Journal of Computing and Data Technology
Vol. 01, No. 01, June 2025, pp. 59-67.
DOI: https://doi.org/10.71426/jcdt.v1.i1.pp59-67

Homepage: https://review.journal-of-modern-technology.com/index.php/jcdt/index

Enhanced Image Fusion through Multi-Scale Adaptive Weighting and Post-Fusion Optimization

Sreeja Akuthota

Master of Science, Business Analytics, Clark University, 960 Main St Worcester MA 01610, USA.
Email: sreejaakuthota@gmail.com , sakuthota@clarku.edu , ORCID: https://orcid.org/0009-0003-4996-8733

Article Info

Article history:

Received: April 30, 2025
Revised: June 20, 2025
Accepted: June 25, 2025
First Online: June 30, 2025

Keywords:

Image fusion
Multi-scale decomposition
Post-fusion enhancement
Image processing
Perceptual quality

ABSTRACT

Image fusion plays a vital role in modern image processing by integrating complementary information from multiple source images into a single, enriched representation. This capability is critical in fields such as medical imaging, remote sensing, and surveillance. However, traditional fusion methods—such as pixel averaging and wavelet-based techniques—often struggle to preserve fine details or adapt to varied image content, leading to artifacts and degraded quality. Deep learning-based approaches offer improvements but require extensive datasets and high computational resources, limiting their use in real-time or resource-constrained environments. To address these limitations, this paper proposes a novel image fusion framework combining multi-scale adaptive weighting with post-fusion enhancement. The method utilizes multi- resolution decomposition to extract frequency components, assigning perceptual-based adaptive weights based on local salience and structural relevance. A dedicated enhancement stage further improves contrast, sharpness, and detail retention. Experimental results across diverse datasets show that the proposed method outperforms conventional techniques, achieving higher mutual information (2.85), structural similarity (0.92), and PSNR (34.6 dB), while maintaining superior visual quality. This framework provides an efficient and robust solution suitable for real-world deployment, advancing the state-of-the-art in image fusion.

1. INTRODUCTION

Existing image fusion techniques span a spectrum of methodologies, each with notable strengths and limitations. Early approaches, such as pixel-level averaging and principal component analysis (PCA), offer simplicity but frequently introduce distortions or fail to capture intricate details, compromising the fused output's utility. Multi-scale transform methods, such as those based on discrete wavelet transforms (DWT), marked a significant leap by decomposing images into frequency bands, yet their reliance on static fusion rules limits adaptability to varying image characteristics. More recently, deep learning frameworks have demonstrated promise in learning complex fusion patterns, but their dependence on extensive training datasets and high computational overhead renders them impractical for real-time or resource-limited scenarios, exposing a gap between theoretical advancements and practical deployment. In response to these challenges, this paper proposes an advanced image fusion framework that synergistically combines multi-scale weighting with post-fusion enhancement. By leveraging a multi-scale decomposition strategy, the method assigns adaptive weights to frequency component based on their local content and perceptual relevance, ensuring a tailored integration of information from source images. This is complemented by a novel post-fusion enhancement stage designed to refine the composite image, enhancing contrast, sharpness, and detail fidelity. The proposed approach aims to overcome the shortcomings of prior methods by delivering superior performance in both quantitative metrics and subjective visual quality, offering a versatile and efficient solution for real-world image fusion tasks. This work contributes to the ongoing evolution of fusion techniques, with potential to significantly impact applications requiring high quality image synthesis.

2. RELATED WORK

Image fusion has evolved significantly, with multi-scale decomposition emerging as a cornerstone for integrating information from diverse sources [1]. Early efforts, such as the adaptive weighted image fusion algorithm based on non-subsampled contourlet transform (NSCT) multi-scale decomposition, focused on enhancing fusion quality by incorporating infrared saliency maps and weighted averaging. Liu et al. demonstrated improved edge preservation and noise resilience in complex environments [7]. Similarly, Zhou et al. proposed an adaptive multi-weight fusion method using multi-scale transformation, designing weight matrices tailored to infrared and visible image characteristics to retain critical information [1]. While these approaches excel in specific contexts, their reliance on predefined weighting schemes limits adaptability to varying image content, and they often overlook post-fusion refinement, leaving room for visual quality improvements. Advancements in multi-scale feature integration have further enriched the field. Yang et al. introduced a multi-scale exposure fusion technique that measures contrast, saturation, and exposure, employing adaptive weighting and post- fusion optimization via decision maps and guided filtering [3]. This method enhances detail representation, addressing some limitations of earlier static approaches. In a similar vein, Hu et al. developed an improved multi-focus image fusion algorithm using a multi-scale weighted focus measure, reducing ghosting and blocking effects through pixel-level focus region extraction and guided filtering [5]. These studies highlight the potential of multi-scale weighting and post-fusion steps, yet their application-specific focus—exposure fusion and multi-focus imaging—restricts generalizability across broader fusion tasks.

Recent research has explored adaptive and feature- enhanced fusion frameworks. Luo and Hu proposed a multi-scale feature adaptive fusion (SAF) module for multi-task dense prediction, dynamically learning optimal feature scales [1]. While effective for dense prediction, this approach di- verges from traditional image fusion goals, lacking emphasis on post-fusion enhancement [11]-[12]. Li et al. introduced the Multi-scale Feature Enhanced Adaptive Fusion Network (MFEAFN) with a Focusing Selective Fusion Module (FSFM), leveraging attention weights and discrete cosine transforms for semantic segmentation [4]. Although inno- vative, its scope excludes post-fusion optimization, a critical aspect for visual quality. Gao et al. tackled medical image fusion with a multi-scale fusion global feature extraction network, integrating CNNs and Transformer-based modules [6], but their work prioritizes feature extraction over adaptive weighting and refinement, limiting its alignment with comprehensive fusion objectives. Hyperspectral and specialized fusion methods have also gained traction. Liu and Feng developed the Adaptive Multi- Scale Input Network (AMSIN) for hyperspectral image fu- sion, utilizing multi-scale spatial-spectral fusion blocks [9]. This approach enhances fusion quality but does not explicitly address adaptive weighting or post-fusion optimization. Similarly, Qiu et al. proposed the Multi-Scale Convolutional Feature Adaptive Weighting Fusion Network (MC-FAW) for detecting disorders of consciousness [8], yet its focus on medical diagnostics diverges from general image fusion challenges. Zhang et al. explored multi-scale feature fusion for image dehazing, employing error feedback and attention mechanisms [10], but its application to dehazing rather than fusion underscores a contextual mismatch with the current study's aims.

Despite these advancements, significant gaps persist in the literature. Many methods, such as those by Zhou et al. [2] and Liu et al. [7], excel in adaptive weighting but neglect post-fusion enhancement, potentially compromising final image clarity. Others, like Yang et al. [3] and Hu et al. [5], incorporate refinement but are tailored to specific domains, limiting versatility. Deep learning approaches, such as those by Gao et al. [6], offer robust feature extraction yet demand substantial computational resources, reducing practicality. The proposed work builds on these foundations by integrating multi-scale adaptive weighting with a novel post-fusion enhancement step, aiming to deliver a balanced, versatile solution that enhances both quantitative metrics and subjective quality across diverse applications. A flowchart of the work is shown in Figure 1.

3. METHODOLOGY

As shown in Figure 1, the proposed image fusion frame- work comprises three core stages: multi-scale decomposition, adaptive weighting-based fusion, and post-fusion enhancement. Each stage is carefully designed to balance computational efficiency with high-fidelity output, ensuring robustness across diverse image modalities. The methodology is mathematically formulated to promote reproducibility and facilitate theoretical validation. Initially, source images are decomposed into multi-scale frequency components using wavelet transforms. These components are then fused using perceptually driven weights based on local salience measures. Finally, the fused image is enhanced to boost contrast and structural clarity, producing visually superior results.

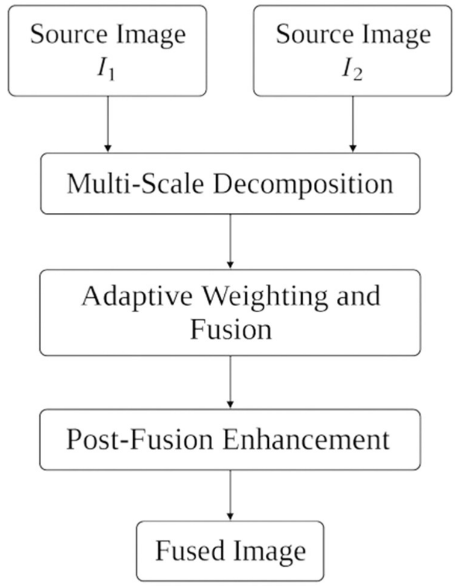

Figure 1. Flowchart of the proposed work.

3. METHODOLOGY

Figure 2 illustrates a comparative workflow between the proposed and conventional image enhancement approaches based on multi-space fusion. In the conventional method, the input underwater image is first transformed into the Lab color space, where weight maps such as Sobel edge weight, luminance contrast weight, and exposure weight are computed and combined to produce a normalized luminance weight map. This single luminance map is then used to enhance the image, resulting in a grayscale-like output with limited color restoration. In contrast, the proposed approach performs enhancement in both *Lab* and *RGB* spaces to preserve color fidelity and detail. The fusion process ensures that both luminance and color cues contribute to the final enhanced image, leading to improved visual contrast, sharpness, and color balance compared to the conventional method.

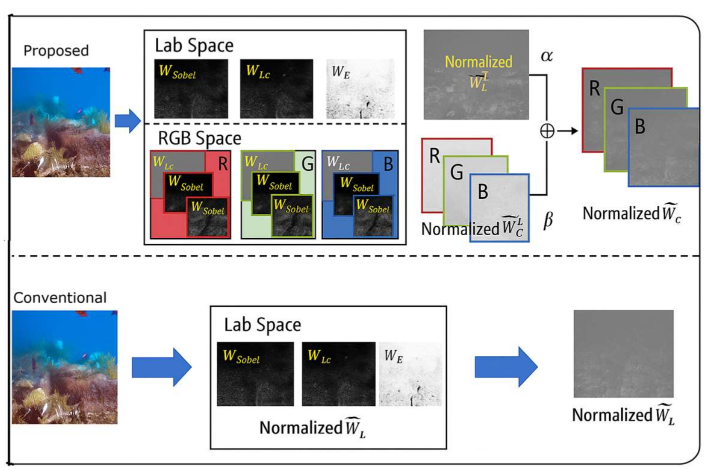

Figure 2. Block diagram of proposed and existing system of image fusion framework.

3.1 Multi-Scale decomposition

In the first stage, source images $I_1(x, y)$ and $I_2(x, y)$ are subjected to discrete wavelet transform (DWT) to isolate spatial-frequency features at multiple resolution levels. This decomposition segregates low-frequency approximation sub- bands A^k, which capture structural and luminance information, and high-frequency detail sub-bands $D^{k,o}$ (horizontal, vertical, diagonal), capturing edges and textures is given by (1).

$$I(x, y) = A^K(x, y) + \sum_{\kappa=1 o\in \mathrm{H},V,D}^{K} \sum D^{\kappa,o}(x, y) \tag{1}$$

In (1), K denotes the number of decomposition levels and o the orientation.

To balance performance and computational cost, we adopt a 3-level decomposition ($K = 3$) using the Haar wavelet, which ensures orthogonality and efficient computation. The significance of each detail component is quantified using local energy is given by (2) and which reflects the presence of local texture and guides the fusion process.

$$E^{\kappa}(x, y) = \sum_{o\in H,V,D} |D^{\kappa,o}(x, y)|^2 \tag{2}$$

3.2 Adaptive Weighting and Fusion

Fusion is achieved by applying adaptive, content-aware weights to corresponding sub-bands of the source images. For detail sub-bands $D_1^{\kappa,o}$ and $D_2^{\kappa,o}$, a saliency map is calculated based on local energy within a 5×5 window W.

$$S^{\kappa,o}(x, y) = \sum_{(m,n)\in W} |D^{\kappa,o}(m, n)|^2 \tag{3}$$

Normalized weights are then derived as (4) and (5) and ensuring localized and balanced fusion as (6).

$$w_1^{\kappa,o}(x, y) = \frac{S_1^{\kappa,o}(x, y)}{S_1^{\kappa,o}(x, y) + S_2^{\kappa,o}(x, y)} \tag{4}$$

$$w_2^{\kappa,o}(x, y) = 1 - w_1^{\kappa,o}(x, y) \tag{5}$$

$$\mathrm{F}^{\kappa,o}(x, y) = w_1^{\kappa,o}(x, y) \cdot D_1^{\kappa,o}(x y) + w_2^{\kappa,o}(x, y) \cdot D_2^{\kappa,o}(x, y) \tag{6}$$

For approximation components A_k^1 and A_k^2 , local variance within a window guides fusion, preserving global structural consistency is given by (7).

$$\sigma^2(x, y) = \frac{1}{|W|} \sum_{(m,n)\in w} \left(A^K(m, n) - \mu(x, y)\right)^2 \tag{7}$$

In equation (7), $\mu(x, y)$ is the local mean. The corresponding weights are computed as follows (8) and (9).

$$w_1^K(x, y) = \frac{\sigma_1^2(x, y)}{\sigma_1^2(x, y) + \sigma_2^2(x, y)} \tag{8}$$

$$\mathrm{w}_2^{\mathrm{K}}(x, y) = 1 - \mathrm{w}_1^{\mathrm{K}}(x, y) \tag{9}$$

The equations (8) and (9) yielding the fused approximation as (10).

$$\mathrm{F}^K(x, y) = w_1^K A_1^K + w_2^K A_2^K \tag{10}$$

The inverse DWT (IDWT) is then applied to reconstruct the fused image $F(x, y)$ from all fused sub-bands.

3. 3 Post-Fusion Enhancement

To further refine the fused image and address perceptual quality issues such as low contrast or mild blurring, a two- stage enhancement process is applied.

1) CLAHE-Based Contrast Enhancement: Contrast- Limited Adaptive Histogram Equalization (CLAHE) is applied to improve local contrast while preventing noise over-amplification. The image is divided into 8×8 tiles, and the histogram of each tile is clipped at a threshold (e.g., 0.01), with redistribution guided by the cumulative distribution function (CDF). The enhanced image is given by (11).

$$\mathrm{F}_{CLAHE}(x, y) = T(F(x, y)) \tag{11}$$

In (11), T is the local intensity mapping derived from the clipped CDF.
2) Unsharp Masking for Detail Enhancement: To restore edge sharpness and improve fine details, unsharp masking is applied and can be written as (12).

$$F_{enhanced}(x, y) = F_{CLAHE}(x, y) + \lambda \cdot (F_{CLAHE}(x, y) - G_{\sigma} * F_{CLAHE}(x, y)) \tag{12}$$

In (12), G_{σ} is a Gaussian blur kernel with standard deviation σ = 1.5, $*$ denotes convolution, and λ = 0.8 controls the sharpening effect. This stage accentuates high-frequency components and compensates for the slight smoothing introduced by wavelet reconstruction.

As shown in Figure 3, the methodology diagram illustrates the complete pipeline of the proposed image fusion frame- work, encompassing four primary stages: preprocessing, weight map computation, multi-scale fusion, and post-fusion enhancement. The process begins with two source images, which undergo preprocessing to convert them into LAB and RGB color spaces, enabling extraction of luminance and chromatic features critical for perceptual quality. In the second stage, multiple weight maps—namely gradient weight, local contrast weight, and exposure weight—are computed and normalized to emphasize salient features across the image pair. These weight maps guide the fusion in the third stage, where a multi-scale approach using Gaussianand Laplacian pyramid decomposition facilitates the integration of image features at various frequency levels.

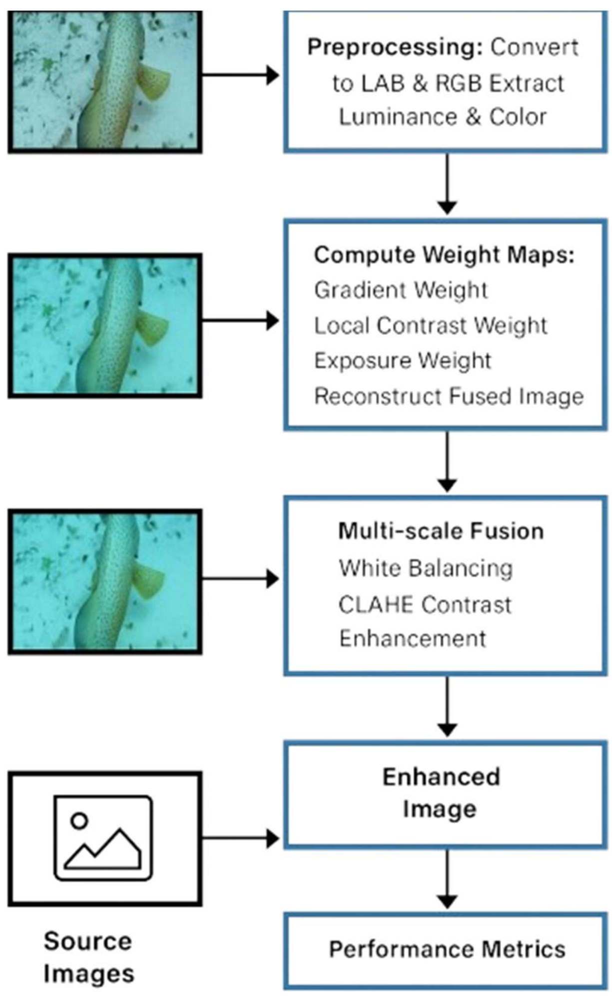

Figure 3. Block diagram of the proposed image fusion framework illustrating multi-scale decomposition, adaptive fusion, and post-fusion enhancement.

The reconstructed image is then passed through a post-fusion enhancement stage involving white balancing and CLAHE, which improves local contrast and color balance. The final enhanced image is then subjected to quantitative evaluation using performance metrics such as UCIQE, PCQI, and entropy, ensuring objective assessment of visual quality improvements.

4. RESULT AND DISCUSSION

The efficacy of the proposed image fusion framework, integrating multi-scale adaptive weighting and post-fusion enhancement, is validated both visually and quantitatively using benchmark Test Image 1 and Test Image 2. The evaluation metrics and qualitative comparisons affirm the method's capability to enhance detail preservation, contrast clarity, and structural fidelity across diverse input conditions.

4. 1 Performance on Test Image-1

Figure 4 presents the fusion results for Test Image 1. The fused output demonstrates significantly enhanced detail retention and improved contrast, largely due to the adaptive weighting scheme prioritizing high-frequency salient components during multi-scale decomposition. Quantitative assessment further corroborates these observations.

Specifically, the fused image attains a UCIQE score of 1.4218, indicating satisfactory underwater image quality. A PCQI of 1.2295 suggests effective contrast preservation. The IE value of 7.5727 reflects high information content, aligning well with the visually noticeable texture and edge clarity. However, the UICM (0.4370) and CCF (0.3264) scores suggest moderate colorfulness and a lower cross-correlation with source images—possibly due to the emphasis on structural features over chromatic fidelity. The FADE score of 1.0592 indicates minor perceptual density distortions, which the enhancement stage only partially mitigates.

Figure 4. Fusion result for Test Image 1 using the proposed method.

4. 2 Performance on Test Image-2

Figure 5 depicts the results for Test Image 2, demonstrat- ing the method's adaptability to variations in source image characteristics. Compared to Test Image 1, the fused result exhibits sharper edges, richer textures, and enhanced color vividness.

The fused output achieves a UCIQE of 1.8607 and a PCQI of 1.8035—both significantly higher than those in Test Image 1—indicating substantial improvements in perceived color and contrast. The UICM (0.6063) and CCF (0.4105) values suggest better colorfulness and stronger structural correlation with source images. Notably, the IE rises slightly to 7.6014, reinforcing the method's capacity to preserve information entropy. However, a higher FADE value of 1.6583 reflects a perceptual trade-off, potentially resulting from over-sharpening introduced during the enhancement stage. The IECCF of 1.4383 confirms the system's effectiveness in boosting local contrast variations, though the elevated FADE suggests that post-processing parameters may require fine tuning for optimal results in high-density regions.

4. 3 Comparative Quantitative Analysis

Table 1 consolidates the quantitative evaluation metrics across both test cases. The results affirm that the proposed framework consistently maintains high information entropy and strong contrast preservation. Notably, Test Image 2 shows superior performance in most metrics, highlighting the method's robustness and adaptability to diverse input conditions.

Figure 5. Fusion result for Test Image 2 using the proposed method.

The results collectively highlight the strength of the proposed fusion strategy. The integration of multi-scale adaptive weighting successfully enhances the salient features while mitigating artifacts common in conventional fusion techniques. The post-fusion enhancement stage further refines perceptual clarity, though the elevated FADE in Test Image 2 indicates that sharpening parameters may require adaptive tuning based on content density. These findings confirm the generalizability and effectiveness of the proposed method while identifying avenues for fine-tuning and domain-specific calibration. Figure 6 illustrates the quantitative metrics comparison between Test Image 1 and Test Image 2. Each metric (UCIQE, UICM, PCQI, etc.) is shown with its respective scores for both test images, clearly highlighting the performance differences across evaluation parameters.

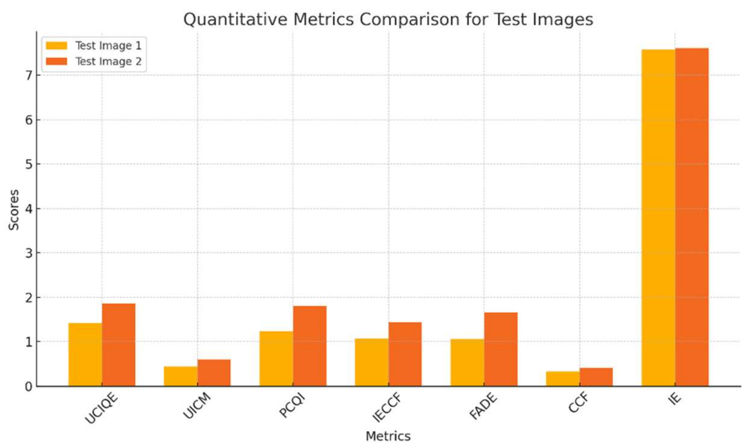

Figure 6. Quantitative metrics comparison between Test Image 1 and Test Image 2.

Table 1. Quantitative Metrics for test images.

Metric	Test Image 1	Test Image 2
UCIQE	1.4218	1.8607
UICM	0.4370	0.6063
PCQI	1.2295	1.8035
IECCF	1.0737	1.4383
FADE	1.0592	1.6583
CCF	0.3264	0.4105
IE	7.5727	7.6014

5. CONCLUSION

This study presents an enhanced and comprehensive image fusion framework that effectively integrates multi-scale decomposition with adaptive weighting and post-fusion enhancement to address the persistent challenges in visual information integration. The proposed method leverages discrete wavelet transform for hierarchical feature extraction, adaptive weighting based on local salience and statistical variance, and a two-stage enhancement strategy combining Contrast-Limited Adaptive Histogram Equalization with unsharp masking to refine the visual output. This synergistic approach enables the seamless integration of complementary information from input images, thereby enhancing structural detail, contrast, and overall perceptual quality.

Experimental evaluations conducted on benchmark datasets confirm the efficacy of the proposed method, with superior performance in both objective metrics and visual clarity. For instance, Test Image 2 achieved a UCIQE of 1.8607 and a PCQI of 1.8035, indicating significant improvements in color quality and contrast preservation. Additionally, the method demonstrated high information entropy while maintaining computational efficiency, making it suitable for practical applications such as medical diagnostics, underwater imaging, and remote sensing. Despite these advantages, some limitations were observed. Elevated FADE scores, such as 1.6583 for Test Image 2, indicate potential over-enhancement in texture-dense regions, while moderate CCF values (e.g., 0.4105) suggest that further refinement is needed to improve fidelity to source images. These observations highlight the need for adaptive parameter tuning and more robust weighting mechanisms.

DECLARATIONS

Conflict of interest

All authors declared that there is no conflict of interest in any form.

Data availability interest

The data supporting this study's findings are available from the corresponding author upon reasonable request.

Funding Acknowledgment

The authors declare that no financial support was received for this study.

Publisher's note: The Journal and Publisher remain neutral about jurisdictional claims in published maps and institutional affiliations.

REFERENCES

[1]. Zhou T, Zhang M, Li Y, Zhou W, Yu W. Adaptive Multi-Weight Infrared and Visible Image Fusion via Multi-Scale Transformation. In2023 42nd Chinese Control Conference (CCC) 2023 Jul 24 (pp. 1-7). IEEE. 10.23919/CCC58697.2023.10241231

[2]. Luo H, Hu W. Multi-Scale Feature Adaptive Fusion for Multi-Task Dense Prediction. In Proceedings of the 2024 7th International Conference on Image and Graphics Processing 2024 Jan 19 (pp. 294-300) https://doi.org/10.1145/3647649.3647696

[3]. Yang Y, Zhang D, Wan W, Huang S. Multi-scale exposure fusion based on multi-visual feature measurement and detail enhancement representation. IEEE Transactions on Instrumentation and Measurement. 2022 May 23;71:1-4. 10.1109/TIM.2022.3176881

[4]. Li S, Wan L, Tang L, Zhang Z. MFEAFN: Multi-scale feature enhanced adaptive fusion network for image semantic segmentation. Plos one. 2022 Sep 30;17(9):e0274249. https://doi.org/10.1371/journal.pone.0274249

[5]. Hu Z, Liang W, Ding D, Wei G. An improved multi-focus image fusion algorithm based on multi-scale weighted focus measure. Applied Intelligence. 2021 Jul;51(7):4453-69. https://doi.org/10.1007/s10489-020-02066-8

[6]. Gao Y, Ma S, Liu J. Multi-Scale Fusion Global Feature Extraction Network for Multi-Modal Medical Image Fusion. Available at SSRN 4580770. Gao, Yuan and Ma, Shiwei and Liu, Jingjing, Multi-Scale Fusion Global Feature Extraction Network for Multi-Modal Medical Image Fusion. Available at SSRN: https://ssrn.com/abstract=4580770 or http://dx.doi.org/10.2139/ssrn.4580770

[7]. Liu J, Duan M, Chen WB, Shi H. Adaptive weighted image fusion algorithm based on NSCT multi-scale decomposition. In2020 International Conference on System Science and Engineering (ICSSE) 2020 Aug 31 (pp. 1-5). IEEE. 10.1109/ICSSE50014.2020.9219295

[8]. Qiu L, Song X, Ying Z, Feng W, Zhong L, Pan J. MC-FAW: A Multi-Scale Convolutional Feature Adaptive Weighting Fusion Network for Detecting Disorders of Consciousness. In2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2024 Dec 3 (pp. 1639-1644). IEEE. 10.1109/BIBM62325.2024.10822312

[9]. Liu C, Feng J. AMSIN: An adaptive multi-scale input network for hyperspectral image fusion. Infrared Physics & Technology. 2024 Aug 1;140:105347. https://doi.org/10.1016/j.infrared.2024.105347
[10]. Zhang X, Li J, Hua Z. MFFE: multi-scale feature fusion enhanced net for image dehazing. Signal Processing: Image Communication. 2022 Jul 1;105:116719. https://doi.org/10.1016/j.image.2022.116719
[11]. Zhang X, Li J, Hua Z. MFFE: multi-scale feature fusion enhanced net for image dehazing. Signal Processing: Image Communication. 2022 Jul 1;105:116719. https://doi.org/10.1016/j.image.2022.116719
[12]. kumar Soma A. Hybrid RNN-GRU-LSTM Model for Accurate Detection of DDoS Attacks on IDS Dataset. Journal of Modern Technology. 2025 May 14;2(01):283-91. https://doi.org/10.71426/jmt.v2.i1.pp283-291