EDITORIAL

Saliency Object Detection-Based Medical Image Fusion: Future Directions for Smart Healthcare Systems

Manoj Diwakar [1,*]

[1] Department of Computer Science and Engineering, Graphic Era University, Dehradun 248002, India

Abstract

Saliency Object Detection-Based Medical Image Fusion concentrates on the most prominent anatomical parts that naturally attract attention. The process begins with a two-step pipeline. First, it identifies the salient regions those structures most critical to a diagnostic decision. These highlighted areas are then fed into a guided-filtering framework. This work blends them with complementary information from a second imaging modality, such as CT and MRI. The method uses total variation regularization to suppress noise while preserving edges. A saliency-based weighting scheme ensures that every key detail is retained. The result is a single, high-quality image that carries the full diagnostic power of both inputs, providing physicians with a clearer, more informative view without sacrificing clinically relevant information.

Keywords: medical image fusion, object detection, total variation, saliency features.

Submitted: 01 October 2025
Accepted: 20 January 2026
Published: 02 February 2026

Vol. 2, **No.** 1, 2026.
10.62762/BISH.2025.918243

***Corresponding author:**
Manoj Diwakar
manoj.diwakar@gmail.com

1 Introduction

Medical image fusion plays a vital role in modern clinical diagnosis by integrating complementary information from multiple imaging modalities into a single, more informative image [1–4]. Here, the methodology for medical image fusion is divided into three segments: dual-image decomposition, map construction, and dual-image reconstruction. The first segment involves segmenting two input images into base-layer images using total variation. The second segment, weight map construction, is a pipeline that includes operations such as saliency object detection (SOD), morphological operations, and conversion to a binary image. This step aims to outline the most visually salient features or regions in the source image. The third stage is a two-stage reconstruction, which uses a fusion technique to combine the segmented images from both stages and produce an output image with the maximum possible information.

2 Methods and Materials

2.1 Total Variation

Total variation [7–9] is used to measure the amount of changes present in an image. Primarily, it provides image denoising or image restoration. In addition to suppressing noise, it helps preserve important edges and the image's structure. This is achieved by minimizing the total variation of the image,

Citation
Diwakar, M. (2026). Saliency Object Detection-Based Medical Image Fusion: Future Directions for Smart Healthcare Systems. *Biomedical Informatics and Smart Healthcare*, 2(1), 1–4.

encouraging smoothness in regions with low variation while allowing sharp transitions at edges. With this method, images with less variation or fewer changes in pixel intensity tend to appear smoother and less noisy.

2.2 Saliency Object Detection (SOD)

Saliency object detection [5, 6] refers to the process of identifying the most visually significant or salient objects within an image. A saliency score is assigned after mapping the image's pixels. This score defines the degree of saliency for each pixel.

2.3 Morphological Operation

Morphological operation involves the manipulation of shape and structure to enhance, extract, or analyze features within an image using techniques such as erosion, dilation, opening, and closing [10].

2.4 Medical Image fusion Dataset

Various publicly available medical imaging datasets [11–13] provide multimodal data such as CT, MRI, PET, and SPECT, which are widely used for medical image fusion research and experimental validation.

3 Medical image fusion Processing

3.1 Dual Decomposition

Two input images I_1 and I_2 first go through total variation. From this, base images B_1 and B_2 are obtained as shown below:

$$B_i = I_i * \mathrm{TV}, \tag{1}$$

where the abbreviations represent the following terminologies: base layer (B), input image (I), and total variation (TV).

By subtracting the base layer from the source image, the detailed layer image can be obtained:

$$D_i = I_i - B_i, \tag{2}$$

where D_i represents the detailed layer images.

Here, the detail image (D_i) captures small variations, while the base layer (B_i) captures large variations.

3.2 Saliency Detection with Morphological Operations and Guided Filtering

Numerous methods exist for applying Salient Object Detection (SOD) to an image. Here, a frequency-based method is applied to the input image. This method analyzes the frequency content of an image to identify salient objects:

$$F_{bn} = \mathrm{FB}(S_i) \tag{3}$$

where F_{bn} represents the frequency-based output of the source image.

Next, shape manipulation and analysis are performed using a morphological operation on the obtained image:

$$M_n = F_{bn} * M_o \tag{4}$$

where M_n, F_{bn}, and M_o represent the output of the morphological operation, the frequency-based image, and the morphological operation, respectively.

Binary image conversion is then performed using maximum entropy thresholding on the obtained image. Converting an image to binary allows clear separation of the foreground and background:

$$E_n = M_n * E_t \tag{5}$$

where E_n, M_n, and E_t represent the output of maximum entropy thresholding, the morphological image, and the maximum entropy thresholding, respectively.

A guided filter is now applied to the binary image, using the source image as the guidance image. Two guided images are obtained: one with base-layer details and the other with detailed-layer details, denoted as G and G_D, respectively:

$$G_n = G_{a_1,b_1}(P_x, l) \tag{6}$$

$$G_{D_n} = G_{a_2,b_2}(P_x, l) \tag{7}$$

where a_1, b_1, a_2, b_2 are parameters of the guided filter, and G_n, G_{D_n} are the resulting images after guided filtering.

3.3 Dual Image Reconstruction

The dual image reconstruction stage is carried out in two sequential steps. In the first step, the base layers and detail layers obtained from the decomposition stage are independently fused using a guided filtering–based strategy, which helps preserve important structural information while suppressing artifacts. In the second step, the fused base and detail images are combined to generate the final fused output. This reconstruction approach is computationally efficient and conceptually simple, while guided filtering effectively leverages pixel-level information for optimization. As a result, the final fused image exhibits improved structural consistency, enhanced detail preservation, and better overall visual quality.

4 Conclusion

The article describes a study called saliency-based fusion that combines incompatible CT and MRI scans to produce a clear image. We can divide each modality into a smooth base and a detail-rich layer, a decomposition known as the total-variation (TV) decomposition. the clinically relevant structure and the background noise. Saliency-directed weight mapping assigns greater weight to regions that are anatomically or diagnostically salient (e.g., lesions, edges), thereby making the specific patches dominate the resulting blend. A guided filter is then applied, but it does not smooth edges, leaving them sharp without halos or bleeding artifacts. The multistage procedure, comprising decomposition, saliency weighting, guided filtering, and recombination, produces fused images that retain the most significant diagnostic information from each source while improving contrast and readability.

Data Availability Statement

Not applicable.

Funding

This work was supported without any funding.

Conflicts of Interest

The author declares no conflicts of interest.

AI Use Statement

The author declares that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Huang, B., Yang, F., Yin, M., Mo, X., & Zhong, C. (2020). A review of multimodal medical image fusion techniques. *Computational and mathematical methods in medicine, 2020*(1), 8279342. [CrossRef]

[2] Karim, S., Tong, G., Li, J., Qadir, A., Farooq, U., & Yu, Y. (2023). Current advances and future perspectives of image fusion: A comprehensive review. *Information Fusion, 90*, 185-217. [CrossRef]

[3] Azam, M. A., Khan, K. B., Salahuddin, S., Rehman, E., Khan, S. A., Khan, M. A., ... & Gandomi, A. H. (2022). A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics. *Computers in biology and medicine, 144*, 105253. [CrossRef]

[4] Tan, W., Tiwari, P., Pandey, H. M., Moreira, C., & Jaiswal, A. K. (2025). Multimodal medical image fusion algorithm in the era of big data. *Neural computing and applications, 37*(28), 22995-23015. [CrossRef]

[5] Li, X., Guo, X., Han, P., Wang, X., Li, H., & Luo, T. (2020). Laplacian redecomposition for multimodal medical image fusion. *IEEE Transactions on Instrumentation and Measurement, 69*(9), 6880-6890. [CrossRef]

[6] Babu, B. S., & Narayana, M. V. (2023). Two stage multi-modal medical image fusion with marine predator algorithm-based cascaded optimal DTCWT and NSST with deep learning. *Biomedical Signal Processing and Control, 85*, 104921. [CrossRef]

[7] Vanitha, K., Satyanarayana, D., & GiriPrasad, M. N. (2024). A novel framework of multimodal medical image fusion using adaptive NSST and optimized deep learning approach. *The Imaging Science Journal, 72*(8), 1015-1042. [CrossRef]

[8] Khan, S. U., Khan, F., Ullah, S., & Lee, B. (2023). Multimodal medical image fusion in NSST domain with structural and spectral features enhancement. *Heliyon, 9*(6). [CrossRef]

[9] Liu, Y., Liu, S., & Wang, Z. (2015). A general framework for image fusion based on multi-scale transform and sparse representation. *Information fusion, 24*, 147-164. [CrossRef]

[10] Kahol, A., & Bhatnagar, G. (2024). Deep learning-based multimodal medical image fusion. *Data Fusion Techniques and Applications for Smart Healthcare*, 251-279. [CrossRef]

[11] *Zero-Learning-Fast-Medical-Image-Fusion (MRI-PET image dataset)*. (2025). Retrieved from https://github.com/bsun0802/Zero-Learning-Fast-Medical-Image-Fusion/tree/master/images/MRI-PET

[12] *Multimodal-image-fusion-to-detect-brain-tumors (Brain tumor fusion dataset)*. (2025). Retrieved from https://github.com/ashna111/multimodal-image-fusion-to-detect-brain-tumors/tree/master/dataset

[13] *Brain Tumor Multimodal Image (CT and MRI)*. (2025). Retrieved from https://www.kaggle.com/datasets/murtozalikhon/brain-tumor-multimodal-image-ct-and-mri

Manoj Diwakar is a Professor in the Department of Computer Science and Engineering at Graphic Era University, Dehradun, with over 20 years of academic and industrial experience. He holds a Ph.D. in Computer Science and is a professional member of IEEE and a life member of IAENG. His research focuses on image processing, medical imaging, and information security. Dr. Diwakar has published over 250 research papers in reputed journals, conferences, and book chapters. (Email: manoj.diwakar@gmail.com)