

**MINISTRY OF
EDUCATION
AND TRAINING**

**VIETNAM ACADEMY OF
SCIENCE AND TECHNOLOGY**

GRADUATE UNIVERSITY OF SCIENCE AND TECHNOLOGY



LE THI HONG HA

**RESEARCH ON MEDICAL IMAGE FUSION
METHODS BASED ON EDGE-PRESERVING
DECOMPOSITION AND ADAPTIVE WEIGHTS**

**SUMMARY OF THE DOCTORAL THESIS IN
COMPUTER SCIENCE**

Major: Computer Science

Code: 9 48 01 01.

Hanoi - 2026

The work was completed at: Graduate University of Science and Technology, Vietnam Academy of Science and Technology

Scientific supervisors:

1. Supervisor 1: Assoc. Prof. Dr. Nguyen Long Giang, Institute of Information Technology, Vietnam Academy of Science and Technology.
2. Supervisor 2: Dr. Dinh Phu Hung, Thuyloi University, Ministry of Agriculture and Environment.

Reviewer 1:

Reviewer 2:

The thesis will be defended before the University-level Doctoral Thesis Evaluation Committee, to be held at the Graduate University of Science and Technology, Vietnam Academy of Science and Technology, at o'clock, on day month year

The thesis can be found at:

1. Library of the Graduate University of Science and Technology
2. National Library of Vietnam

INTRODUCTION

1. Rationale for the topic

Medical image fusion is the process of integrating information from multiple images acquired by one or several diagnostic modalities in order to produce an output image that contains more useful information than any individual source image. Integrating anatomical and functional information at the same time improves the representation of lesions and supports diagnosis.

Owing to differences in imaging principles, each modality reflects only a certain aspect of the object. MRI images provide anatomical information with good soft-tissue representation and high spatial resolution; PET and SPECT images reflect the functional and metabolic characteristics of tissue through the distribution of radioactive tracers. Therefore, using each type of image alone often fails to fully reflect the pathological condition.

Current methods fall into two directions: transform-based and deep-learning-based. Traditional approaches depend considerably on the decomposition technique, the fusion strategy and the choice of parameters, so maintaining brightness, contrast and edge features at the same time has not met expectations. Meanwhile, deep learning methods usually require high computational cost, large amounts of training data, and are limited in interpretability.

On this basis, the thesis aims to develop an effective image fusion method that ensures image quality and good controllability. The proposed approach combines image decomposition based on edge-preserving filters with optimization algorithms, focusing on: (i) constructing an adaptive fusion rule for the base component; and (ii) designing an effective fusion strategy for the detail component to balance structural and detail information.

2. Research objectives

2.1. General objective

- To propose a medical image fusion method based on image decomposition and an adaptive fusion strategy, in order to improve the quality of the fused image and enhance the preservation of edge and structural information from the source images.

2.2. Specific objectives

- To study and develop a medical image fusion method that combines edge-preserving filters and optimization algorithms in order to improve the brightness, contrast and edge preservation of the source images.
- To study and build a medical image fusion model based on three-component decomposition and an adaptive fusion strategy in order to improve fusion effectiveness and preserve important information.

3. Research content

- To study the image fusion problem, medical imaging modalities, image decomposition and transformation methods, fusion strategies, optimization algorithms, energy functions and image quality evaluation criteria.
- To review related studies, identify their strengths, weaknesses and research gaps, and thereby determine a suitable approach.
- To propose a medical image fusion method based on edge-preserving filters combined with optimization algorithms in order to improve the quality of the fused image and the preservation of information from the source images.
- To carry out experiments on the MRI–PET and MRI–SPECT datasets.
- To evaluate and compare the proposed methods with a number of representative image fusion methods published from 2020 to 2025.

4. Research subjects and scope

4.1. Research subjects

- The medical image fusion problem; the MRI, PET and SPECT imaging modalities.
- Image decomposition, feature extraction and fusion strategy techniques.
- Optimization algorithms, edge-preserving filters, structure tensors and energy functions; image quality evaluation criteria.

4.2. Research scope

- The multimodal medical image fusion problem for MRI–PET and MRI–SPECT.

- A survey of image decomposition techniques and fusion strategies aimed at improving the quality of the output image and preserving information from the source images.

- Experiments on two-dimensional brain images collected from The Whole Brain Atlas; the results are compared with representative medical image fusion methods.

5. Research methods

Theoretical research method; experimental research method; expert method

6. Main contributions of the thesis

- Proposing the medical image fusion method IEXDoG_COA (Improved Extended Difference-of-Gaussians – Coati Optimization Algorithm), which combines the edge-preserving filter WMCF (Weighted Mean Curvature Filter) for image decomposition, the ASR_COA algorithm, which uses COA optimization to adaptively fuse the base component, and the XDoG_ST_LE algorithm, based on an improved XDoG filter, to fuse the detail component, in order to improve brightness and contrast and enhance the preservation of edge and structural features of the source images. The results have been published in [P1], [P2], [P4].

- Proposing the medical image fusion method SNM (SSFRG + NRBOFA + MLRFA), which combines the SSFRG decomposition technique with the NRBOFA algorithm for base component fusion and the MLRFA algorithm for detail component fusion, in order to improve the quality of the fused image and enhance the preservation of important information from the source images. The results have been published in [P3].

7. Structure of the thesis

Besides the Introduction, Conclusions and List of references, the thesis is organized into three chapters:

Chapter 1. Overview and theoretical background.

Chapter 2. Medical image fusion method based on two-component edge-preserving decomposition.

Chapter 3. Medical image fusion method based on three-component decomposition and adaptive weights.

Chapter 1. OVERVIEW AND THEORETICAL BACKGROUND

1.1. Medical image fusion

Medical image fusion is a technique that integrates information from multiple images acquired by one or several diagnostic imaging modalities in order to produce an output image containing more information while preserving features important for diagnosis. The four commonly used imaging modalities are CT, MRI, PET and SPECT, of which CT and MRI are structural images, whereas PET and SPECT are functional images. Research focuses on fusing multimodal image pairs such as MRI–CT, MRI–PET and MRI–SPECT in order to effectively combine anatomical and functional information.

Medical image fusion has many applications, typically supporting the diagnosis of Alzheimer’s disease and tumor detection, segmentation and classification tasks. Combining structural and functional features at the same time improves the analytical capability of modern computer-aided diagnosis systems.

1.2. Related work on image fusion

Medical image fusion methods have developed mainly in two directions: transform-based methods and deep-learning-based methods.

1.2.1. Transform-based methods

The general idea is to decompose the source images in different representation domains, fuse the feature components, and reconstruct the image through the inverse transform. Decomposition methods are usually divided into three groups: multi-scale transforms (MST), including pyramid transforms, wavelets and multi-scale geometric analysis (NSCT, NSST); filter-based methods (two- or three-component decomposition); and sparse representation (SR)-based methods. The fusion of base and detail components is carried out with various strategies, ranging from simple rules (averaging, maximum) to energy-based methods, optimization, and biologically inspired neural models (PCNN and its variants).

1.2.2. Deep-learning-based methods

CNN, autoencoder, GAN and Transformer architectures are used to exploit information from multimodal image data. CNNs are strong at local features, Transformers are strong at global context, and GANs generate highly realistic images; hybrid CNN–Transformer models combine the advantages of several architectures. However, deep learning models usually require large-scale training data, have high computational cost, and are difficult to interpret. Therefore, methods based on image decomposition combined with optimization remain a promising research direction.

1.3. Limitations of some published methods

The survey reveals three key issues that require further research. First, decomposition methods based on mean filters and bilateral filters find it difficult to fully preserve geometric details and structural boundaries, while NSCT and NSST have high computational cost. Second, base component fusion strategies based on simple rules or energy are limited in maintaining brightness, contrast and structural consistency at the same time. Third, detail component fusion rules such as ABS-Max or PCNN variants do not preserve edge and structural information well.

Accordingly, the thesis builds and develops two medical image fusion models, IEXDoG_COA and SNM, which combine decomposition based on edge-preserving filters with adaptive fusion mechanisms based on optimization algorithms, structure tensors, multi-level morphological gradients and local energy.

1.4. Evaluation metrics

The evaluation metrics are divided into four main groups: (1) visual perception (the mean luminance index Q_{MLI} and the contrast index Q_{CI}); (2) information theory (the mutual information index Q_{MI} and the nonlinear correlation information entropy index Q_{NCIE}); (3) structural (the weighted structural preservation index Q_W); (4) feature-based (the edge information preservation index Q_G , the feature similarity index Q_{FSIM} and the phase-

congruency-based feature index Q_P). Most metrics take values in $[0, 1]$; the higher the value, the better the fused image preserves information and features from the source images.

1.5. Background knowledge

This chapter presents the background knowledge used to build the proposed models, including: two optimization algorithms, the Coati Optimization Algorithm (COA) and the Newton-Raphson-Based Optimizer (NRBO); the filters WMCF, XDoG, RoG, SSIF and RGF; and structures and energy functions such as the Structure Tensor (ST), Local Energy (LE), and the Multi-Level Morphological Gradient (MLMG) operator.

1.6. Experimental datasets

The experiments were carried out on three brain image datasets collected from the Harvard Whole Brain Atlas (AANLIB): the Normal Anatomy with MRI/PET set (normal brain, three planes T, S, C), the Glioma MRI-SPECT set (glioma), and the Alzheimer's disease dataset (MRI-SPECT). All images are co-registered, ensuring spatial correspondence between modalities.

1.7. Conclusion of Chapter 1

Chapter 1 presents an overview of the multimodal medical image fusion problem, classifies the groups of methods and analyzes their strengths and weaknesses, and thereby identifies challenges and research directions. Image quality evaluation metrics, optimization algorithms, filters, structure tensors and energy functions are also introduced, providing the theoretical basis for the following chapters.

Chapter 2. MEDICAL IMAGE FUSION METHOD BASED ON TWO-COMPONENT EDGE-PRESERVING DECOMPOSITION

2.1. Introduction

Simultaneously preserving brightness, contrast, edge information and structure remains a challenge. For the base component, averaging or maximum rules easily degrade contrast or cause an imbalance of information. For the detail component, the ABS-Max rule finds it difficult to maintain details from low-intensity source images. On this basis, this chapter proposes the IEXDoG_COA model with the following contributions: the adaptive fusion algorithm ASR_COA for the base component; the XDoG_ST_LE algorithm for the detail component; and the IEXDoG_COA model integrating the two strategies above.

2.2. Medical image fusion model IEXDoG_COA

The IEXDoG_COA model is built on the principle of separately processing the different kinds of information in medical images, with an overall procedure consisting of four main stages (illustrated in Figure 2.1): preprocessing (Step 1 and Step 2), decomposition of the input images (Step 3), fusion of the components (Step 4) and reconstruction of the fused image (Step 5 and Step 6).

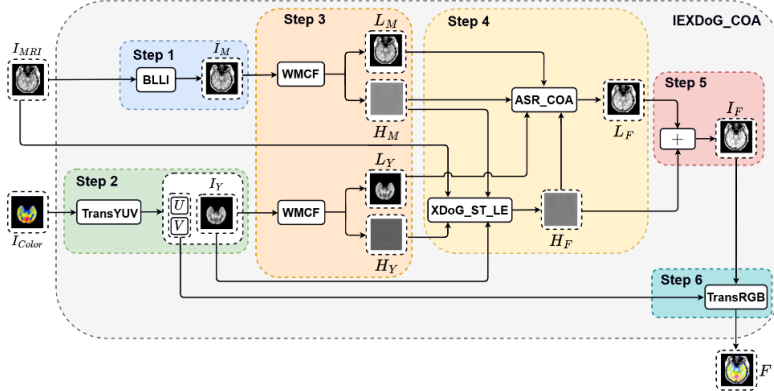


Figure 2.1. Block diagram of the model IEXDoG_COA

2.2.1. The WMCF decomposition algorithm

The base and detail components are determined by: $L_I = \text{WMCF}(I)$ and $H_I = I - L_I$

2.2.2. The XDoG_ST_LE detail component fusion algorithm

XDoG_ST_LE combines the XDoG filter (which detects salient edges), the structure tensor ST (which describes the orientation and saliency of local structures) and the local energy LE (which reflects the degree of intensity variation). The features are combined through the Hadamard product to build a decision map; a guided filter (GIF) is applied to smooth the map and ensure spatial continuity, limiting noise and artifacts.

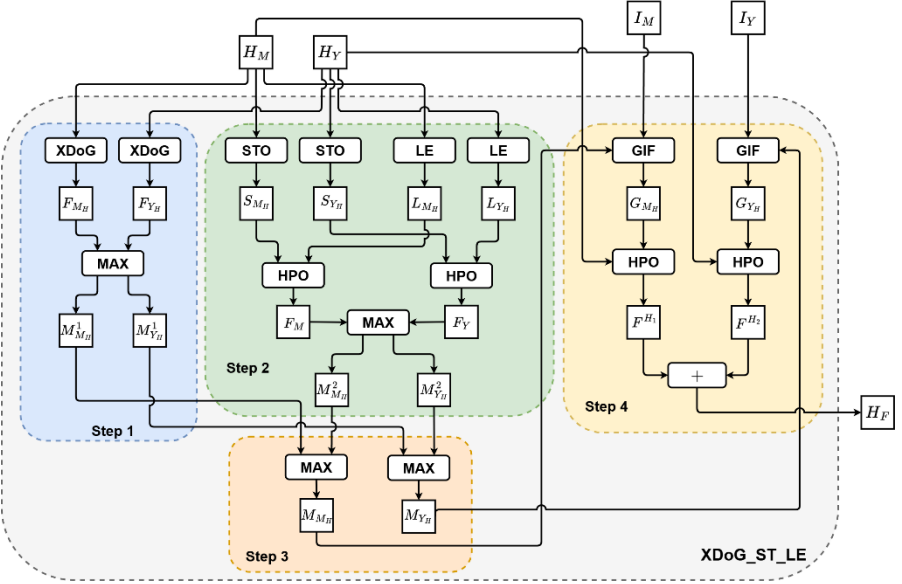


Figure 2.2. Flowchart of the detail component fusion algorithm

XDoG_ST_LE

2.2.3. The ASR_COA base component fusion algorithm

The fused base component is defined as a linear combination $L_F = \alpha \times L_M + \beta \times L_Y$ where L_M and L_Y are the base components of the MRI image and the PET/SPECT image after conversion to the grayscale channel; α and β are two parameters to be optimized within the ranges $\alpha \in [0.9, 0.999]$, $\beta \in [0.001, 0.3]$. ASR_COA uses COA to determine a suitable parameter set for each image pair.

The objective function CF is built by improving objective functions from previous studies, and simultaneously combines four criteria: information content (entropy), mean brightness, contrast (standard deviation) and edge information (Sobel):

$$CF(x) = \frac{\log(\log(E_{in}(x) + e) \times n_e(x)) \times (E_F(x) - E_M) \times C_F(x)}{H \times W \times B_F(x)}$$

where: E_M, E_F are the entropies of I_M and I_F ; B_F is the mean value of I_F ; H, W are the dimensions of I_M ; n_e is the number of edge points and E_{in} is the total edge strength (determined by Sobel); C_F is the contrast, $\exp(1)$ helps prevent the argument of the logarithm from being zero when the image has very few or no edge points. COA searches for (α, β) such that CF attains its maximum value.

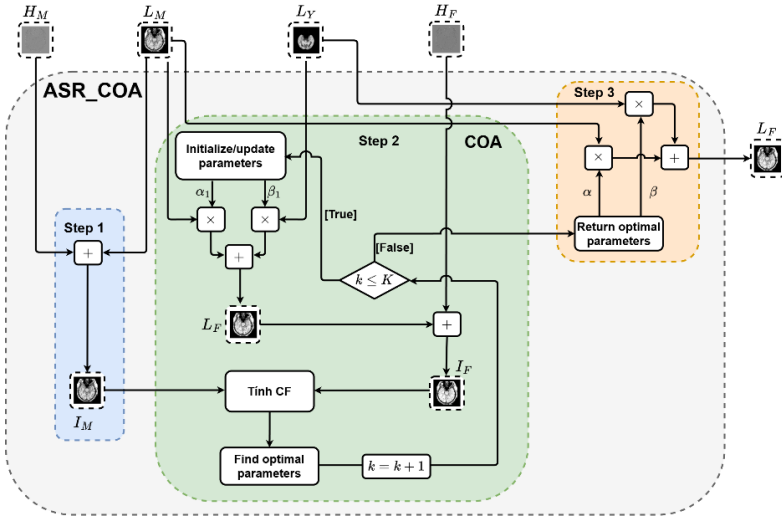


Figure 2.1. Flowchart of the base component fusion algorithm *ASR_COA*

2.2.4. Complexity of the IEXDoG_COA algorithm

The preprocessing stage, WMCF decomposition and detail fusion each have a complexity of $O(HW)$. With a population of NC individuals and R iterations, the complexity of *ASR_COA* is $O(NC \times R \times HW)$. Therefore, the complexity of *IEXDoG_COA* is $O(NC \times R \times HW)$; for a set of p image pairs, it is $O(p \times NC \times R \times HW)$.

2.3. Experimental setup and evaluation

A total of 90 MRI–PET image pairs of size 256×256 , taken from the Normal Anatomy set, were divided into the following subsets: D (3 pairs: 083T, 057S, 068C), D_T , D_S and D_C (each with 30 pairs along the T, S, C axes, respectively). IEXDoG_COA (denoted G_8) was compared with seven representative methods published in 2019–2024 ($G_1 - G_7$); COA was compared with six optimization algorithms: DA, GWO, MVO, ROA, SCA, WOA.

Experiment 1 - Effectiveness of COA

Table 2.6. Mean and standard deviation (Std) values

Method	Mean ($\times 10^{-5}$)	Std ($\times 10^{-5}$)	Mean ($\times 10^{-5}$)	Std ($\times 10^{-5}$)	Mean ($\times 10^{-5}$)	Std ($\times 10^{-5}$)
Dataset	83T		57S		68C	
COA	5.573	0.008	5.584	0.011	5.544	0.037
DA	5.568	0.012	5.583	0.015	5.418	0.079
GWO	5.555	0.010	5.566	0.020	5.403	0.040
MVO	5.517	0.038	5.515	0.051	5.247	0.051
ROA	5.537	0.027	5.552	0.029	5.348	0.086
SCA	5.526	0.023	5.528	0.036	5.296	0.062
WOA	5.542	0.047	5.561	0.028	5.434	0.065

Table 2.6 presents the mean value and standard deviation of CF over 30 independent runs. COA achieves the highest mean value and the smallest standard deviation, demonstrating good convergence and high stability.

Experiment 2 - Effectiveness of WMCF

WMCF was replaced by four other edge-preserving filters (Table 2.7); WMCF achieves the best values in Q_W , Q_G , Q_{FSIM} and Q_P on all three datasets, confirming the experimental basis for selecting WMCF.

Experiment 3 - Effectiveness of ASR_COA

ASR_COA was replaced by the ABS-Max, INSML+WLE, CNP, nuclear-norm and averaging rules ($T_1 - T_5$), evaluated with Q_{MLI} and Q_{CI} . T_5 (corresponding to the averaging rule) gives the lowest values; G_8 achieves the best results on all of D_T , D_S and D_C (Table 2.8), confirming the role of ASR_COA in improving brightness and contrast.

Experiment 4 - Effectiveness of XDoG_ST_LE

XDoG_ST_LE was replaced by ABS-Max, INSML, SR, local energy and structure tensor ($C_1 - C_5$), evaluated with Q_W , Q_G , Q_{FSIM} and Q_P . G_8 achieves the highest values on all metrics and all three datasets (Table 2.9), demonstrating superior preservation of details and structure.

Table 2.1. Results of Experiment 2 (replacing the WMCF filter)

Dataset	Model	Q_{MLI}	Q_{CI}	Q_W	Q_G	Q_{FSIM}	Q_P
D_T	W_1	0.36759	0.39883	0.85382	0.69055	0.91621	0.51960
	W_2	0.36528	0.39709	0.86200	0.69615	0.92664	0.53360
	W_3	0.31572	0.34947	0.88010	0.71510	0.95996	0.58239
	W_4	0.30772	0.33715	0.87497	0.71218	0.96792	0.57756
	G_8	0.35802	0.39144	0.88170	0.72528	0.96905	0.65548
D_S	W_1	0.40143	0.36883	0.86787	0.71750	0.93746	0.57675
	W_2	0.40608	0.37174	0.87281	0.71942	0.94173	0.57804
	W_3	0.37665	0.34930	0.88352	0.73267	0.95880	0.60859
	W_4	0.37376	0.34383	0.88057	0.72700	0.96197	0.60553
	G_8	0.40242	0.36827	0.88743	0.75346	0.97995	0.68973
D_C	W_1	0.35532	0.37043	0.90109	0.75448	0.94244	0.63305
	W_2	0.34909	0.36350	0.90560	0.75891	0.95004	0.64227
	W_3	0.31892	0.33493	0.91548	0.77294	0.97061	0.68190
	W_4	0.32201	0.33410	0.91371	0.76761	0.97267	0.67717
	G_8	0.35574	0.36937	0.92164	0.79170	0.98423	0.74619

Table 2.8. Results of Experiment 3 (replacing ASR_COA)

Model	Q_{MLI}	Q_{CI}	Q_{MLI}	Q_{CI}	Q_{MLI}	Q_{CI}
Data	D_T		D_S		D_C	
T_1	0.31603	0.35002	0.35801	0.33488	0.32075	0.34011
T_2	0.31148	0.34592	0.35408	0.33162	0.31755	0.33733
T_3	0.29149	0.32848	0.33153	0.31322	0.29881	0.32150
T_4	0.31417	0.34882	0.35633	0.33417	0.31810	0.33912
T_5	0.27571	0.32640	0.29231	0.31014	0.26725	0.31313
G_8	0.35802	0.39144	0.40242	0.36827	0.35574	0.36937

Experiment 5 - Overall evaluation

Table 2.10 shows that α is consistently high and stable, while β is considerably smaller. This indicates that the base component of the MRI image plays the dominant role in low-frequency-domain fusion. The stability of α and β across different datasets confirms the consistency of the optimization mechanism in ASR_COA.

Table 2.9. *Results of Experiment 4 (replacing XDoG ST LE)*

Data	Algorithm	Q_W	Q_G	Q_{FSIM}	Q_P
D_T	C_1	0.80013	0.67751	0.96292	0.57018
	C_2	0.08734	0.20572	0.60356	0.14056
	C_3	0.83628	0.66895	0.95509	0.52030
	C_4	0.00166	0.03403	0.44915	0.08430
	C_5	0.72306	0.59954	0.92194	0.40283
	G_8	0.88170	0.72528	0.96905	0.65548
D_S	C_1	0.81484	0.68789	0.96491	0.57382
	C_2	0.12652	0.23222	0.60045	0.15825
	C_3	0.81208	0.65431	0.94954	0.48917
	C_4	0.00196	0.01746	0.36795	0.17783
	C_5	0.74912	0.60171	0.92951	0.40515
	G_8	0.88743	0.75346	0.97995	0.68973
D_C	C_1	0.81231	0.70741	0.96344	0.62209
	C_2	0.05018	0.15460	0.49013	0.12784
	C_3	0.88224	0.71912	0.95535	0.58716
	C_4	0.00090	0.00351	0.31062	0.29702
	C_5	0.67710	0.57089	0.90654	0.36971
	G_8	0.92164	0.79170	0.98423	0.74619

Table 2.10. *Mean values of the optimal parameters α and β*

No.	Data	α	β
1	D_T	0.99358	0.22262
2	D_S	0.99661	0.20913
3	D_C	0.99380	0.17241

Figure 2.6 shows that G_8 produces fused images with markedly higher brightness and contrast than the remaining models, whereas $G_4 - G_6$ yield

darker images. Figure 2.7 (zoomed-in regions) further demonstrates the structure preservation and feature enhancement capability of G_8 .

Table 2.11. *Results of Experiment 5.*

DL	No.	Q_{MLI}	Q_{CI}	Q_W	Q_G	Q_{FSIM}	Q_P
D_T	G_1	0.31612	0.35352	0.87461	0.69883	0.91902	0.52821
	G_2	0.30922	0.34452	0.86114	0.70447	0.94706	0.58701
	G_3	0.32141	0.32573	0.79109	0.51351	0.91053	0.36329
	G_4	0.30029	0.33691	0.88046	0.69627	0.93300	0.52996
	G_5	0.27012	0.32319	0.76661	0.63623	0.86708	0.39640
	G_6	0.29009	0.33124	0.87650	0.69137	0.92085	0.48694
	G_7	0.31139	0.34641	0.83508	0.69574	0.93421	0.50906
	G_8	0.35802	0.39144	0.88170	0.72528	0.96905	0.65548
D_S	G_1	0.35758	0.33787	0.87807	0.72783	0.93633	0.59199
	G_2	0.35184	0.33058	0.86305	0.73760	0.95496	0.63104
	G_3	0.36022	0.31377	0.79231	0.55566	0.91891	0.45258
	G_4	0.34181	0.32277	0.88065	0.72026	0.94145	0.56889
	G_5	0.28748	0.30880	0.73235	0.63448	0.87186	0.45026
	G_6	0.33045	0.31566	0.87671	0.71459	0.93094	0.54906
	G_7	0.35376	0.33205	0.83926	0.73085	0.94039	0.57085
	G_8	0.40242	0.36827	0.88743	0.75346	0.97995	0.68973
D_C	G_1	0.32067	0.34338	0.91123	0.76085	0.94556	0.64527
	G_2	0.31425	0.33513	0.90991	0.77269	0.96756	0.68857
	G_3	0.32546	0.31710	0.85383	0.60094	0.93577	0.55070
	G_4	0.30623	0.32799	0.91396	0.76041	0.95029	0.62159
	G_5	0.26082	0.30937	0.79611	0.68186	0.88633	0.47291
	G_6	0.29809	0.32355	0.90724	0.75096	0.93969	0.59308
	G_7	0.30488	0.32541	0.83233	0.69591	0.93248	0.50502
	G_8	0.35574	0.36937	0.92164	0.79170	0.98423	0.74619

Table 2.11 shows that model G_8 (IEXDoG_COA) achieves the highest values on all metrics Q_{MLI} , Q_{CI} , Q_W , Q_G , Q_{FSIM} and Q_P on all three subsets D_T , D_S and D_C . For the visual perception metrics (Q_{MLI} , Q_{CI}), G_8 shows a marked improvement over the remaining methods, with a gain of approximately 7–12%, indicating its ability to enhance brightness and

contrast. For the structural metric Q_W , G_8 maintains high and stable values, reflecting good preservation of edges and structure. For the feature metrics (Q_G , Q_{FSIM} and Q_P), the proposed model continues to deliver superior results, particularly for Q_P with a substantial improvement (approximately 8–12%), indicating its ability to preserve important information from the input images.

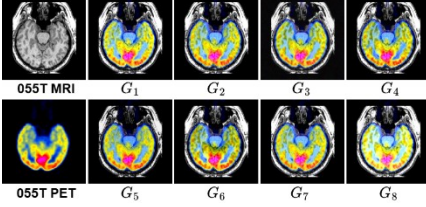


Figure 2.6. *Fused images*

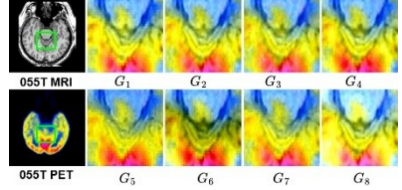


Figure 2.7. *Regions extracted from Figure 2.6*

Computational efficiency

The mean execution time of G_8 (IEXDoG_COA) is approximately 7.27–7.37 seconds, ranking 6th out of 8. Although it is not the fastest, G_8 achieves a reasonable balance between computational cost and image quality, and is considerably lower than methods based on NSCT ($G_4 \approx 13$ –14 seconds) and SR ($G_7 \approx 10$ seconds).

The Friedman test combined with the Wilcoxon Signed-Rank test (Tables 2.12, 2.13) on D_T , D_S and D_C shows that 125 of 126 comparisons are statistically significant ($p \ll 0.001$), confirming the consistent superiority of IEXDoG_COA.

2.4. Conclusion of Chapter 2

Chapter 2 proposes the IEXDoG_COA model, which uses WMCF for image decomposition, ASR_COA for the base component and XDoG_ST_LE for the detail component. Experimental results on D_T , D_S and D_C show that G_8 achieves superior performance, markedly improving brightness, contrast and feature preservation while maintaining stability. However, the two-component decomposition mechanism, the objective function without constraint terms, and detail processing at a single decomposition level do not fully represent the different levels of detail, which motivates the development of the SNM model in Chapter 3.

Chapter 3. MEDICAL IMAGE FUSION METHOD BASED ON THREE-COMPONENT DECOMPOSITION AND ADAPTIVE WEIGHTS

3.1. Introduction

Decomposition methods that use mean filters or bilateral filters do not preserve background information well and easily cause blurring and loss of detail. Meanwhile, NSST and NSCT have high computational cost. In the base component fusion step, rules based on energy or the improved Roberts operator often reduce the brightness and contrast of the fused image. For the detail component, methods based on PCNN and its variants do not preserve structural information well in complex or low-contrast regions and often lose edges.

Although IEXDoG_COA considerably improves brightness, contrast and edge preservation, the two-component decomposition mechanism of WMCF does not fully represent the different levels of detail in an image; the detail fusion process does not sufficiently consider the relationships between levels; and COA increases cost as the number of iterations grows. On this basis, this chapter proposes the SNM model, consisting of three algorithms: the SSFRG decomposition, the NRBOFA base component fusion and the MLRFA detail component fusion.

The main contributions are: (1) the SSFRG decomposition algorithm, combining the SSIF and RoG filters; (2) the NRBOFA algorithm, using the NRBO optimizer; (3) the MLRFA algorithm, based on MLMG, LE and the RGF filter; (4) the integration into the SNM model and its validation on 111 MRI–PET and MRI–SPECT image pairs.

3.2. Medical image fusion model SNM

Unlike IEXDoG_COA, SNM uses a three-component decomposition mechanism to clearly separate background information, important structural details and weak texture components. Procedure: the MRI and PET/SPECT images are decomposed by SSFRG into one base component and two detail components; the detail components are fused by MLRFA; NRBOFA

determines the optimal weights for combining the base components; finally, the fused image is reconstructed.

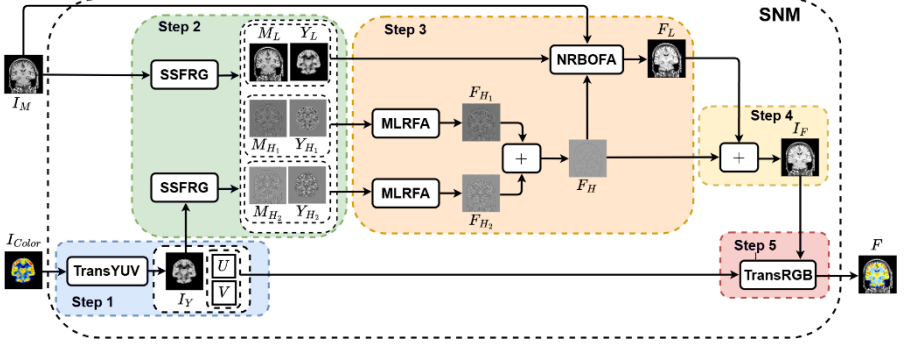


Figure 3.1. The SNM algorithm

3.2.1. The SSFRG decomposition algorithm

SSFRG combines the SSIF and RoG filters and decomposes the input image into three components: I^L (the base component obtained from SSIF), $I^{H_1} = I - I^R$ (the first detail component, reflecting weak textures/noise) and $I^{H_2} = I^R - I^L$ (the second detail component, retaining prominent structural boundaries), where I^R is the RoG filtering result.

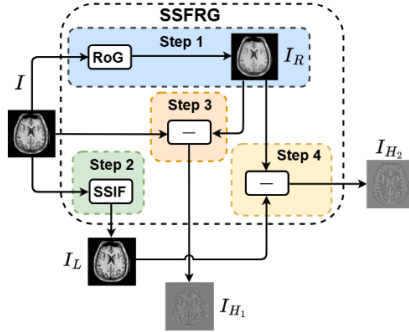


Figure 3.2. SSFRG decomposition algorithm

3.2.2. The MLRFA detail fusion algorithm

MLRFA combines the multi-level morphological gradient MLMG (which exploits edges at multiple scales), the local energy LE (which evaluates the degree of intensity variation) and the rolling guidance filter RGF. The edge and energy features are combined through the Hadamard

product; RGF removes small oscillations while preserving important boundaries; a decision map is built with the maximum rule to form the fused detail component.

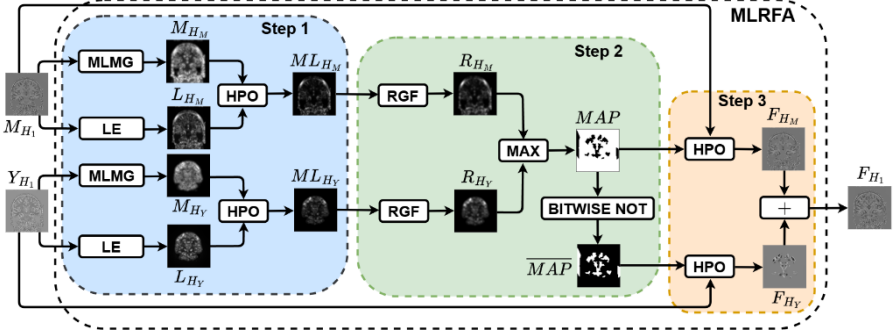


Figure 3.3. *MLRFA fusion algorithm*

3.2.3. The NRBOFA algorithm

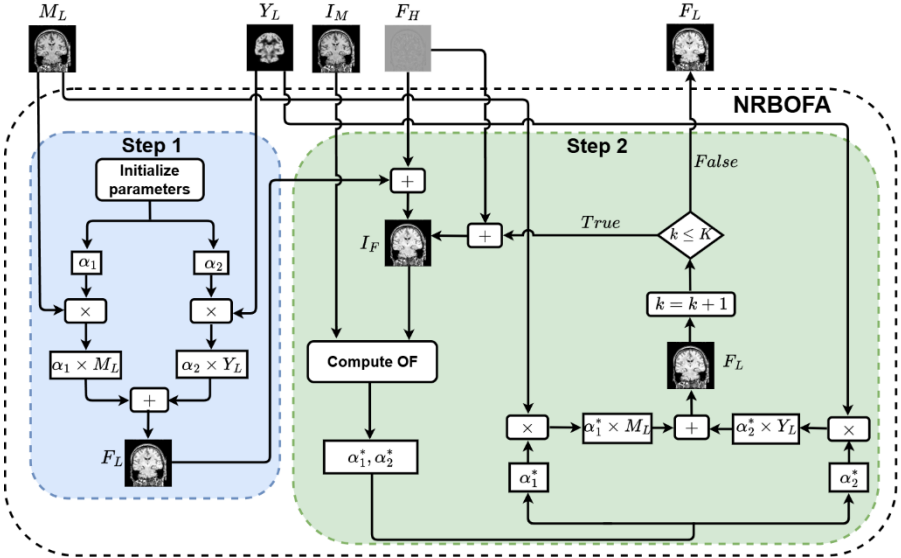


Figure 3.4. *The NRBOFA algorithm*

The temporary base component is defined as $F_L = \alpha_1 \times M_L + \alpha_2 \times Y_L$, with $\alpha_1 \in [0.9, 0.999]$, $\alpha_2 \in [0.001, 0.3]$. Unlike ASR_COA, NRBOFA builds a multi-criteria objective function with a penalty term to constrain and balance brightness, contrast, structure and information content:

$$OF(x) = \frac{f_B(x) \times f_C(x) \times P(x) \times SSIM(I_F(x), I_M) \times e^{E_F(x) - E_M} \times \sqrt[4]{E_{in}(x)}}{H \times W \times B_F(x)}$$

where f_B, f_C respectively evaluate the fidelity of brightness and contrast with respect to the MRI image; P is a penalty term based on the variance of the three normalized features (brightness, contrast, entropy); SSIM is the structural similarity; E_{in} is the total edge intensity (Sobel). The hyperparameters are: $\sigma_B = 0.06$, $\beta = 0.055$, $\rho = 2.2$. NRBO searches for (α_1, α_2) that optimize OF .

3.2.4. Complexity of the SNM algorithm

SSFRG and MLRFA have a complexity of $O(HW)$ (with the window parameters being constants); NRBOFA has a complexity of $O(K \times (N \times d + N \times H \times W))$. Overall, the complexity of SNM is $O(K \times N \times H \times W)$; for a set of p image pairs, it is $O(p \times K \times N \times H \times W)$.

3.3. Experimental setup and evaluation

Table 3.6. *Values obtained from 30 runs on D_1*

Data	59T		68S		72C		28S		36S	
Algorithm	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
	$\times 10^{-6}$	$\times 10^{-7}$	$\times 10^{-5}$	$\times 10^{-6}$	$\times 10^{-4}$	$\times 10^{-7}$	$\times 10^{-4}$	$\times 10^{-6}$	$\times 10^{-4}$	$\times 10^{-8}$
GWO	9.854	2.96	5.245	1.41	2.1681	1.09	4.762	3.11	8.4229	3.52
MVO	9.624	4.19	5.265	1.58	2.1666	1.63	4.710	8.69	8.4225	3.88
SCA	9.212	3.79	5.128	2.16	2.1687	0.79	4.720	6.09	8.3959	178.24
TSA	9.796	3.54	5.321	1.05	2.1675	2.52	4.603	13.72	8.4056	114.22
ROA	9.563	3.54	5.210	2.15	2.1678	1.07	4.757	5.80	8.4228	3.60
GJO	9.536	4.25	5.119	1.33	2.1674	1.42	4.639	11.67	8.4125	129.97
WSO	9.772	3.73	5.197	1.50	2.1689	0.71	4.650	8.79	8.3973	159.56
COA	9.994	4.54	5.309	1.49	2.1676	2.63	4.777	2.28	8.4207	28.31
NRBO	10.195	1.55	5.373	0.10	2.1698	0.48	4.783	1.82	8.4229	1.67

Experiments were carried out on 111 MRI–PET image pairs (normal brain), 36 MRI–SPECT pairs (glioma) and 27 MRI–SPECT pairs (Alzheimer’s disease), of size 256×256 , divided into four groups $D_1 - D_4$. SNM is compared with 11 representative methods published in 2021–2025 ($A_1 - A_{11}$); NRBO is compared with eight optimization algorithms (GWO, MVO, SCA, TSA, ROA, GJO, WSO, COA). Main parameters: SSIF with

$r = 11, \theta = 0.01, k = 1.5, s = 1$; RoG with $\delta_1 = 1, \delta_2 = 2, \beta = 0.01, R = 3$; NRBOFA $N = 50, d = 2, K = 50$; MLRFA $n = 5, LE w = 3, RGF \sigma_s = 3, \sigma_r = 0.05, T = 1$. The deep learning models were trained on the PyTorch platform with an NVIDIA A100 GPU for 100 epochs, with a learning rate of 0.005, using the Adam optimizer. The remaining experiments were run in MATLAB R2022b on an Intel® Core™ i5-12400 CPU with 16 GB of RAM.

Experiment 1 - Effectiveness of NRBO

Table 3.6 shows that NRBO achieves the highest Mean and the smallest Std of the objective function OF , outperforming the eight compared algorithms and confirming that selecting NRBO as the core optimization component in SNM is reasonable.

Experiment 2 - Effectiveness of NRBOFA

Table 3.7. Results of Experiment 2 (replacing NRBOFA).

Dataset	D_2		D_3		D_4	
Algorithm	Q_{MLI}	Q_{CI}	Q_{MLI}	Q_{CI}	Q_{MLI}	Q_{CI}
L_1	0.23167	0.30759	0.08404	0.18183	0.17772	0.25685
L_2	0.28217	0.32118	0.11991	0.20497	0.24267	0.31181
L_3	0.30320	0.32140	0.13751	0.22681	0.22204	0.28928
L_4	0.31451	0.34761	0.13473	0.22730	0.24243	0.31256
L_5	0.32713	0.34181	0.14212	0.22829	0.24761	0.31059
L_6	0.32799	0.34328	0.14124	0.22866	0.24830	0.31250
L_7	0.32993	0.34427	0.14277	0.22887	0.25187	0.31221
L_8	0.33184	0.34361	0.14378	0.22865	0.25394	0.31150
SNM	0.33529	0.35602	0.14723	0.23786	0.27311	0.33545

NRBOFA was replaced by eight base component fusion strategies ($L_1 - L_8$), evaluated with Q_{MLI} and Q_{CI} on D_2 , D_3 and D_4 . SNM achieves the highest values on all three subsets; compared with L_1 , the improvement reaches 44.7–75.2% in terms of Q_{MLI} . L_1 (improved Roberts) and L_2 (energy-based) give the lowest results, reflecting the limitations of traditional rules.

Experiment 3 - Effectiveness of MLRFA

MLRFA was replaced by six detail fusion strategies (H_1 – H_6), evaluated with Q_{MI} , Q_{NCIE} , Q_G and Q_{FSIM} on D_2 , D_3 and D_4 . SNM achieves the highest values on most metrics; only Q_{FSIM} on D_3 and Q_G on D_3 are approximately equal to the best method. Compared with H_1 , the improvement in Q_G reaches 82.9–157.8%, confirming the effectiveness and stability of MLRFA.

Table 3.8. *Results of Experiment 3 (replacing MLRFA)*

Data	Algorithm	Q_{MI}	Q_{NCIE}	Q_G	Q_{FSIM}
D_2	H_1	2.55021	0.80611	0.40478	0.76284
	H_2	4.24714	0.81350	0.73070	0.96281
	H_3	4.39664	0.81433	0.72929	0.96320
	H_4	4.43955	0.81445	0.72693	0.96156
	H_5	4.40478	0.81444	0.72528	0.95826
	H_6	4.44284	0.81447	0.73448	0.96576
	SNM	4.54896	0.81540	0.74016	0.96978
D_3	H_1	2.04558	0.80458	0.27792	0.78894
	H_2	2.76597	0.80703	0.70807	0.97416
	H_3	3.11460	0.80768	0.70918	0.97670
	H_4	3.07267	0.80755	0.70291	0.97523
	H_5	3.06278	0.80765	0.70804	0.97399
	H_6	3.09734	0.80767	0.71086	0.97895
	SNM	3.17790	0.80788	0.71657	0.97893
D_4	H_1	2.61402	0.80629	0.38172	0.77298
	H_2	3.71181	0.80997	0.77596	0.97753
	H_3	4.08485	0.81104	0.75189	0.96893
	H_4	4.12495	0.81115	0.75537	0.97053
	H_5	4.30318	0.81175	0.77884	0.97706
	H_6	4.11184	0.81113	0.76343	0.97289
	SNM	4.42834	0.81238	0.78517	0.98170

Experiment 4 - Overall evaluation

The comparison was carried out on the subsets D_2 , D_3 and D_4 using representative image pairs and zoomed-in regions to clarify features such as

edges, brightness and contrast. The results show that A_2 and A_6 tend to reduce brightness and contrast, whereas A_4 , A_{10} , A_{11} improve these aspects better. The SNM model shows a clear advantage by simultaneously enhancing brightness and contrast across the datasets.

Table 3.9 shows that the parameter α_1^* is always much larger than α_2^* , reflecting the dominant role of the MRI base component in the fusion process.

Table 3.9. Mean values of the optimal parameters α_1^* and α_2^*

No.	Dataset	α_1^*	α_2^*
1	D_2	0.97878	0.25723
2	D_3	0.99311	0.28609
3	D_4	0.96687	0.25970

Table 3.10. Results of SNM and the compared models on D_3

Data	Algorithm	Q_{MLI}	Q_{CI}	Q_{MI}	Q_{NCIE}	Q_G	Q_{FSIM}
D_3	A_1	0.13631	0.22238	2.58248	0.80614	0.70514	0.94898
	A_2	0.11641	0.19926	2.56885	0.80606	0.69345	0.95696
	A_3	0.13473	0.22294	2.19712	0.80475	0.67492	0.93187
	A_4	0.14189	0.22838	2.90465	0.80675	0.71927	0.94910
	A_5	0.13749	0.22428	2.56813	0.80612	0.69807	0.95218
	A_6	0.08933	0.17129	1.46387	0.80311	0.55816	0.88729
	A_7	0.13366	0.22020	2.14506	0.80463	0.67434	0.93030
	A_8	0.13577	0.22511	2.23972	0.80500	0.67997	0.93648
	A_9	0.13673	0.22491	2.43918	0.80556	0.68753	0.93904
	A_{10}	0.13428	0.21933	2.64994	0.80613	0.71554	0.96600
	A_{11}	0.14462	0.18928	2.74869	0.80635	0.72228	0.94575
	SNM	0.14723	0.23786	3.17790	0.80788	0.71657	0.97893

The results in Table 3.10 show that SNM achieves superior and stable performance on all of D_2 , D_3 and D_4 in terms of Q_{MLI} , Q_{CI} , Q_{MI} , Q_{NCIE} and Q_{FSIM} ; for Q_G , SNM ranks second (D_2), third (D_3) and is close to the best (D_4), indicating some room for improvement in local edge preservation. Overall, SNM strikes a good balance between enhanced brightness, information preservation and perceptual quality.

The Friedman test and the Wilcoxon Signed-Rank test (with Holm–Bonferroni correction) on the three datasets D_2 , D_3 and D_4 across six evaluation metrics confirm that the differences among the 12 algorithms are highly statistically significant ($p \ll 0.001$). The results show that 188 of 198 comparisons (94.9%) are statistically significant, of which 181 favor SNM, confirming that the improvement of SNM is stable and not a random difference between experimental runs.

Computational efficiency

Table 3.14. *Parameter sensitivity results (N , K) of NRBOFA (bold rows are the original configuration used for Table 3.10)*

Data	N	K	Q_{MLI}	Q_{CI}	Q_{MI}	Q_{NCIE}	Q_G	Q_{FSIM}	Time (s)
D_2	50	50	0.33529	0.35602	4.54896	0.81540	0.74016	0.96978	8.623
	30	50	0.32850	0.34903	4.53773	0.81524	0.73995	0.97053	5.173
	50	30	0.33044	0.35120	4.51806	0.81509	0.73869	0.96898	5.189
	30	30	0.32963	0.35016	4.52910	0.81520	0.73964	0.97013	3.447
D_3	50	50	0.14723	0.23786	3.17790	0.80788	0.71657	0.97893	8.167
	30	50	0.14715	0.23773	3.17511	0.80786	0.71623	0.97877	5.665
	50	30	0.14697	0.23741	3.17440	0.80785	0.71612	0.97865	5.484
	30	30	0.14715	0.23766	3.17251	0.80784	0.71602	0.97867	3.498
D_4	50	50	0.27311	0.33545	4.42834	0.81228	0.78517	0.98170	8.045
	30	50	0.27159	0.33377	4.42576	0.81225	0.78543	0.98183	5.585
	50	30	0.27267	0.33493	4.42780	0.81227	0.78525	0.98157	5.629
	30	30	0.27532	0.33779	4.37214	0.81199	0.78095	0.97963	3.573

The mean execution time of SNM is approximately 8.04–8.62 seconds, ranking 8th out of 12. The cost mainly comes from the weight optimization in NRBOFA and the MLMG operator. Although it is not the fastest, SNM still achieves top performance on the evaluation metrics, showing a reasonable balance between cost and quality.

Parameter sensitivity analysis

The sensitivity analysis of the parameters N and K of NRBOFA (Tables 3.14, 3.15) shows that when they are reduced from 50 to 30, image quality varies by less than 2% while processing time decreases by 30–60%, confirming the stability of the model.

Comparison and discussion of the two proposed models (Tables 3.16 and 3.17)

On the brightness and contrast metrics (Q_{MLI} and Q_{CI}), SNM gives better results on 5 of 6 datasets thanks to NRBOFA, whose objective function includes a constraint mechanism. Conversely, on the mutual information metrics (Q_{MI} and Q_{NCIE}) and the edge and structure preservation metrics (Q_W , Q_G , Q_{FSIM} and Q_P), IEXDoG_COA shows an advantage thanks to the XDoG_ST_LE algorithm, which is dedicated to the detail component. SNM is a successor and extension of IEXDoG_COA; in practice, the appropriate model can be chosen according to the diagnostic goal: IEXDoG_COA when preserving edge and structural features is the priority, and SNM when brightness and contrast are the priority.

Table 3.16. Results of IEXDoG and SNM on D_T , D_S and D_C

DL	Algorithm	Q_{MLI}	Q_{CI}	Q_W	Q_G	Q_{FSIM}	Q_P
D_T	IEXDoG_COA	0.35802	0.39144	0.88170	0.72528	0.96905	0.65548
	SNM	0.36362	0.39768	0.87546	0.71389	0.95214	0.60606
D_S	IEXDoG_COA	0.40242	0.36827	0.88743	0.75346	0.97995	0.68973
	SNM	0.40260	0.36997	0.88164	0.74861	0.97094	0.66348
D_C	IEXDoG_COA	0.35574	0.36937	0.92164	0.79170	0.98423	0.74619
	SNM	0.36036	0.37498	0.91861	0.78674	0.97442	0.71892

Table 3.17. Results of IEXDoG and SNM on D_2 , D_3 and D_4

DL	Algorithm	Q_{MLI}	Q_{CI}	Q_{MI}	Q_{NCIE}	Q_G	Q_{FSIM}
D_2	IEXDoG_COA	0.33212	0.35166	4.29338	0.81344	0.74164	0.97722
	SNM	0.33529	0.35602	4.54896	0.81500	0.74016	0.96978
D_3	IEXDoG_COA	0.14254	0.23174	3.28270	0.80831	0.72832	0.98481
	SNM	0.14723	0.23786	3.17790	0.80788	0.71657	0.97893
D_4	IEXDoG_COA	0.27345	0.33702	4.46820	0.81252	0.79178	0.98666
	SNM	0.27311	0.33545	4.42834	0.81228	0.78517	0.98170

3.4. Conclusion of Chapter 3

Chapter 3 proposes the SNM model consisting of three algorithms: SSFRG, NRBOFA and MLRFA. Experimental results show that SNM achieves the best values in Q_{MLI} , Q_{CI} , Q_{MI} and Q_{NCIE} , while maintaining

competitive results in Q_G and Q_{FSIM} , and preserves well the anatomical information from MRI and the functional information from PET/SPECT. The model still has limitations in computational cost and in preserving some fine edge features under low-contrast conditions.

CONCLUSIONS AND RECOMMENDATIONS

The thesis studies the multimodal medical image fusion problem in order to improve the quality of fused images and preserve important features from the source images. The main results are as follows:

First, the IEXDoG_COA model is proposed: it uses WMCF for image decomposition, ASR_COA (based on COA) for the base component and XDoG_ST_LE (combining XDoG, ST and LE) for the detail component, improving both visual quality and information preservation.

Second, the SNM model, consisting of SSFRG, NRBOFA and MLRFA, is proposed, which simultaneously exploits intensity, structure and local energy information at multiple levels to improve the quality of the fused image.

Third, the models are validated on multiple MRI–PET and MRI–SPECT datasets (including glioma and Alzheimer’s disease). SNM achieves the best results in Q_{MLI} and Q_{CI} , while IEXDoG_COA shows an advantage in edge preservation.

Future research directions: studying new decomposition techniques and fusion strategies; developing more effective objective functions and optimization algorithms; combining deep learning models (CNN, Transformer, hybrid architectures) with traditional image decomposition; extending the evaluation to large-scale CT–MRI and PET–CT datasets; and building fusion models that operate in real time.

LIST OF PUBLISHED PAPERS RELATED TO THE THESIS

1. **T. H. H. Le**, P. H. Dinh*, V. H. Vu, N. L. Giang, 2024, “A new approach to medical image fusion based on the improved Extended difference - of - Gaussians combined with the Coati optimization algorithm,” *Biomedical Signal Processing and Control*, vol. 93, pp. 106175. (SCIE, Q1, IF = 4.9).
2. P. H. Dinh*, **T. H. H. Le**, N. L. Giang, 2025, “An effective medical image fusion method utilizing moth - flame optimization and coupled neural P systems”, *Neural Computing and Applications*, vol. 37, pp. 10037–10077. (Scopus, Q1, IF=4.5).
3. H. L. Nguyen, **T. H. H. Le***, H. L. Nguyen, P. H. Dinh, N. L. Giang, 2026, “SNM: A medical image fusion method using SSFRG decomposition, NRBOFA base fusion, and multi - level morphological gradient”, *Biomedical Signal Processing and Control*, vol. 113, pp. 109080. (SCIE, Q1, IF = 4.9).
4. **Le Thi Hong Ha***, Dinh Phu Hung, Vu Van Hieu, Nguyen Long Giang, 2024, “Medical image fusion based on PCNN and the MFO algorithm”, *Proceedings of the 27th National Conference: Selected Issues of Information and Communication Technology*, Science and Technology Publishing House, pp. 641–646. (in Vietnamese)