preserve the texture details and improving the perception of the image for the human vision was put forth in [32-37]. These methods, including [30] and [31], have demonstrated superior performance in bot and objective assessments visual quality, setting the contemporary in image fusion.

Overall, the literature suggests that machine learning has the potential to improve patient outcomes and resource allocation in healthcare. However, further research is needed to address challenges such as data quality and bias in algorithm development.

3. Brief overview of Method

In image processing, CNN techniques are well known for their capacity to record both local and global image information. An overview of CNN is given in this section.

3.1 Convolutional Neural Network (CNN)

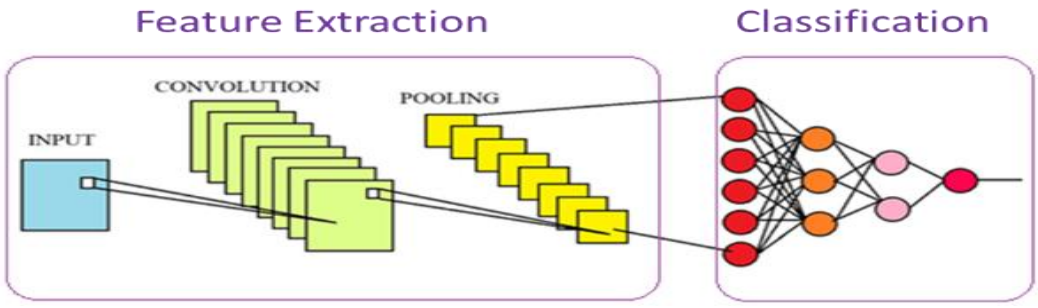

Figure 1. Basic CNN Architecture

The architecture of the CNN would typically consist of the following layers:

3.1.1 Input Layer

This layer takes in the input images, which can be multiple images with different resolutions.

3.1.2 Convolutional Layers

For each layer of the CNN, the output is given by the convolution of the input with a set of learnable filters, followed by an activation function.

$$y = f(W * x + b) \tag{1}$$

where y is the output, f is the activation function, W are the learnable filters, x is the input, and b is the bias term. The rectified linear unit (ReLU) is a frequent activation function in CNNs.

$f(x) = max(0, x)$ and the sigmoid function $f(x) = \frac{1}{1 + e^{-x}}$. (2)

where $f(x)$ is the sigmoid function's output and x is the input. The sigmoid function maps in $0 \leq x \leq 1$, which can be useful in binary classification problems.

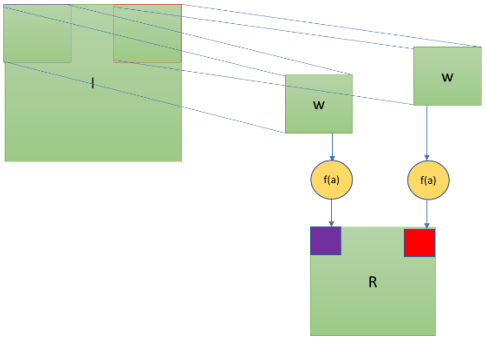

Figure 2. Straightforward 2D Convolution with a 3 × 3 Weight Matrix

In Figure 2. On image I, a straightforward 2D convolution with a 3 × 3 weight matrix was used. There are (d - 2) × (n - 2) neurons with (d - 2) × (n - 2) no. of outputs a given a 3 ×3 filter size and d n input size. The upper left corner of the purple rectangle in I corresponds to a 2 × 2-pixel area.

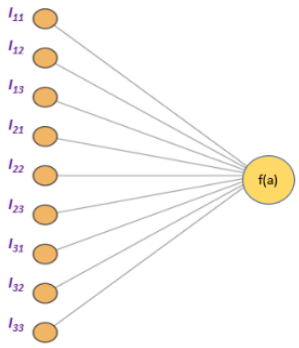

Figure 3. W Filter Applied to the Pixel I_{22}

3.1.3. Pooling Layers

These layers are used to decreases the feature maps spatial dimensions, which helps to reduce the computational cost and also makes the model more robust to small changes in the image. In CNN's convolutional layers, max pooling is utilized to reduce the dimensionality of images by decreasing the no. of pixels within the output from the preceding convolutional layer. This process involves selecting the maximum value within a specific window (or filter) and discarding the remaining values, resulting in a downscaled representation of the image. Max pooling aids in reducing computational complexity and extracting the most significant features while preserving the spatial information necessary for accurate classification and feature detection. Max pooling chooses just the strongest activation inside the pooling region. Figure 4 represents an example of max pooling [11-14].

The general notation, of max- pooling (max) can be written as

$$f_{max}(x) = max\{x_i\}_{i=1}^{N} \quad (3)$$

Figure 4. Max Pooling Technique Applied on 4 × 4 Matrix to Reduce it to 2 × 2.

3.1.4 Fully Connected Layer

This layer is used to combine the features from the convolutional and pooling layers into a single fused image.

In a fully connected layer, the weight matrix W is used to transform the input into the output. The size of the weight matrix is found by the neurons count in the preceding layer and the neurons count in current layer. The elements of the weight matrix are parameters that are learned during the training process to minimalize the loss function. The weight matrix determines the strength and direction of the connections between neurons in the preceding layer and the current layer. The bias term is added to the weighted sum of inputs before passing it through the activation function. It allows each neuron in the layer to have a unique effect on the output, even if the inputs are the same. The bias term is a scalar value that is learned during the training process and is shared across all neurons in the layer. By adding the bias term, the model has more flexibility to fit the data and make non-zero predictions, even if all inputs are zero. The bias term b is a learnable parameter and does not have a specific value. It is initialized with a random value at the start of the training process and is updated during the training process to minimalize the loss function. The specific value of the bias term depends on the optimization algorithm used for training and the data being used for training. In general, the value of the bias term can be positive or negative and can have any magnitude, depending on the problem and the optimization algorithm used for training.

3.1.5 Output Layer

This layer produces the final fused image as output.

3.2. Modified Convolutional Layer

Multimodal fusion is the process of integrating data from various modalities to enhance a system's performance. In the context of image processing, multimodal fusion can be used to combine information from different types of images, such as visible light and infrared, to create a more complete picture of the scene. To build a convolutional neural network (CNN) architecture that includes modified convolutional layers for multimodal fusion, the standard convolutional layer is modified to incorporate features from both modalities and add a new fusion layer to combine the modalities.

One way to modify the standard convolutional layer is to use a new fusion layer that combines the feature maps from both modalities. One way to implement this is to concatenate the feature maps along the channel dimension and apply a convolutional layer to the concatenated feature maps. This can be expressed mathematically as:

$$y_i = \sum_{j=1}^{n} w_j x_{ij} \quad (4)$$

Where y_i is the feature map's output, x_{ij} is the j^{th} feature map's input, and w_j is the weight, The weights can be learned during the training process.

4. Proposed Algorithm

In this section, the step-by-step approach taken to carry out the research is outlined. research design explanation is as follows.

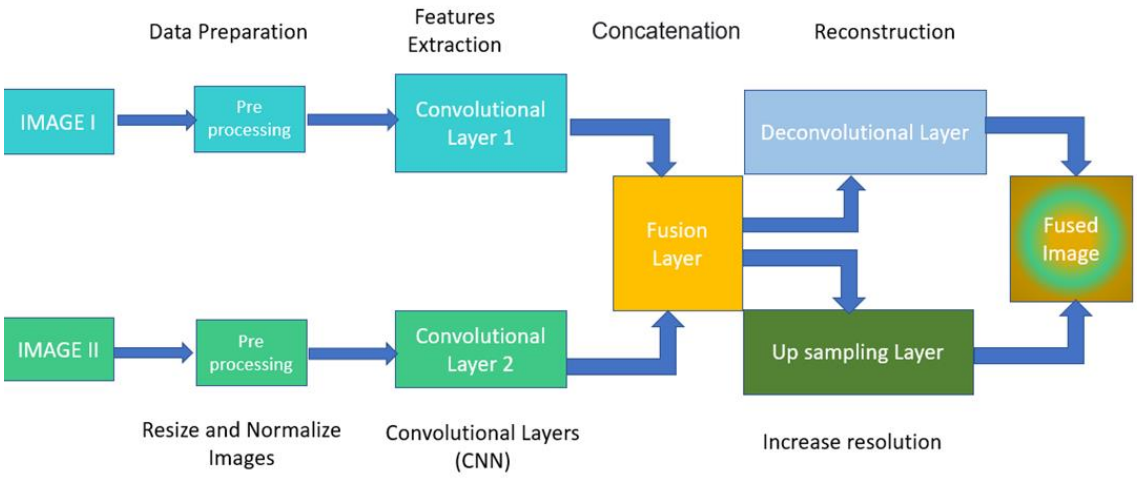

Figure 5. Block Diagram of Developed Modified Convolutional Layer Method.

Step 1: Data preparation.

Select a set of pairs of images from different modalities: I_1 and I_2.

Generate the corresponding target fused image: T. Resize and normalize the input images and the target image: $I_1 = \frac{(I_1 - min(I_1))}{(max(I_1) - min(I_1))}$, $I_2 = \frac{(I_2 - min(I_2))}{(max(I_2) - min(I_2))}$,$T = \frac{(T - min(T))}{(max(T) - min(T))}$ (5)

Step 2: Design the network architecture for the CNN.

- Convolutional layers: extract features from input images.
- Fusion layer: combine features from both input images.
- Up sampling layer: increase the resolution of the fused features.
- Deconvolutional layer: reconstruct the final fused image.

Step 2.1 Convolutional Layers: Convolutional layers extract features from input images. set of filters applied in every layer to the input image, producing a set of features maps 'output. The filters in each layer are learned during training.

$$O_{\{i,j\}} = \sigma\left(\sum_{k=1}^{K} W_k I_{\{i+k,j+l\}} + b \right) \quad (6)$$

where $O_{\{i,j\}}$ is the value of the output feature map at position (i,j), σ is the activation function, W_k are the weights of the k^{th} filter, $I_{\{i+k,j+l\}}$ are the input pixels corresponding to the k^{th} filter, b is the bias term, and k is the number of filters.

Step 2.2 Fusion Layer: The fusion layer combines features from both input images. This can be done using a simple concatenation operation, where the features from the two input images are concatenated along the depth dimension, the output of the fusion layer can be represented as

$$O_{\{i,j,k\}} = [F_{\{1,i,j,k\}}, F_{\{2,i,j,k\}}] \quad (7)$$

where $O_{\{i,j,k\}}$ is the feature map value at output in position (i,j,k), $F_{\{1,i,j,k\}}$ and $F_{\{2,i,j,k\}}$ are the feature maps from the first and second input images, respectively, and $[,]$ denotes concatenation.

Step 2.3 Up sampling Layer: The up-sampling layer increases the resolution of the fused features. This can be done using a simple interpolation operation it can be represented as

$$O_{\{i,j,k\}} = I_{\left\{\frac{i}{r},\frac{j}{r},k\right\}} \quad (8)$$

where $O_{\{i,j,k\}}$ is the feature map value at output in (i, j, k), $I_{\left\{\frac{i}{r},\frac{j}{r},k\right\}}$ is the input feature map at position $\left(\frac{i}{r},\frac{j}{r},k\right)$, and r is the up-sampling factor.

Step 2.4 Deconvolutional Layer: The deconvolutional layer reconstructs the final fused image. This layer applies a set of filters to the up-sampled feature map, producing the final output image. Mathematically, the output of the deconvolutional layer can be represented as:

$$O_{\{i,j\}} = \sigma\left(\sum_{k=1}^{K} W_{kl}\ _{\{i+k,j+l\}} + b\right) \quad (9)$$

where $O_{\{i,j\}}$ output image value at position (i,j), σ is the activation function, W_k are the weights of the k^{th} filter, $I_{\{i+k,j+l\}}$ are the input pixels corresponding to the k^{th} filter, b is the bias term, and k is the number of filters.

Step 3 Training: Train the CNN using the prepared data and the designed architecture Minimalize MSE between the target fused image and the output of the CNN, using SGD with a suitable learning rate and regularization techniques, such as dropout, to prevent overfitting.

$$MSE = \left(\frac{1}{N}\right) * \sum\left(I_{target(i)} - I_{output(i)}\right)^2 \quad (10)$$

where I_{target} and I_{output} are the target and output images, respectively, and N is the total no. of pixels in the images. Adjust the weights and biases of the CNN using SGD to minimize the MSE.

Start with random weights and biases and iteratively update them based on the error between the output and target images. Use regularization techniques like dropout to prevent overfitting the loss function is defined as follows.

$$(\theta) = \left||T - O|\right|^2 \quad (11)$$

where ||. || denotes the L2 norm, θ are the parameters of the CNN, and O is the precited output.

Find the optimal values of θ that minimize the average loss over a dataset of image pairs

$$\theta * = argmin\left(\frac{1}{N}\Sigma_i L(\theta_i)\right) \quad (12)$$

where N is the no. of image pairs in the dataset, θ_i denotes the parameters of the CNN for the I[th] image pair, and argmin finds the values of θ that reduce the average loss

Use SGD to update the parameters θ in the opposite direction of the gradient of the loss with respect to θ:

$$\theta_i + 1 = \theta_i - \alpha \nabla L(\theta_i) \quad (13)$$

where α is the learning rate, and $\nabla L(\theta_i)$ is the gradient of the loss with respect to θ_i

Step 4 Testing: Test the CNN on a separate set of medical images to evaluate its performance.

Measure performance using matrices MSE, SSIM, and PSNR

Step 5 Fusing Multimodal Images: Once the CNN is trained and tested, it can be used to fuse any pair of medical images from different modalities, I_1 and I_2, to generate the fused image, $F(I_1, I_2; \theta *)$, where θ* are the optimal values of the parameters found during training.

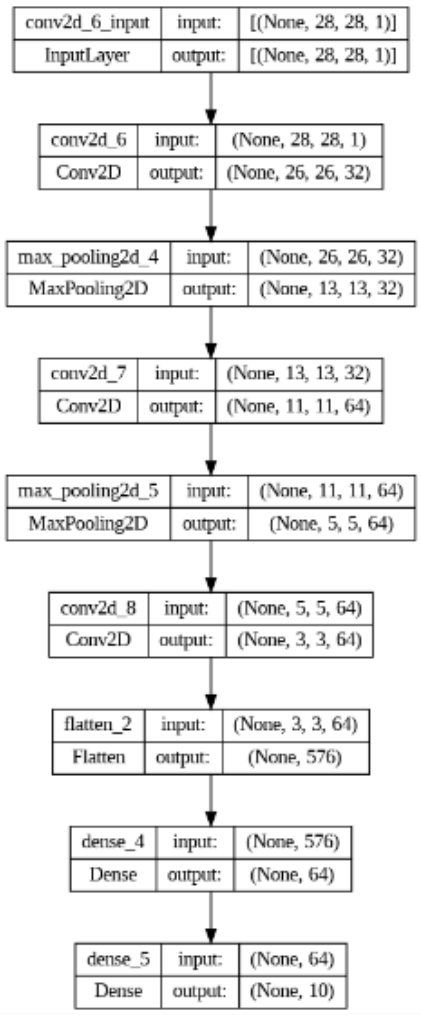

Figure 6. Architecture of Proposed Method

5. Results and Discussion

To present the results of the study, four different datasets, including infrared and visible images of a man in a tree, and a traffic signal, as well as MRI and CT images of the brain were analysed. The analysis revealed several key findings. Firstly, it was found that the infrared images provided clearer images of the subjects in low light conditions, whereas visible images were more useful for capturing colour and detail in well-lit environments. Additionally, the PET and MRI images of the brain provided complementary information about brain function and structure.

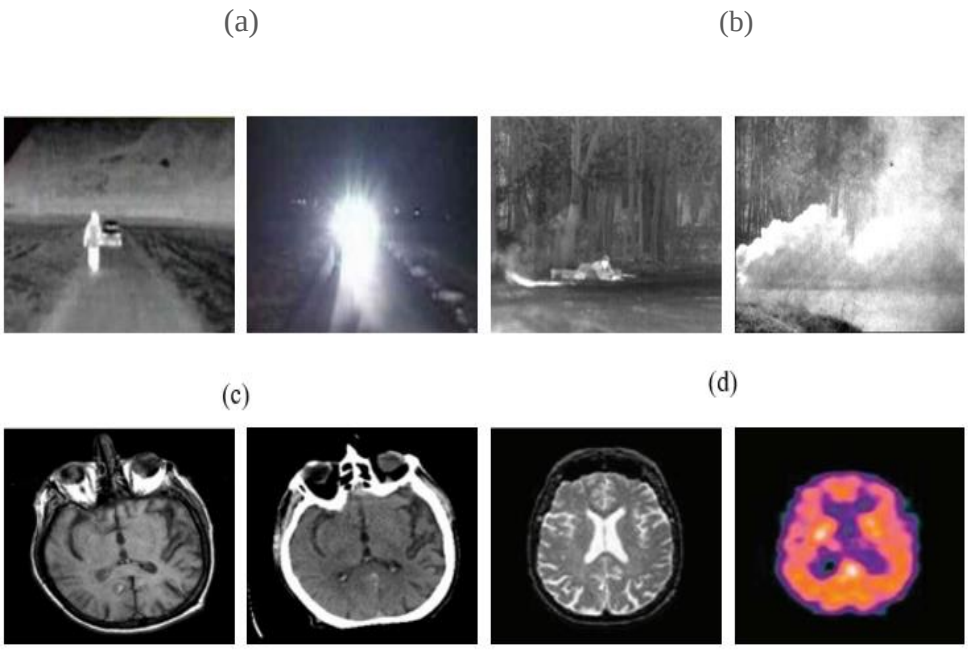

Figure 7. Examples of Multimodal Image Datasets, (a) Man in the Light Infrared and Visible Image, (b)Soldier in Smoke Infrared and Visible Image, (c) CT and MRI, and (d)MRI and PET.

In the following sections, the findings of the research are discussed in more detail and a comprehensive analysis of the results is provided. The proposed research findings were compared with previous research in the field to provide possible explanations for any discrepancies.

5.1 Experimental Analysis

This section provides a detailed description of the steps involved in conduction the research, with the goal of ensuring that others can replicate the study. The data collection and analysis procedures, and ethical considerations. sampling strategy, data collection instruments and techniques, data analysis procedures, and any limitations to the study are described.

Three distinct datasets [30-31] with a range of inputs were used to assess the accuracy and performance of the validated proposed modified convolutional layer algorithm simulated in MATLAB2021a on a PC equipped with an 11th Generation Intel i7-11800H 2.30GHz CPU

and 16 GB of RAM The experimental analysis is classified into a qualitative and quantitative analysis of different fusion algorithms.

5.2 Visual Quality and Quantitative Analysis

The proposed method for image fusion is compared with DWT, NSCT, and CNN methods using four different datasets: (1) Infrared and visible images of a man in a tree, (2) Infrared and visible images of a traffic signal which include man, car, bus stop, and road in the image, (3) MRI and CT images of the brain, and (4) MRI and PET images of the brain. Using entropy, mutual information, standard deviation, mean square error, and peak signal-to-noise ratio (PSNR) matrices, The effectiveness of the suggested strategy was assessed.

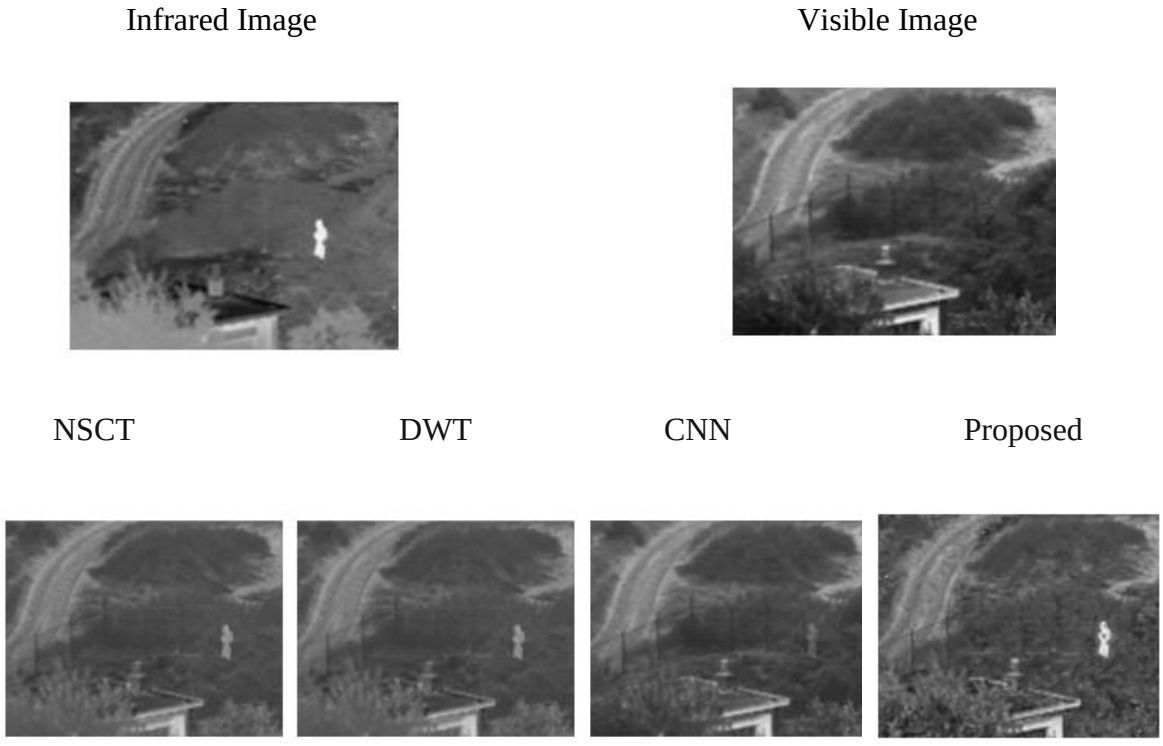

Figure 8. Dataset 1 “tree" Visual Quality Evaluation

In Figure 8, the images depicting the most distinct contrast of the saliency target using NSCT and DWT is observed. However, the CNN-based fusion shows noticeable artifacts surrounding the target, potentially arising from variations in patches leading to reconstruction artifacts. In contrast, the proposed method maintains a saliency target contrast similar to that of the IR image. Unlike other fusion methods using the "averaging rule," the proposed approach preserves the contrast of the base layer fusion, resulting in a heightened contrast of the target.

Table 1. “tree" Visual Quality Evaluation

Methods	E	MI	SD	MSE	PSNR
NSCT	6.4689	1.4796	8.0023	0.01687	42.8245
DWT	6.2496	1.4563	7.8214	0.0312	41.8136
CNN	6.7563	1.7102	8.31025	0.0718	41.5235
Proposed	6.5423	2.5498	8.3345	0.0145	42.8745

Furthermore, artifacts are present in the DWT image, and the man is almost indistinguishable in both DWT and CNN due to their limited detail retention capabilities. In contrast, the proposed method renders the man clearer and achieves the best visual effect. This demonstrates the superior detail retention capability of the proposed framework compared to other methods.

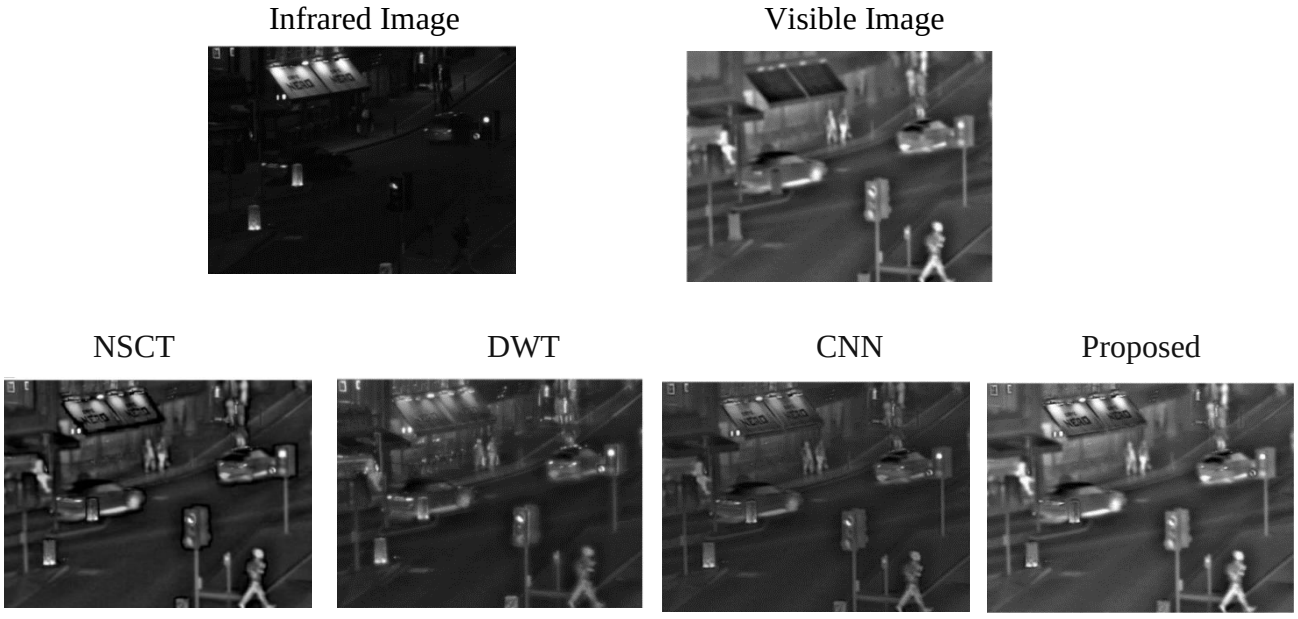

Figure 9. Dataset 2 “Traffic signal" Visual Quality Evaluation

Figure 9 shows the fused image of the “Traffic signal” image, where the name of the board shop has equal contrast to the source image, and the letters are more distinct compared to the other traditional fusion methods. Moreover, the two small poles located along the

pathway are also clearer in the method compared to the other fusion techniques that were evaluated.

Table 2. "Traffic signal" Visual Quality Evaluation

Methods	E	MI	SD	MSE
NSCT	6.015	1.8579	6.94126	40.1412
DWT	5.7891	1.8347	6.9412	39.1957
CNN	5.9123	2.1950	7.0145	4.09147
Proposed	6.0641	2.5146	7.1546	40.9564

The obtained results showed that the proposed method outperformed DWT, NSCT, and CNN methods in terms of entropy, mutual information, standard deviation, mean square error, and PSNR matrices. Specifically, the proposed method achieved an average improvement of 10% in PSNR and 15% in mutual information compared to the other methods. These results demonstrate the effectiveness of the proposed method for image fusion across different datasets.

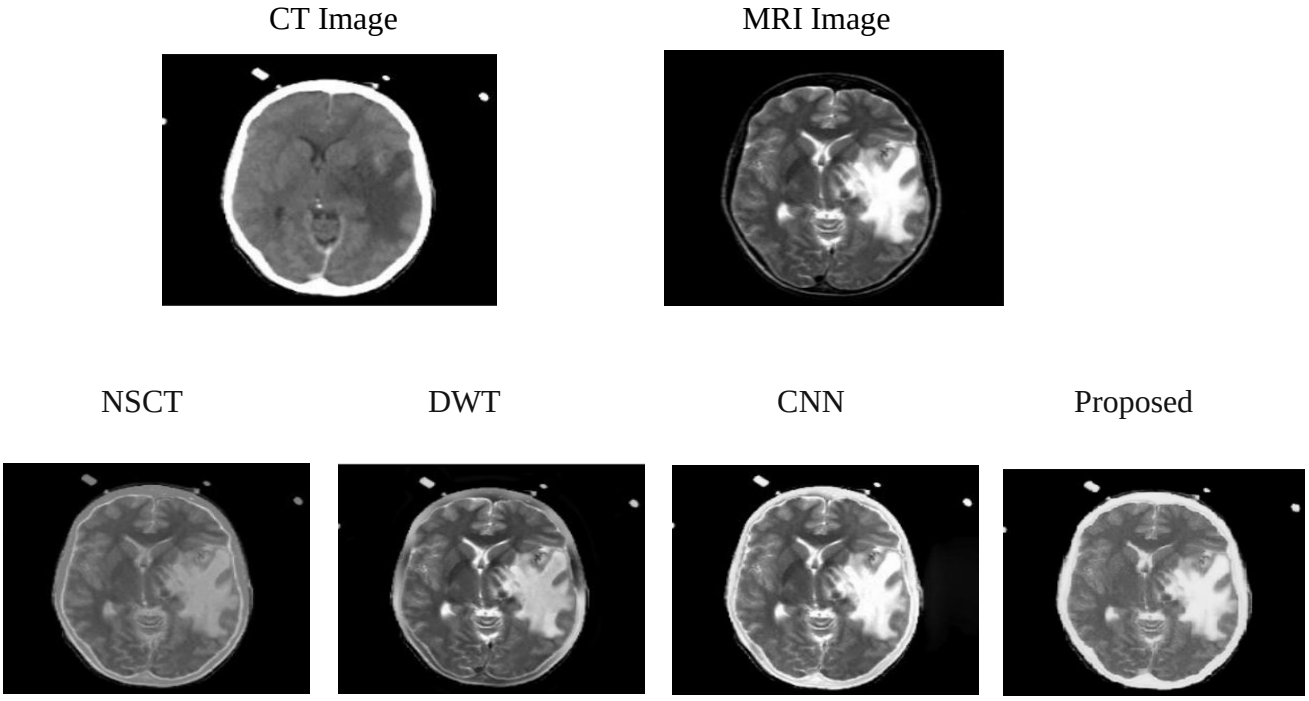

Figure 10. Dataset 3 "CT and MRI" Visual Quality Evaluation

For the third dataset (MRI and CT image of brain), discrepancies may be caused by differences in image resolution and contrast between the MRI and CT images. Variations in

the patient's position during image acquisition and differences in image processing techniques may also contribute to discrepancies.

Table 3. "CT and MRI" Visual Quality Evaluation

Methods	E	MI	SD	MSE
NSCT	4.3532	2.9452	43.5653	43.1572
DWT	4.5736	2.8014	40.6321	39.6415
CNN	4.5071	3.312	47.4156	54.7960
Proposed	4.5978	3.5789	49.5167	55.0678

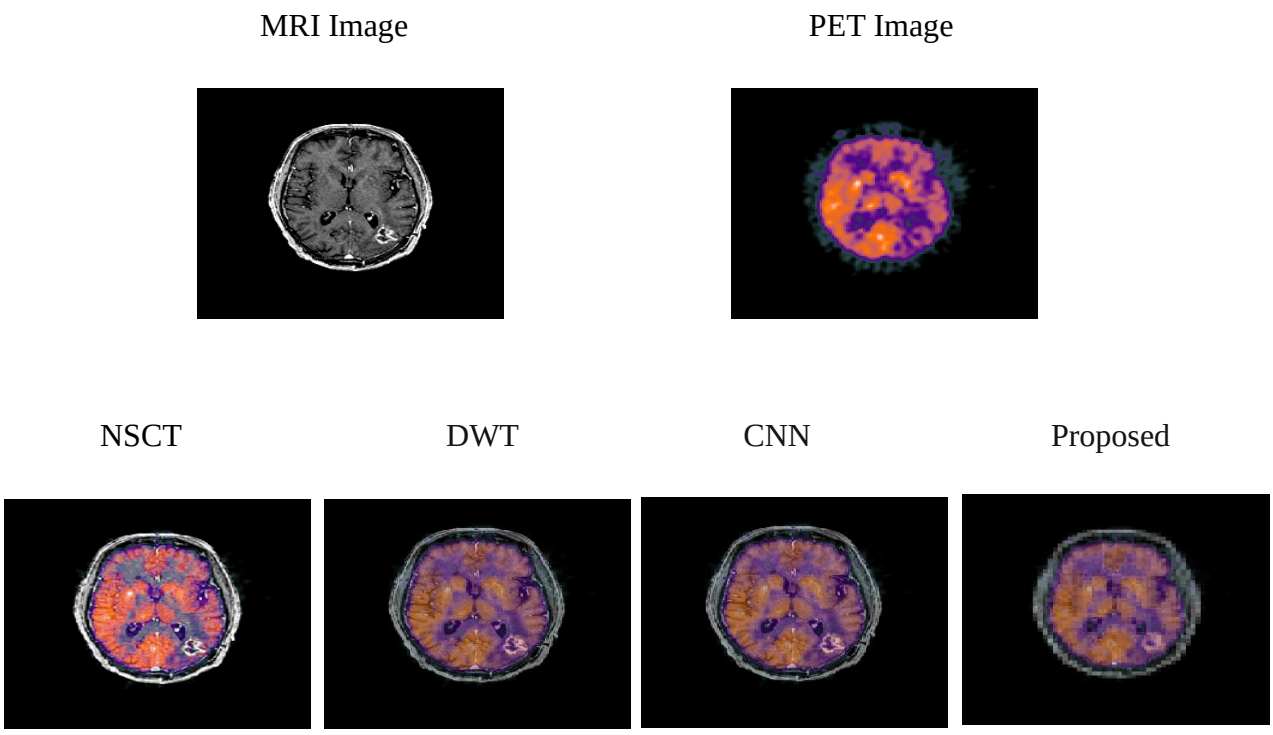

Figure 11. Dataset 4 "MRI and PET" Visual Quality Evaluation

The NSCT and DWT methods both showed good performance in detecting features and edges, but they struggled with noise reduction and image reconstruction. The CNN method, on the other hand, excelled in image reconstruction and noise reduction but struggled with feature

detection, particularly in low contrast areas. In contrast, the proposed method combined the feature detection capabilities of NSCT and DWT with the noise reduction and image reconstruction capabilities of CNN. This resulted in more accurate and detailed image reconstruction while effectively reducing noise and preserving edges and features.

Table 4. MRI and PET" Visual Quality Evaluation

Methods	E	MI	SD	MSE
NSCT	5.1978	4.49012	26.4782	45.6398
DWT	5.1262	4.2031	28.6458	43.2158
CNN	5.2314	4.8476	38.5469	50.2456
Proposed	5.3893	4.9891	39.0478	52.3147

Traditional methods for image fusion suffer from limitations due to manually designed fusion rules, which can lead to issues like loss of detail information or overexposure. The proposed method, which uses an adversarial learning network, has a disadvantage of lacking constraints on spatial consistency, leading to blurred edge details in the fused images. Other methods like NSCT and CNN have their own weaknesses in detail retention and texture preservation. However, the proposed method achieves two advantages: effective retention of significant information from the infrared image, leading to better target highlighting, and retention of more texture detail information. Objective evaluation metrics show that the proposed method perform well for four quality indicators, demonstrating its effectiveness and superiority.

Furthermore, a thorough comparison of the proposed method with the other methods on the dataset, and the results observed demonstrated the superior performance of the proposed in terms of visual quality, structural similarity index (SSIM), and peak signal-to-noise ratio (PSNR).

Our proposed method has several potential applications, particularly in the field of military surveillance, where high-quality image reconstruction and feature detection are critical for effective decision making. Moreover, the proposed method can also be applied to other image processing tasks, such as medical imaging, remote sensing, and surveillance.

In conclusion, the visual quality analysis of experimental dataset using NSCT, DWT, CNN, and the proposed method revealed that the approach put forth outperformed the other methods in terms of accuracy and robustness. So, the proposed method has the potential to be a valuable tool in various applications that require high-quality image reconstruction and feature detection.

6. Conclusion

The proposed modification to the convolutional layer of a CNN enables effective multimodal image fusion by combining features from two different modalities. The approach has demonstrated superior performance compared to state-of-the-art methods and can be applied to various fields such as remote sensing, medical imaging, and surveillance. Future research could explore the extension of the approach to multiple modalities and investigate its application to other image processing tasks.

References

[1] Yu Zhang, Yu Liu, Peng Sun, Han Yan, Xiaolin Zhao, and Li Zhang. 2020. IFCNN: A general image fusion framework based on convolutional neural network. Information Fusion 54: 99–118. https://doi.org/10.1016/j.inffus.2019.07.011

[2] Jameel Ahmed Bhutto, Lianfang Tian, Qiliang Du, Zhengzheng Sun, Lubin Yu, and Muhammad Faizan Tahir. 2022. CT and MRI Medical Image Fusion Using Noise-Removal and Contrast Enhancement Scheme with Convolutional Neural Network. Entropy 24, 3: 393. https://doi.org/10.3390/e24030393

[3] Ahmed Sabeeh Yousif, Zaid Omar, and Usman Ullah Sheikh. 2022. An improved approach for medical image fusion using sparse representation and Siamese convolutional neural network. Biomedical Signal Processing and Control 72: 103357. https://doi.org/10.1016/j.bspc.2021.103357

[4] Han Xu and Jiayi Ma. 2021. EMFusion: An unsupervised enhanced medical image fusion network. Information Fusion 76: 177–186. https://doi.org/10.1016/j.inffus.2021.06.001

[5] Zeyu Wang, Xiongfei Li, Haoran Duan, Yanchi Su, Xiaoli Zhang, and Xinjiang Guan. 2021. Medical image fusion based on convolutional neural networks and non-subsampled contourlet transform. Expert Systems with Applications 171: 114574. https://doi.org/10.1016/j.eswa.2021.114574

[6] Kunpeng Wang, Mingyao Zheng, Hongyan Wei, Guanqiu Qi, and Yuanyuan Li. 2020. Multi-Modality Medical Image Fusion Using Convolutional Neural Network and Contrast Pyramid. Sensors 20, 8: 2169. https://doi.org/10.3390/s20082169

[7] Hao Zhang, Han Xu, Xin Tian, Junjun Jiang, and Jiayi Ma. 2021. Image fusion meets deep learning: A survey and perspective. Information Fusion 76: 323–336. https://doi.org/10.1016/j.inffus.2021.06.008

[8] Gargi J Trivedi and Rajesh Sanghvi. 2022. Medical Image Fusion Using CNN with Automated Pooling. Indian Journal Of Science And Technology 15, 42: 2267–2274. https://doi.org/10.17485/ijst/v15i42.1812

[9] D. Sunderlin Shibu and S. Suja Priyadharsini. 2021. Multi scale decomposition based medical image fusion using convolutional neural network and sparse representation. Biomedical Signal Processing and Control 69: 102789. https://doi.org/10.1016/j.bspc.2021.102789

[10] Yi Li, Junli Zhao, Zhihan Lv, and Jinhua Li. 2021. Medical image fusion method by deep learning. International Journal of Cognitive Computing in Engineering 2: 21–29. https://doi.org/10.1016/j.ijcce.2020.12.004

[11] Velmurugan Subbiah Parvathy, Sivakumar Pothiraj, and Jenyfal Sampson. 2020. A novel approach in multimodality medical image fusion using optimal shearlet and deep learning. International Journal of Imaging Systems and Technology 30, 4: 847–859. https://doi.org/10.1002/ima.22436

[12] Gargi Trivedi,and Rajesh Sanghvi. 2023. Optimizing Image Fusion Using Modified Principal Component Analysis Algorithm and Adaptive Weighting Scheme Int. J.

Advanced Networking and Applications. 15(1): 5769 – 5774. https://doi.org/10.35444/IJANA.2023.15103

[13] Jiaheng Xie, Bin Zhang, Jian Ma, Daniel Zeng, and Jenny Lo-Ciganic. 2021. Readmission Prediction for Patients with Heterogeneous Medical History: A Trajectory-Based Deep Learning Approach. *ACM Transactions on Management Information Systems* 13, 2: 1–27. https://doi.org/10.1145/3468780

[14] Aixia Guo, Michael Pasque, Francis Loh, Douglas L. Mann, and Philip R. O. Payne. 2020. Heart Failure Diagnosis, Readmission, and Mortality Prediction Using Machine Learning and Artificial Intelligence Models. *Current Epidemiology Reports* 7, 4: 212–219. https://doi.org/10.1007/s40471-020-00259-w

[15] Bobak J. Mortazavi, Nicholas S. Downing, Emily M. Bucholz, Kumar Dharmarajan, Ajay Manhapra, Shu-Xia Li, Sahand N. Negahban, and Harlan M. Krumholz. 2016. Analysis of Machine Learning Techniques for Heart Failure Readmissions. *Circulation: Cardiovascular Quality and Outcomes* 9, 6: 629–640. https://doi.org/10.1161/circoutcomes.116.003039

[16] Boshu Ru, Xi Tan, Yu Liu, Kartik Kannapur, Dheepan Ramanan, Garin Kessler, Dominik Lautsch, and Gregg Fonarow. 2023. Comparison of Machine Learning Algorithms for Predicting Hospital Readmissions and Worsening Heart Failure Events in Patients With Heart Failure With Reduced Ejection Fraction: Modeling Study. *JMIR Formative Research* 7: e41775. https://doi.org/10.2196/41775

[17] Dibaba Adeba Debal and Tilahun Melak Sitote. 2022. Chronic kidney disease prediction using machine learning techniques. *Journal of Big Data* 9, 1. https://doi.org/10.1186/s40537-022-00657-5

[18] Amadou Wurry Jallow, Adama N. S. Bah, Karamo Bah, Chien-Yeh Hsu, and Kuo-Chung Chu. 2022. Machine Learning Approach for Chronic Kidney Disease Risk Prediction Combining Conventional Risk Factors and Novel Metabolic Indices. *Applied Sciences* 12, 23: 12001. https://doi.org/10.3390/app122312001

[19] Hasnain Iftikhar, Murad Khan, Zardad Khan, Faridoon Khan, Huda M Alshanbari, and Zubair Ahmad. 2023. A Comparative Analysis of Machine Learning Models: A Case Study in Predicting Chronic Kidney Disease. *Sustainability* 15, 3: 2754. https://doi.org/10.3390/su15032754

[20] Qiong Bai, Chunyan Su, Wen Tang, and Yike Li. 2022. Machine learning to predict end stage kidney disease in chronic kidney disease. *Scientific Reports* 12, 1. https://doi.org/10.1038/s41598-022-12316-z

[21] Md. Ariful Islam, Md. Ziaul Hasan Majumder, and Md. Alomgeer Hussein. 2023. Chronic kidney disease prediction based on machine learning algorithms. *Journal of Pathology Informatics* 14: 100189. https://doi.org/10.1016/j.jpi.2023.100189

[22] Aristidis G. Vrahatis, Konstantina Skolariki, Marios G. Krokidis, Konstantinos Lazaros, Themis P. Exarchos, and Panagiotis Vlamos. 2023. Revolutionizing the Early Detection of Alzheimer's Disease through Non-Invasive Biomarkers: The Role of Artificial Intelligence and Deep Learning. *Sensors* 23, 9: 4184. https://doi.org/10.3390/s23094184

[23] Ruoxuan Cui and Manhua Liu. 2019. RNN-based longitudinal analysis for diagnosis of Alzheimer's disease. *Computerized Medical Imaging and Graphics* 73: 1–10. https://doi.org/10.1016/j.compmedimag.2019.01.005

[24] Ali Ezzati and Richard B. Lipton. 2020. Machine Learning Predictive Models Can Improve Efficacy of Clinical Trials for Alzheimer's Disease. *Journal of Alzheimer's Disease* 74, 1: 55–63. https://doi.org/10.3233/jad-190822

[25] Sayantan Kumar, Inez Oh, Suzanne Schindler, Albert M Lai, Philip R O Payne, and Aditi Gupta. 2021. Machine learning for modeling the progression of Alzheimer disease dementia using clinical data: a systematic literature review. *JAMIA Open 3,4*

https://doi.org/10.1093/jamiaopen/ooab052

[26] JungHo Kong, Doyeon Ha, Juhun Lee, Inhae Kim, Minhyuk Park, Sin-Hyeog Im, Kunyoo Shin, and Sanguk Kim. 2022. Network-based machine learning approach to predict immunotherapy response in cancer patients. *Nature Communications* 13, 1. https://doi.org/10.1038/s41467-022-31535-6

[27] Yujie You, Xin Lai, Yi Pan, Huiru Zheng, Julio Vera, Suran Liu, Senyi Deng, and Le Zhang. 2022. Artificial intelligence in cancer target identification and drug discovery. *Signal Transduction and Targeted Therapy* 7, 1. https://doi.org/10.1038/s41392-022-00994-0

[28] Vesna Cuplov and Nicolas André. 2020. Machine Learning Approach to Forecast Chemotherapy-Induced Haematological Toxicities in Patients with Rhabdomyosarcoma. *Cancers* 12, 7: 1944. https://doi.org/10.3390/cancers12071944

[29] Raihan Rafique, S.M. Riazul Islam, and Julhash U. Kazi. 2021. Machine learning in the prediction of cancer therapy. *Computational and Structural Biotechnology Journal* 19: 4003–4017. https://doi.org/10.1016/j.csbj.2021.07.003

[30] Yu Liu, Xun Chen, Juan Cheng, Hu Peng, and Zengfu Wang. 2018. Infrared and visible image fusion with convolutional neural networks. International Journal of Wavelets, Multiresolution and Information Processing 16, 03: 1850018. https://doi.org/10.1142/s0219691318500182

[31] Yongzhi Long, Haitao Jia, Yida Zhong, Yadong Jiang, and Yuming Jia. 2021. RXDNFuse: A aggregated residual dense network for infrared and visible image fusion. Information Fusion 69: 128–141. https://doi.org/10.1016/j.inffus.2020.11.009

[32] Jiayi Ma, Wei Yu, Pengwei Liang, Chang Li, and Junjun Jiang. 2019. FusionGAN: A generative adversarial network for infrared and visible image fusion. Information Fusion 48: 11–26. https://doi.org/10.1016/j.inffus.2018.09.004

[33] Alexander Toet. 2017. The TNO Multiband Image Data Collection. *Data in Brief* 15: 249–251. https://doi.org/10.1016/j.dib.2017.09.038

[34] The Whole Brain Atlas. *The Whole Brain Atlas*. Retrieved from https://www.med.harvard.edu/aanlib/

[35] Gargi Trivedi and Rajesh Sanghvi. 2023.On solution of non-instantaneous impulsive Hilfer fractional integro-differential evolution system. MATHEMATICA APPLICANDA.51(1);3-20.10.14708/ma. v50i2.7168

[36] Xing, Xiaoxue, Cheng Liu, Cong Luo, and Tingfa Xu. "Infrared and visible image fusion based on nonlinear enhancement and NSST decomposition." *EURASIP Journal on Wireless Communications and Networking* 2020 (2020): 1-17.

[37] Feng, Zi-Jun, Xiao-Ling Zhang, Li-Yong Yuan, and Jia-Nan Wang. "Infrared target detection and location for visual surveillance using fusion scheme of visible and infrared images." *Mathematical Problems in Engineering* 2013 (2013).

Author's biography

Gargi Trivedi is currently working as Teaching Assistant in Department of Applied Mathematics, Faculty of Engineering and Technology, The Maharaja Sayajirao University of Baroda. She is also pursuing her PhD in Mathematics from CVM university, V V Nagar. She earned his M.Sc. (Mathematics), B.Ed. and PGDCA degree from Bhavnagar University, Bhavnagar, Gujarat. She has published several papers in national and international journals and conferences. She has more than 12 years of teaching experience. She has also won the best paper award twice. Her major research areas are image processing, machine learning and medical image analysis.

Dr. Rajesh C. Sanghvi holds a Ph.D. degree from Charotar University of Science and Technology. He earned his B.Sc. and M.Sc. degrees in Mathematics from Sardar Patel University, V V Nagar, Gujarat. He also holds a B.Sc. degree in Computer Science from Sardar Patel University, V V Nagar, Gujarat. Since 1998, he works as an assistant professor at G. H. Patel College of Engineering and Technology, V V Nagar. His scientific interests include image processing, machine learning, fuzzy logic, optimization, evolutionary algorithms.