



Article

Multi-Focus Image Fusion Based on Dual-Channel Rybak Neural Network and Consistency Verification in NSCT Domain

Ming Lv ^{1,2}, Sensen Song ^{1,2,3,*}, Zhenhong Jia ^{1,2}, Liangliang Li ^{4,*} and Hongbing Ma ⁵

¹ School of Computer Science and Technology, Xinjiang University, Urumqi 830046, China; lvming09@163.com (M.L.); jzhh@xju.edu.cn (Z.J.)

² Key Laboratory of Signal Detection and Processing, Xinjiang University, Urumqi 830046, China

³ College of Mathematics and System Science, Xinjiang University, Urumqi 830046, China

⁴ School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

⁵ Department of Electronic Engineering, Tsinghua University, Beijing 100084, China; hbma@tsinghua.edu.cn

* Correspondence: songsensen@xju.edu.cn (S.S.); leeliangliang@163.com (L.L.)

Abstract

In multi-focus image fusion, accurately detecting and extracting focused regions remains a key challenge. Some existing methods suffer from misjudgment of focus areas, resulting in incorrect focus information or the unintended retention of blurred regions in the fused image. To address these issues, this paper proposes a novel multi-focus image fusion method that leverages a dual-channel Rybak neural network combined with consistency verification in the nonsubsampled contourlet transform (NSCT) domain. Specifically, the high-frequency sub-bands produced by NSCT decomposition are processed using the dual-channel Rybak neural network and a consistency verification strategy, allowing for more accurate extraction and integration of salient details. Meanwhile, the low-frequency sub-bands are fused using a simple averaging approach to preserve the overall structure and brightness information. The effectiveness of the proposed method has been thoroughly evaluated through comprehensive qualitative and quantitative experiments conducted on three widely used public datasets: Lytro, MFFW, and MFI-WHU. Experimental results show that our method consistently outperforms several state-of-the-art image fusion techniques, including both traditional algorithms and deep learning-based approaches, in terms of visual quality and objective performance metrics ($Q_{AB/F}$, Q_{CB} , Q_E , Q_{FMI} , Q_{MI} , Q_{MSE} , Q_{NCIE} , Q_{NMI} , Q_P , and Q_{PSNR}). These results clearly demonstrate the robustness and superiority of the proposed fusion framework in handling multi-focus image fusion tasks.



Academic Editors: Costas Psychalinos and Yongguang Yu

Received: 6 May 2025

Revised: 17 June 2025

Accepted: 29 June 2025

Published: 30 June 2025

Citation: Lv, M.; Song, S.; Jia, Z.; Li, L.; Ma, H. Multi-Focus Image Fusion Based on Dual-Channel Rybak Neural Network and Consistency Verification in NSCT Domain. *Fractal Fract.* **2025**, *9*, 432. <https://doi.org/10.3390/fractfract9070432>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Multi-source image fusion [1,2] refers to the technology of comprehensively processing multiple images acquired from different sensors, modalities, or time points to generate a fused image containing richer and more comprehensive information. This technique is widely applied in fields such as remote sensing [3,4], medical imaging [5–7], computer vision [8–10], and military surveillance [11–13].

In real-world imaging scenarios, optical devices often struggle to capture an entirely clear scene in a single shot due to inherent limitations in depth of field (DoF) [14]. This results in multi-focus images, where different regions of the same scene are selectively focused while others remain blurred. Multi-focus image fusion (MFIF) addresses this

challenge by integrating complementary information from a series of partially focused images to synthesize a comprehensive, fully focused representation [15,16]. Such fused images are critical for enhancing visual quality and enabling high-level tasks, including medical diagnosis, microscopic imaging, surveillance systems, and autonomous robotics, where precision in detail is paramount [17,18].

Multi-focus image fusion algorithms can be categorized based on their core techniques and implementation approaches as follows: transform domain-based methods, spatial domain-based methods, deep learning-based methods, and hybrid methods [19,20]. In terms of transform domain, these methods transform input images into frequency or multi-scale domains, fuse coefficients (e.g., maximum selection, weighted averaging) to retain focused regions, and reconstruct the fused image; these methods include Laplacian pyramid, discrete wavelet transform, curvelet, contourlet, shearlet, etc. [21,22]. Li et al. [23] introduced the image fusion technique via sparse representation and guided filtering in the Laplacian pyramid domain, this algorithm demonstrates outstanding performance in infrared-visible image fusion and multi-focus image fusion. Li et al. [24] introduced the multi-focus image fusion approach using fractal dimension in curvelet domain. Lv et al. [25] introduced the image fusion via distance-weighted regional energy and structure tensor in the nonsubsampled contourlet transform (NSCT) domain; this algorithm has achieved remarkable results in both subjective and objective metrics for image fusion. Li et al. [26] proposed the image fusion approach via local energy in the shearlet domain; this algorithm demonstrates favorable performance in multi-focus image fusion. Nevertheless, the shearlet transform necessitates equal height and width for input images, thereby constraining its applicability to image fusion tasks involving varying sizes. The algorithms demonstrate robust multi-scale decomposition capabilities while achieving superior performance in edge and texture preservation. However, these methods are subject to certain limitations: they may introduce potential artifacts (particularly Gibbs oscillations), and their fusion processes often rely on heuristic rule design rather than theoretically optimized approaches [27–29].

Fractal and fractional methods have been widely applied in image fusion. For instance, Li et al. [15] developed a multi-focus image fusion method using fractal dimension within the NSCT domain, implemented via coupled neural P systems. Panigrahy et al. [30] presented a parameter-adaptive dual-channel PCNN model for multi-focus image fusion, incorporating fractal dimension. Joshua et al. [31] introduced an adaptive low-light image enhancement algorithm combining a novel intuitionistic fuzzy generator with fractal-fractional derivatives. Xian et al. [32] proposed a multi-focus fusion approach leveraging visual depth and fractional-order differentiation operators integrated with convolution norms. Zhang et al. [10] employed fractional-order differentiation and closed image matting for multi-focus fusion, while Zhang et al. [14] utilized fractional-order derivatives and intuitionistic fuzzy sets. Lu et al. [33] enhanced multi-focus fusion by incorporating residual removal and fractional-order differentiation focus measures. Additionally, Li et al. [34] introduced an adaptive fractional differential method with guided filtering for fusion, and Yu et al. [35] proposed a sparse representation framework based on fractional-order differentiation for multi-focus image fusion.

Spatial domain-based methods operate directly on pixel or local patch levels by selecting focused regions using sharpness metrics (e.g., gradient, energy, variance). These approaches offer computational simplicity, making them particularly suitable for real-time applications. However, they exhibit certain limitations: their performance is sensitive to block size selection and may introduce undesirable blocking artifacts during the fusion process [36].

Deep learning-based methods leverage neural networks to autonomously extract focus-related features and learn optimal fusion rules, encompassing both end-to-end and

multi-stage architectures. Representative algorithms in this domain include CNNs, GANs, Transformers, Mamba, and their variants, which demonstrate remarkable adaptability and consistently deliver cutting-edge performance [37,38]. Zhang et al. [39] proposed the unsupervised generative adversarial network with adaptive and gradient joint constraints for image fusion. While these approaches have revolutionized image fusion through their data-driven nature, they inherently require extensive training datasets and demand significant computational resources for both training and deployment.

Hybrid methods strategically combine transform-domain, spatial-domain, and deep learning techniques to harness their complementary advantages. As exemplified by Avci et al.'s [40] innovative approach integrating discrete wavelet transform with deep convolutional neural networks for multi-focus image fusion, these hybrid architectures consistently demonstrate superior performance when benchmarked against conventional methods. While offering enhanced robustness for complex scene analysis, these methods inevitably introduce higher algorithmic complexity and present significant parameter optimization challenges due to their sophisticated multi-component nature.

In this paper, a novel multi-focus image fusion method based on the dual-channel Rybak neural network (DRYNN) and consistency verification (CV) in the NSCT domain is proposed. The dual-channel Rybak neural network and consistency verification models are used to process the high-frequency sub-bands generated by NSCT, and the average method is used to deal with the low-frequency sub-bands generated by NSCT. One of the key innovations of this paper is the application of consistency verification to the processing of high-frequency coefficients, which significantly improves the image fusion performance. The proposed algorithm was evaluated through a series of qualitative and quantitative analyses on three public datasets: Lytro, MFFW, and MFI-WHU [24]. Comparisons with state-of-the-art traditional and deep learning-based image fusion methods demonstrate the superiority of the proposed approach.

The remainder of this paper is organized as follows: Section 2 reviews the dual-channel Rybak neural network, Section 3 details the proposed methodology, Section 4 presents experimental results and analysis, and Section 5 concludes with future directions.

2. Dual-Channel Rybak Neural Network

The dual-channel Rybak neural network (DRYNN) [41], illustrated in Figure 1, processes two input images, A and B . After up to N iterations, it outputs a binary decision matrix D , which indicates the neuron with the higher internal activity between the corresponding neurons from the two inputs. The DRYNN model can be mathematically formulated as follows:

$$XS^A(i, j) = V_S \sum_{x=-2}^2 \sum_{y=-2}^2 F_S(x+3, y+3) S^A(i+x, j+y) \quad (1)$$

$$XS^B(i, j) = V_S \sum_{x=-2}^2 \sum_{y=-2}^2 F_S(x+3, y+3) S^B(i+x, j+y) \quad (2)$$

$$XI_n^A(i, j) = \alpha \sum_{x=-2}^2 \sum_{y=-2}^2 F_I(x+3, y+3) Z_{n-1}(i+x, j+y) + S^A(i, j) \quad (3)$$

$$XI_n^B(i, j) = \alpha \sum_{x=-2}^2 \sum_{y=-2}^2 F_I(x+3, y+3) Z_{n-1}(i+x, j+y) + S^B(i, j) \quad (4)$$

where $XS^X|X \in \{A, B\}$ denotes the direct inputs, while $XI^X|X \in \{A, B\}$ refers to the feedback inputs. $S^X|X \in \{A, B\}$ represents the external stimulus associated with input

X. Feedback inputs also receive their corresponding stimuli to ensure sufficient feature extraction. The variable n indicates the iteration number, whereas (i, j) represents the (i, j) -th neuron. The parameter α denotes the coefficient of backward inhibition, and V_S represents the linking amplitude. F_S and F_I are the two 5×5 matrices that show the on-center/off-surround and oriented local connection receptive fields of the DRYNN model, respectively. The values of F_S and F_I are determined by Equations (5) and (6), respectively. Z is the binary output matrix, calculated using Equation (7).

$$F_S = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (5)$$

$$F_I = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (6)$$

$$Z_n(i, j) = \begin{cases} 1, & \text{if } P_n(i, j) > E_{n-1}(i, j) \\ 0, & \text{else} \end{cases} \quad (7)$$

where P and E show the internal activity and threshold function of the DRYNN model, respectively. The internal activity is calculated using the following equation:

$$P_n(i, j) = e^{-\alpha_p} P_{n-1}(i, j) + \max\{P_n^A(i, j), P_n^B(i, j)\} \quad (8)$$

where α_p represents the decay constant governing the internal activity in the DRYNN model. The variables $P_n^X(i, j)|X \in \{A, B\}$ denote the internal states of the (i, j) -th neuron corresponding to the input of the DRYNN model at iteration n th. These internal states are computed as follows.

$$P_n^A(i, j) = X S^A(i, j) - \left(\frac{1}{\Phi}\right) X I_n^A(i, j) + h \quad (9)$$

$$P_n^B(i, j) = X S^B(i, j) - \left(\frac{1}{\Phi}\right) X I_n^B(i, j) + h \quad (10)$$

where Φ denotes the time constant associated with the internal blocks, while h serves as the threshold constant. The dynamic threshold is determined using Equation (11).

$$E_n(i, j) = e^{-\alpha_e} E_{n-1}(i, j) + V_E Z_n(i, j) \quad (11)$$

where α_e and V_E denote the decay and amplitude constants for E , respectively. The constant parameters of the DRYNN model are determined according to Equation (12).

$$\begin{cases} V_S = 1, \alpha = 1.5, \alpha_p = 0.7, \Phi = 1.2, \\ h = 0.28, \alpha_e = 0.001, V_E = 30 \end{cases} \quad (12)$$

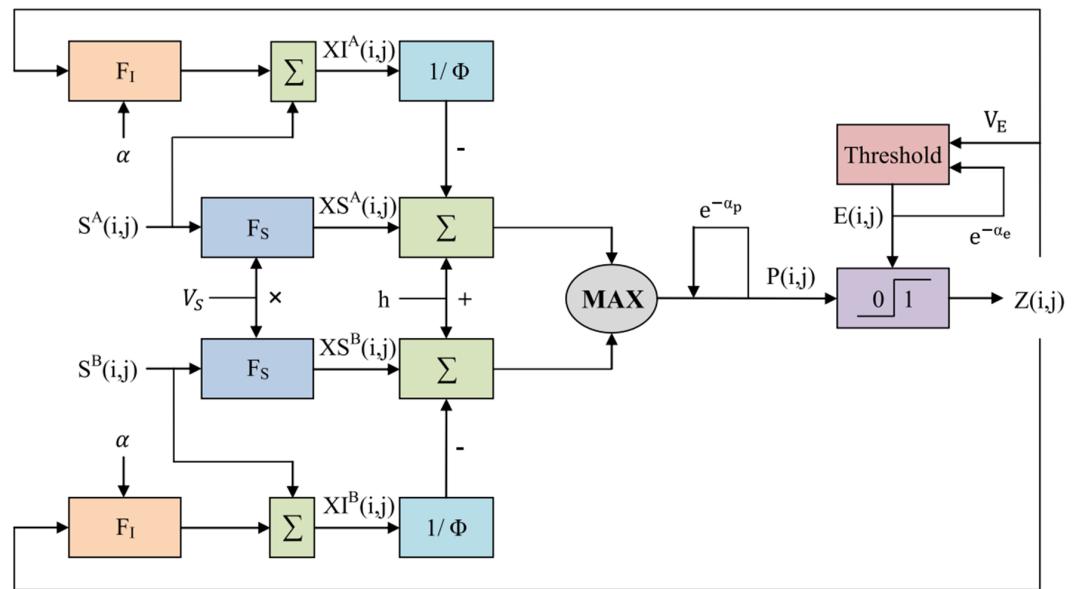


Figure 1. The DRYNN model.

The parameters V_S , α , Φ , h , α_e , and V_E are empirically determined based on the literature by RYNN [42], while the parameter α_p is set using the concept of decay constant for the internal activity of PCNN-based models [43]. The significance of a neuron is reflected in its activity, which represents its internal state.

3. Proposed Fusion Method

We present a multi-focus image fusion algorithm that combines a dual-channel Rybak neural network with consistency verification in the NSCT domain. The overall framework of the proposed method is illustrated in Figure 2. The fusion process consists of four main stages: (1) NSCT decomposition of the input images, (2) fusion of high-frequency sub-bands using the DRYNN model with consistency verification (CV), (3) fusion of low-frequency sub-bands via a simple averaging strategy, and (4) reconstruction of the final fused image through the inverse NSCT. The detailed steps of the algorithm are described as follows.

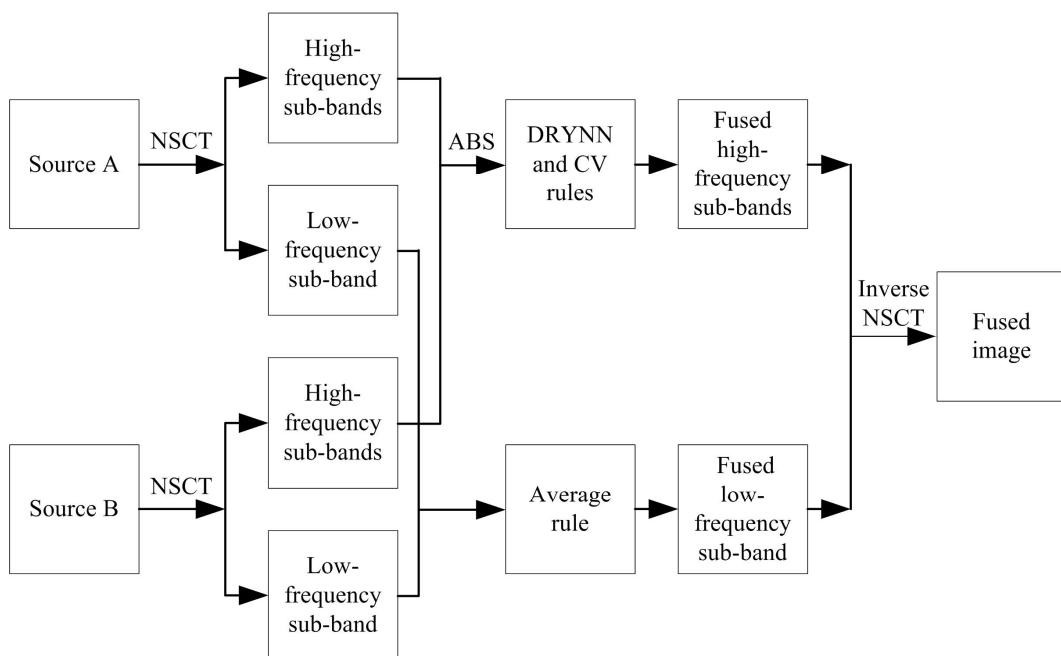


Figure 2. The framework of the proposed algorithm.

3.1. NSCT Decomposition

Owing to its shift-invariant, multi-scale, and multi-directional properties, NSCT exhibits excellent performance in tasks such as multi-focus image fusion [25]. In this section, the NSCT is utilized to perform multi-scale and multi-directional decomposition on the input source images A and B. Through this process, both low-pass $\{LP_A, LP_B\}$ and high-pass $\{HP_A^{l,d}, HP_B^{l,d}\}$ components are extracted from each image. Specifically, the low-frequency component, denoted as $LP_X|X \in \{A, B\}$, captures the coarse structural information, while the high-frequency component, denoted as $HP_X^{l,d}|X \in \{A, B\}$, represents the detailed features of X at the d -th direction corresponding to l -th decomposition level. This decomposition enables the effective separation of image features across various scales and orientations, which is essential for high-quality image fusion.

3.2. High-Frequency Sub-Band Fusion

The DRYNN model is employed to extract and integrate fine details from the high-frequency sub-bands. To generate the fused high-frequency sub-bands, the model's constant parameters are first configured according to Equation (12), followed by initialization using Equation (16). In this process, the absolute values (ABS) of the corresponding high-frequency sub-bands serve as external stimuli for the model, i.e., $S^X = |HP_X^{l,d}|$, where $X \in \{A, B\}$.

$$XS^A(i, j) = V_S \sum_{x=-2}^2 \sum_{y=-2}^2 F_S(x+3, y+3) |HP_A^{l,d}(i+x, j+y)| \quad (13)$$

$$XS^B(i, j) = V_S \sum_{x=-2}^2 \sum_{y=-2}^2 F_S(x+3, y+3) |HP_B^{l,d}(i+x, j+y)| \quad (14)$$

$$P_0(i, j) = 0 \quad (15)$$

$$Z_0(i, j) = 0 \quad (16)$$

$$E_0(i, j) = 1 \quad (17)$$

The computational process begins by initializing parameter E to 1. The DRYNN model is then executed for $N = 200$ optimization cycles. Upon completion, the internal latent states derived from both input streams are aggregated to construct the decision matrix D , generated through the following procedure:

$$D(i, j) = \begin{cases} 1, & \text{if } P_N^{HP_A^{l,d}}(i, j) \geq P_N^{HP_B^{l,d}}(i, j) \\ 0, & \text{else} \end{cases} \quad (18)$$

where $P_N^{HP_X^{l,d}}|X \in \{A, B\}$ is the internal state of $HP_X^{l,d}$ attained after the N -th iteration.

In view of the integrity of the object, the decision map $D(i, j)$ can be refined through a consistency verification (CV) operation [44]:

$$CVD(i, j) = \begin{cases} 1 & \text{if } \sum_{(a,b) \in \theta} D(i+a, j+b) \\ 0 & \text{else} \end{cases} \quad (19)$$

where $CVD(i, j)$ shows the final decision map at position (i, j) , and θ presents a square neighborhood centered at (i, j) with a size of 9×9 .

The fused high-frequency sub-band $HP_F^{l,d}$ can be constructed from D as follows:

$$HP_F^{l,d}(i,j) = \begin{cases} HP_A^{l,d}(i,j), & \text{if } CVD(i,j) = 1 \\ HP_B^{l,d}(i,j), & \text{else} \end{cases} \quad (20)$$

3.3. Low-Frequency Sub-Band Fusion

In this section, the fused low-frequency sub-band image LP_F is generated using the averaging technique, which computes the mean of the corresponding low-frequency coefficients from the source images. This approach ensures a balanced preservation of overall brightness and contrast, and the mathematical formulation is provided as follows [45]:

$$LP_F(i,j) = \frac{LP_A(i,j) + LP_B(i,j)}{2} \quad (21)$$

3.4. Inverse NSCT

The fused image F is reconstructed by applying the inverse NSCT to the fused low-frequency and high-frequency components. This process integrates the fused sub-bands into a single image, effectively preserving both coarse and detailed information, and the corresponding equation is defined as follows:

$$F(i,j) = \text{Inverse NSCT}\left(LP_F(i,j), HP_F^{l,d}(i,j)\right) \quad (22)$$

4. Experimental Results and Analysis

In this section, we validate the effectiveness of the proposed algorithm through comprehensive simulation experiments conducted on three publicly available benchmark datasets: Lytro [46], MFFW [47], and MFI-WHU [39]. The Lytro dataset comprises 20 multi-focus image pairs, the MFFW dataset includes 13 image pairs with challenging focus settings, and a total of 30 representative image pairs were selected from the MFI-WHU dataset for performance evaluation. Figure 3 illustrates sample image pairs drawn from each dataset, showcasing the diversity and complexity of the data used in our experiments.

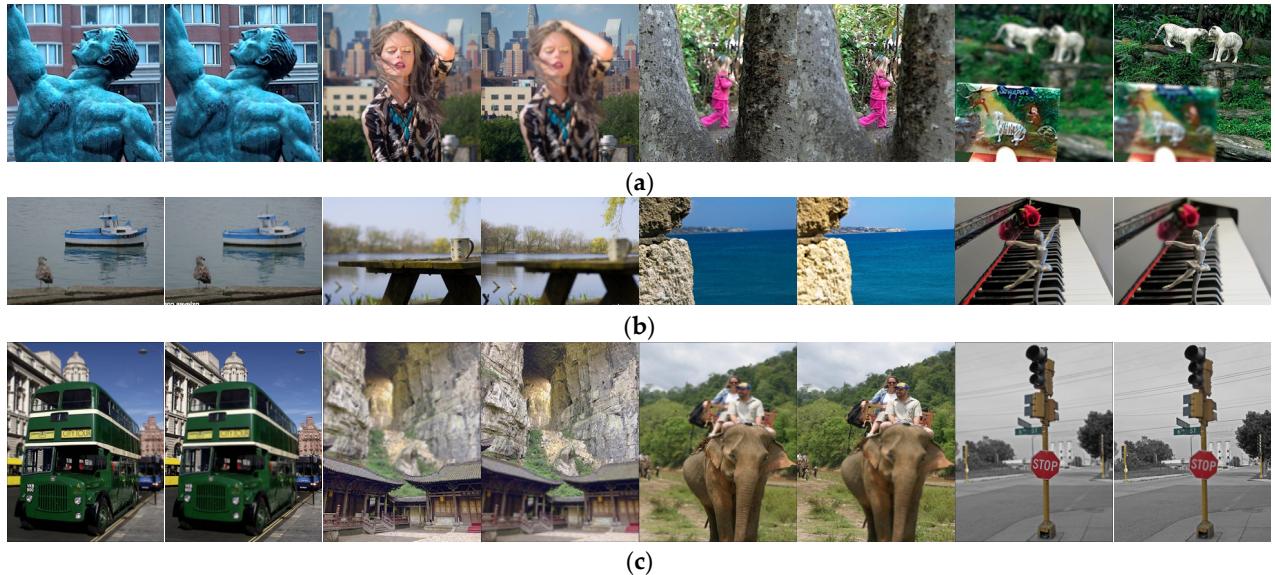


Figure 3. The selection of image pairs from the three datasets. (a) Lytro; (b) MFFW; (c) MFI-WHU.

To ensure a robust and fair comparison, we evaluated our method against eight state-of-the-art image fusion techniques: PMGI [48], MFFGAN [39], U2Fusion [49], XDoG [7],

NSCTST [25], EgeFusion [50], FDFusion [51], and CVTFD [24]. These methods represent a range of classical, transform-based, and deep learning-based approaches.

Furthermore, to objectively assess the quality of the fused images, we employed a suite of widely recognized evaluation metrics. These metrics comprehensively measure various aspects of fusion performance, including $Q_{AB/F}$ [52], Q_{CB} [53], Q_E [53], Q_{FMI} [54], Q_{MI} [52], Q_{MSE} [55–58], Q_{NCIE} [53], Q_{NMI} [53], Q_P [53], and Q_{PSNR} [59–62]. Except for the Q_{MSE} metric, larger values of the other metrics indicate better image fusion performance. In our method, the NSCT decomposition level is 4, and the direction numbers are 4, 8, 8, and 16; the pyramidal and directional filters are defined as “9-7” and “pkva,” respectively.

4.1. Results on Lytro Dataset

Figures 4 and 5 present the simulation results and difference maps of various algorithms applied to a sample dataset from the Lytro collection. The difference maps, generated by subtracting one source input image from the fused output, highlight the fusion performance. As illustrated in Figure 5i, the proposed method produces fusion results with sharp boundaries between focused and defocused regions while preserving color fidelity. In contrast, other methods exhibit limitations: NSCTST and CVTFD demonstrate boundary blurring (Figure 5e,h), while PMGI, MFFGAN, U2Fusion, XDoG, EdgeFusion, and FDFusion introduce significant artifacts (Figure 5a–d,f,g).

Table 1 provides a quantitative comparison of different methods on the Lytro-01 dataset (Figure 4). The proposed algorithm outperforms others, achieving optimal scores in 10 evaluation metrics. Both subjective and objective assessments confirm its superior performance.

Table 1. Quantitative comparative analysis of different methods on the Lytro-01 dataset in Figure 4.

Year	$Q_{AB/F}$	Q_{CB}	Q_E	Q_{FMI}	Q_{MI}	Q_{MSE}	Q_{NCIE}	Q_{NMI}	Q_P	Q_{PSNR}	
PMGI	2020	0.5711	0.6649	0.6497	0.9142	6.0308	133.2364	0.8231	0.8011	0.6796	26.8846
MFFGAN	2021	0.6815	0.6900	0.8314	0.9171	5.6035	30.4630	0.8209	0.7276	0.8118	33.2931
U2Fusion	2022	0.6314	0.5478	0.7841	0.9080	5.4628	19.5930	0.8202	0.7097	0.7443	35.2098
XDoG	2023	0.6829	0.7004	0.8064	0.9147	5.3467	60.4093	0.8197	0.6935	0.8205	30.3198
NSCTST	2023	0.7691	0.8113	0.8783	0.9220	7.1417	26.0130	0.8301	0.9302	0.8994	33.9789
EgeFusion	2024	0.3043	0.3744	0.4831	0.8674	2.4199	93.0763	0.8100	0.3124	0.5134	28.4424
FDFusion	2025	0.7168	0.7686	0.8546	0.9179	6.1315	33.0450	0.8236	0.7983	0.8240	32.9397
CVTFD	2025	0.7597	0.8054	0.8793	0.9219	7.0831	19.6925	0.8296	0.9231	0.8902	35.1878
Proposed		0.7721	0.8268	0.8799	0.9229	7.4256	18.9023	0.8320	0.9675	0.9044	35.3657

Notes: The optimal values are in bold.

Figure 6 provides insight into the performance of different fusion methods on the Lytro dataset. For each metric, the results for individual image pairs are connected into curves, with average scores shown in the legend. Most methods show consistent trends and stable performance, with few outliers. Therefore, the average values in Table 2 are representative. As shown in Table 2, our method achieves the highest scores across all metrics on the Lytro dataset.



Figure 4. Results of Lytro-01. (a) PMGI; (b) MFFGAN; (c) U2Fusion; (d) XDoG; (e) NSCTST; (f) EgeFusion; (g) FDFusion; (h) CVTFD; (i) Proposed.

Table 2. Quantitative average comparative analysis of different methods on the Lytro dataset.

	Year	$Q_{AB/F}$	Q_{CB}	Q_E	Q_{FMI}	Q_{MI}	Q_{MSE}	Q_{NCIE}	Q_{NMI}	Q_P	Q_{PSNR}
PMGI	2020	0.3901	0.5656	0.4736	0.8815	5.8641	75.3956	0.8225	0.8004	0.4620	32.4782
MFFGAN	2021	0.6642	0.6457	0.8409	0.8915	6.0604	34.3748	0.8237	0.8047	0.7125	33.5508
U2Fusion	2022	0.6143	0.5682	0.7835	0.8844	5.7765	59.4424	0.8221	0.7725	0.6657	31.2098
XDoG	2023	0.6885	0.6467	0.8349	0.8926	5.6685	61.2410	0.8216	0.7538	0.7444	30.5692
NSCTST	2023	0.7388	0.7288	0.8745	0.8986	6.7128	31.3985	0.8279	0.8946	0.8100	33.6375
EgeFusion	2024	0.3576	0.4034	0.5032	0.8472	3.2191	77.8597	0.8120	0.4248	0.5405	29.2757
FDFusion	2025	0.6586	0.6127	0.8075	0.8898	6.1834	35.9713	0.8244	0.8252	0.6179	32.9763
CVTFD	2025	0.7285	0.7213	0.8773	0.8988	6.7296	22.2336	0.8279	0.8964	0.7984	35.0684
Proposed		0.7479	0.7515	0.8820	0.9004	7.0950	21.4427	0.8304	0.9455	0.8338	35.2352

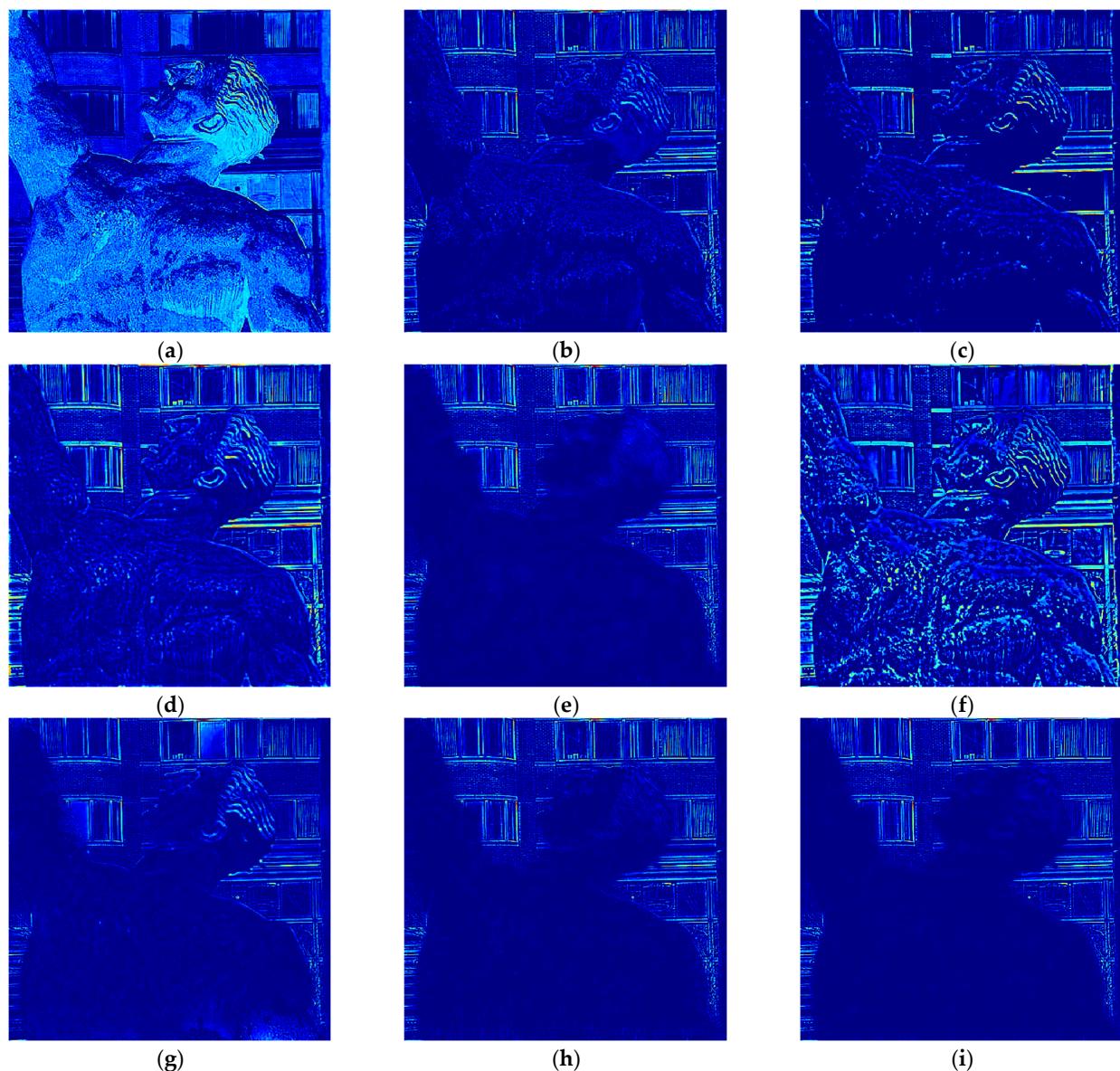
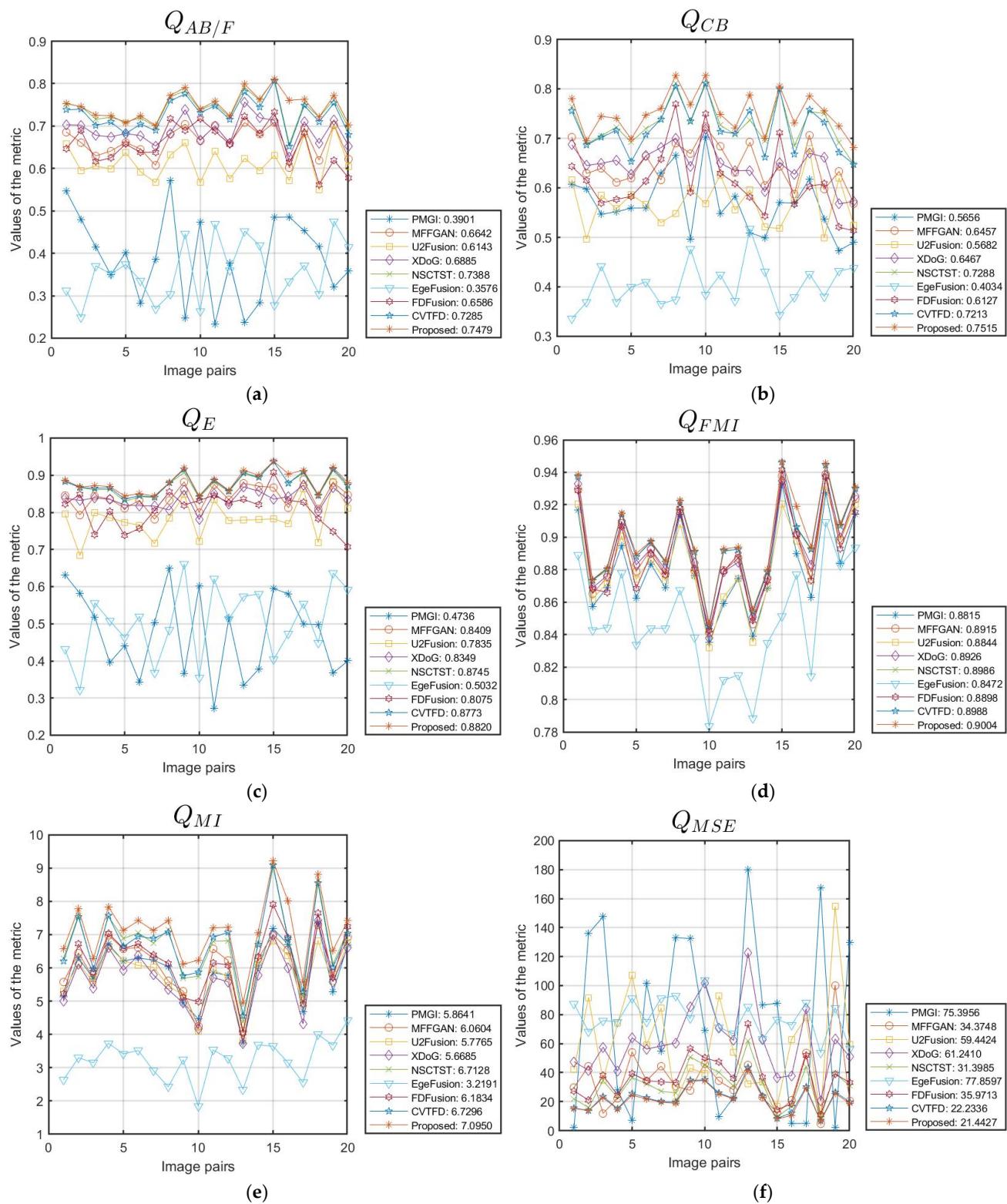


Figure 5. The difference maps of fused images and source image A in Figure 4. (a) PMGI; (b) MFF-GAN; (c) U2Fusion; (d) XDoG; (e) NSCTST; (f) EgeFusion; (g) FDFusion; (h) CVTFD; (i) Proposed.

**Figure 6.** *Cont.*

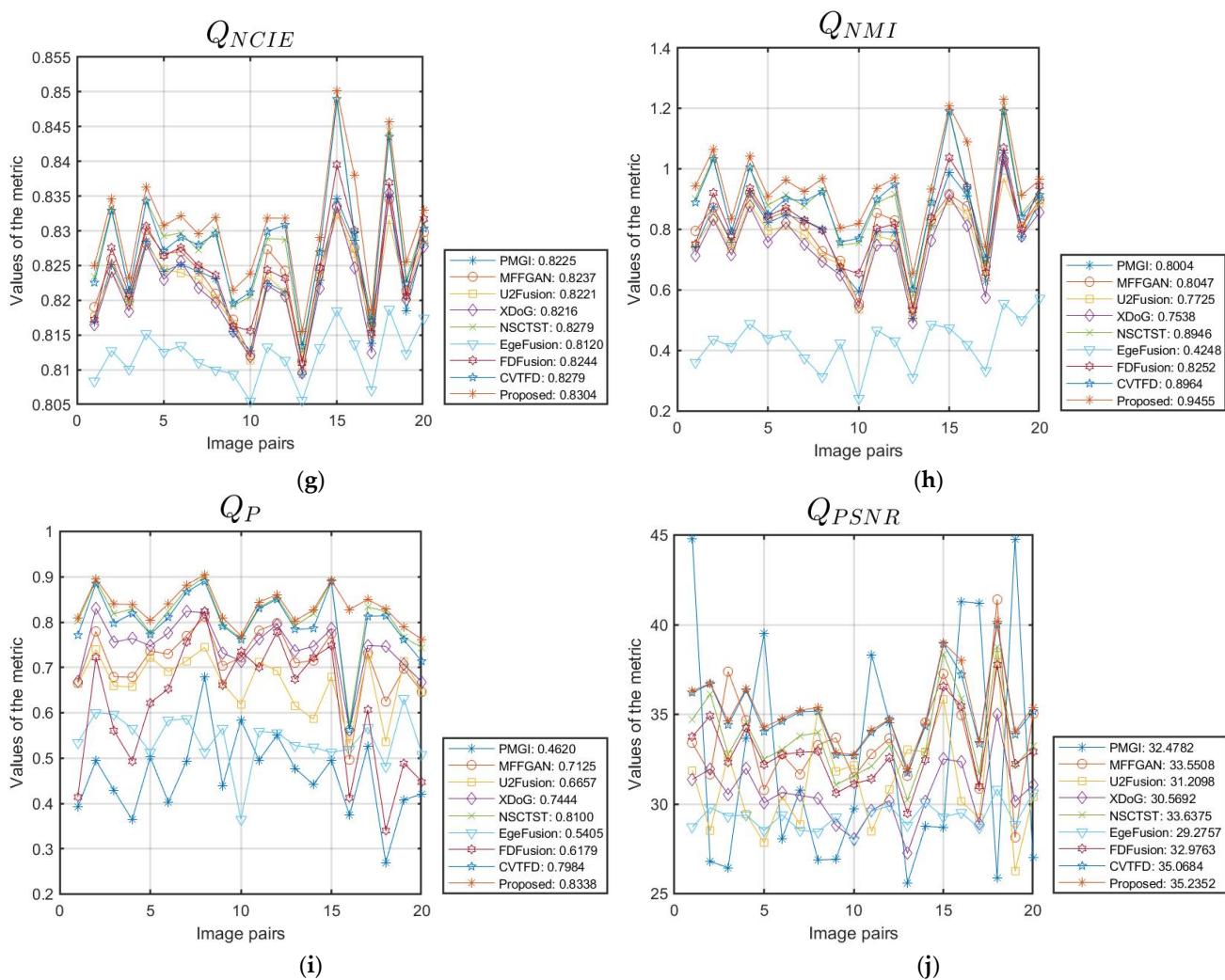


Figure 6. The line chart displays the evaluation metrics across various data from the Lytro dataset.
(a) $Q_{AB/F}$; **(b)** Q_{CB} ; **(c)** Q_E ; **(d)** Q_{FMI} ; **(e)** Q_{MI} ; **(f)** Q_{MSE} ; **(g)** Q_{NCIE} ; **(h)** Q_{NMI} ; **(i)** Q_P ; **(j)** Q_{PSNR} .

4.2. Results on MFFW Dataset

Figures 7 and 8 show that all methods are capable of generating fused images. We can focus on examining the fusion performance of different algorithms in the regions containing the boat and the bird. However, the quality of the difference images varies notably. Specifically, methods such as PMGI, MFFGAN, U2Fusion, XDoG, and EgeFusion struggle to accurately detect near-focus regions due to insufficient extraction of information from the source images. In contrast, NSCTST, FDFusion, and CVTFD perform better in identifying near-focused areas, although some unfocused artifacts remain within these regions. The proposed method, however, demonstrates superior performance in accurately capturing both the central and edge regions of the focus area.

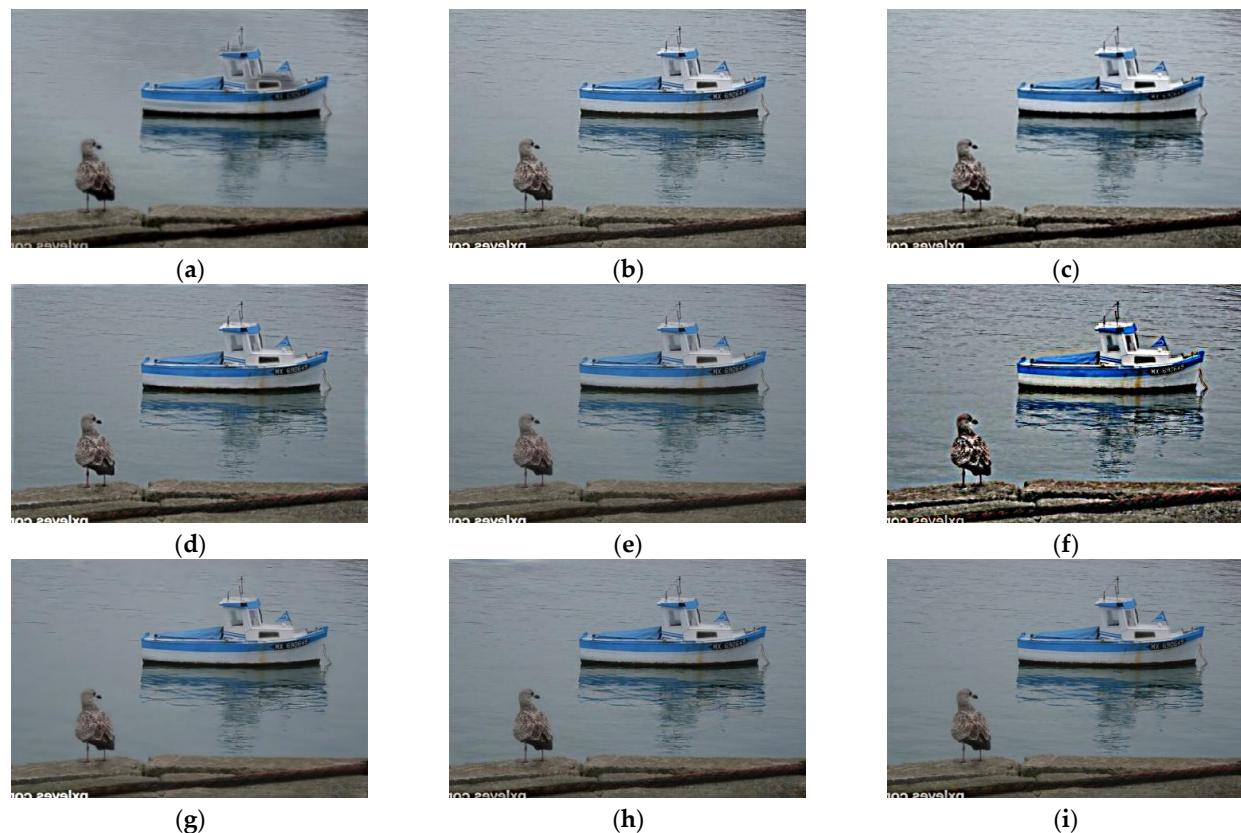


Figure 7. Results of MFFW-01. (a) PMGI; (b) MFFGAN; (c) U2Fusion (d) XDoG; (e) NSCTST; (f) EgeFusion; (g) FDFusion; (h) CVTFD; (i) Proposed.

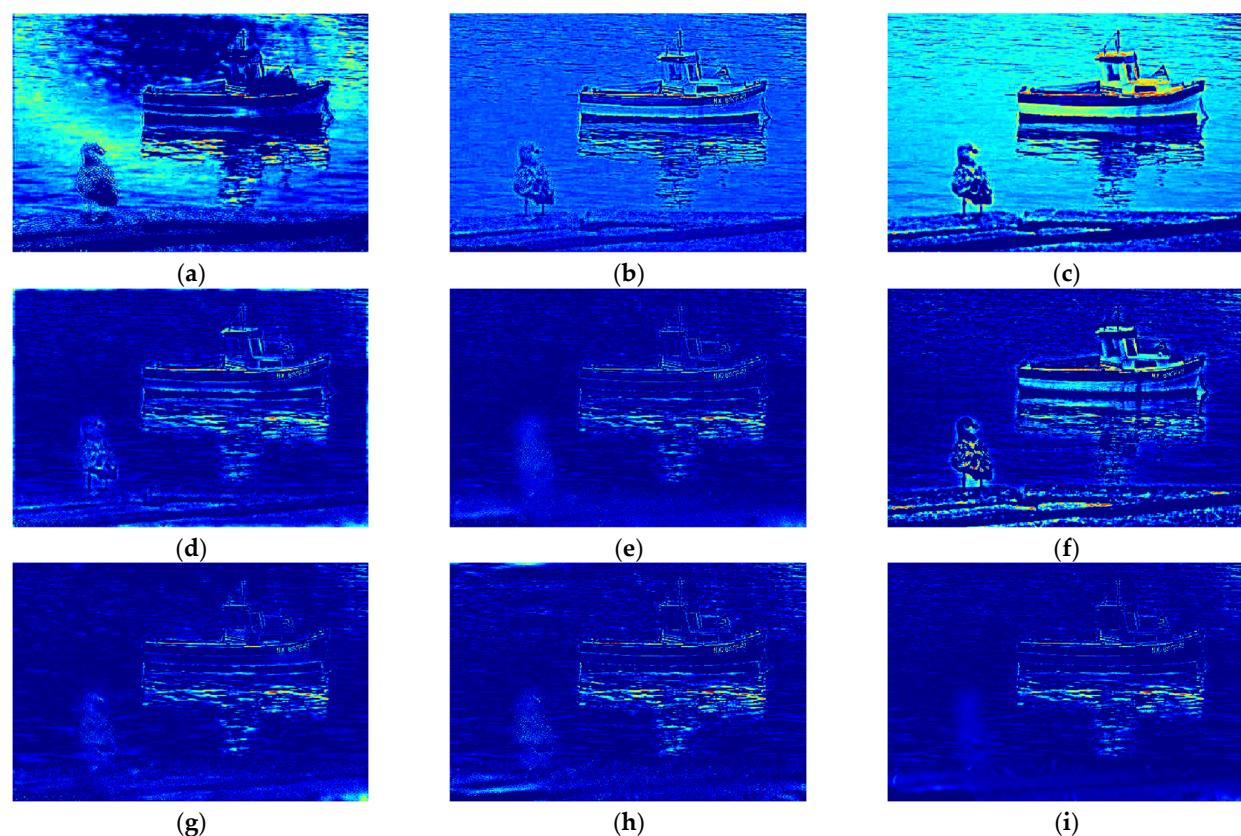


Figure 8. The difference maps of fused images and source image B in Figure 7. (a) PMGI; (b) MFFGAN; (c) U2Fusion; (d) XDoG; (e) NSCTST; (f) EgeFusion; (g) FDFusion; (h) CVTFD; (i) Proposed.

Table 3 presents a quantitative comparison of various methods on the MFFW-01 dataset, as shown in Figure 7. The results indicate that our method achieves the highest performance across 10 evaluation metrics. Both subjective visual assessments and objective measurements validate the effectiveness of our approach for image fusion on the MFFW dataset. Furthermore, the average performance of each method is depicted in Figure 9 and summarized in Table 4, where our method consistently delivers the best overall results, with the exception of the Q_{PSNR} metric.

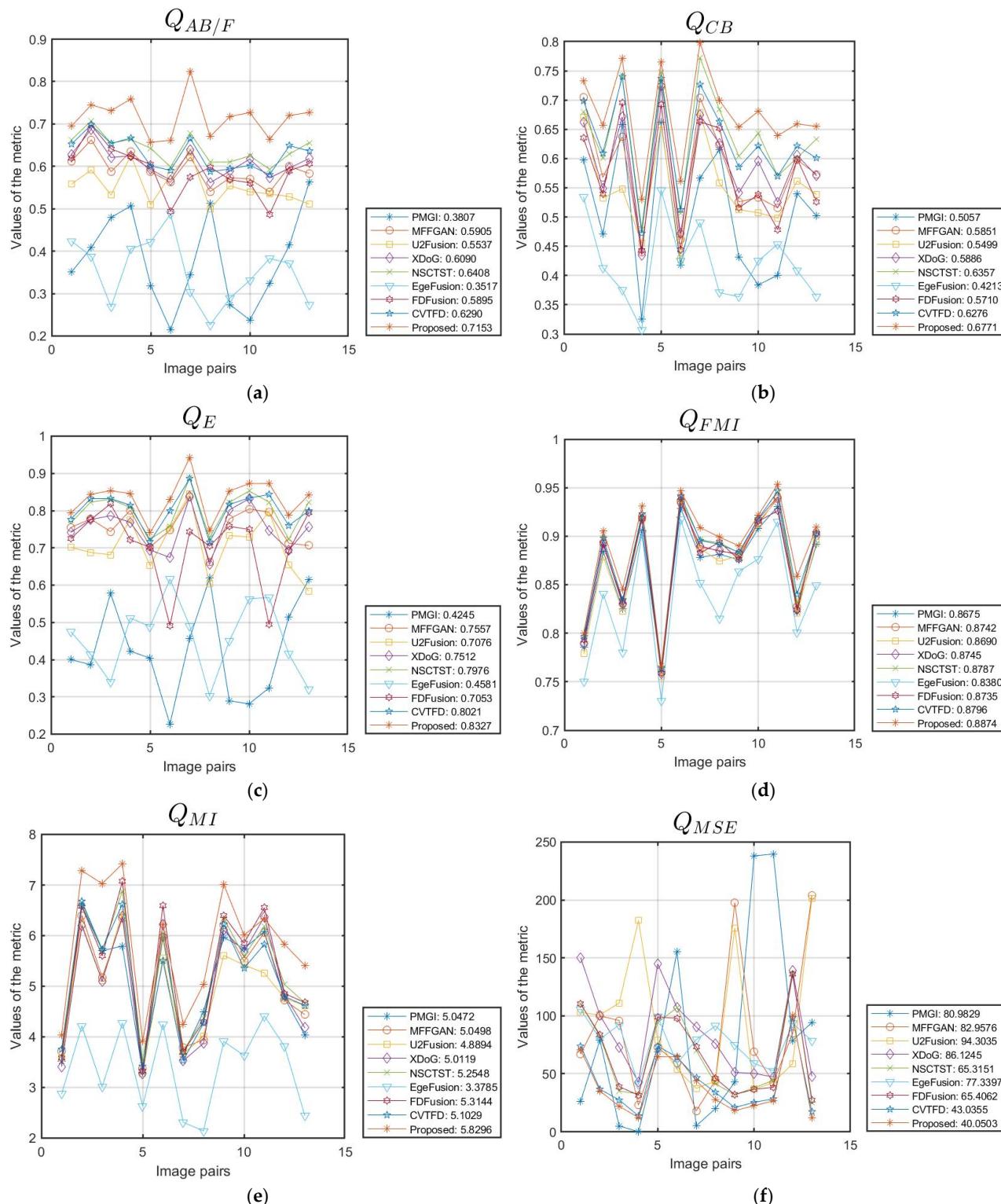


Figure 9. Cont.

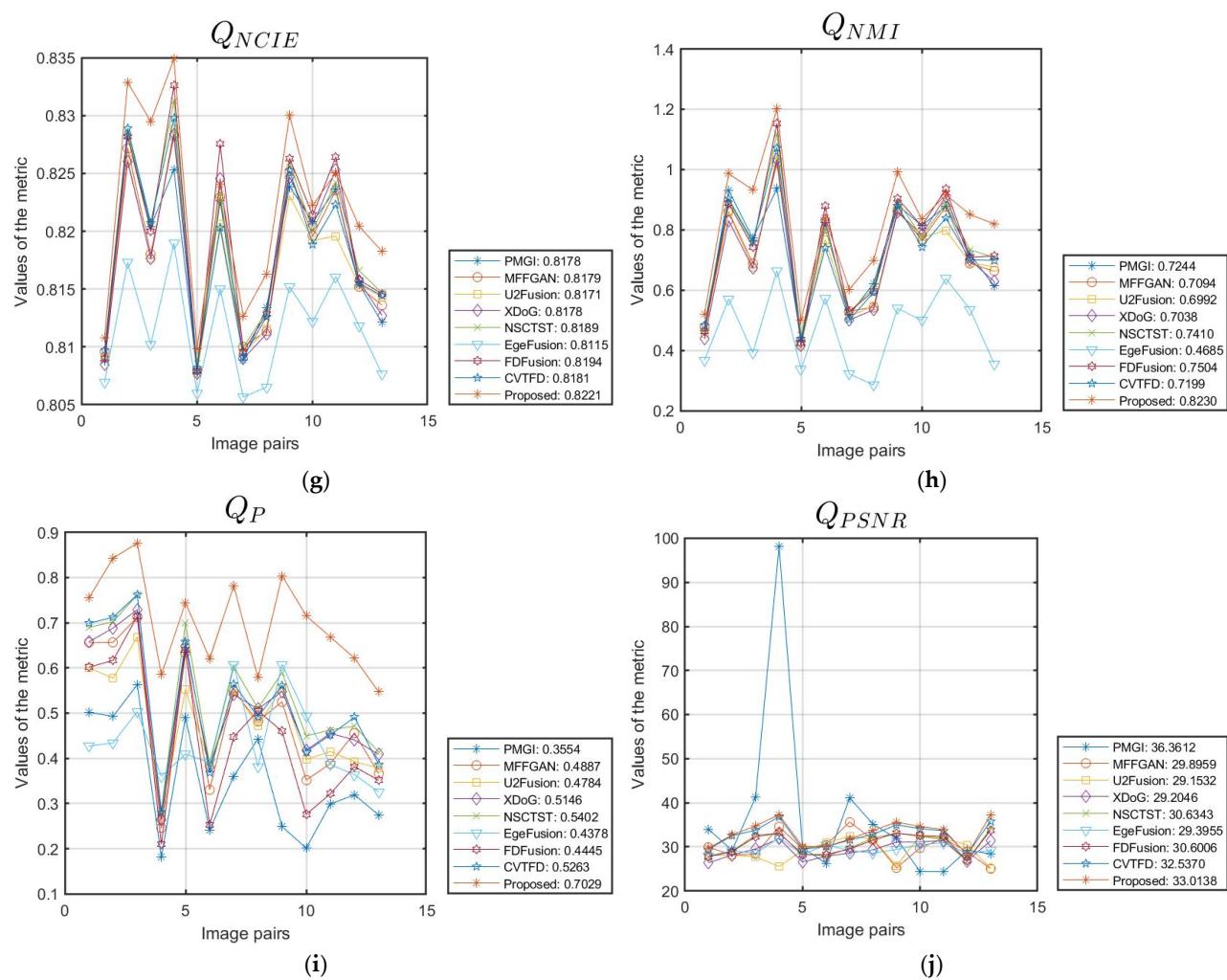


Figure 9. The line chart displays the evaluation metrics across various data from the MFFW dataset. (a) $Q_{AB/F}$; (b) Q_{CB} ; (c) Q_E ; (d) Q_{FMI} ; (e) Q_{MI} ; (f) Q_{MSE} ; (g) Q_{NCIE} ; (h) Q_{NMI} ; (i) Q_P ; (j) Q_{PSNR} .

Table 3. Quantitative comparative analysis of different methods on the MFFW-01 dataset in Figure 6.

	Year	$Q_{AB/F}$	Q_{CB}	Q_E	Q_{FMI}	Q_{MI}	Q_{MSE}	Q_{NCIE}	Q_{NMI}	Q_P	Q_{PSNR}
PMGI	2020	0.5632	0.5018	0.6143	0.8917	4.0362	94.2825	0.8121	0.6143	0.2748	28.3865
MFFGAN	2021	0.5832	0.5732	0.7064	0.9005	4.4393	204.2317	0.8136	0.6631	0.3664	25.0296
U2Fusion	2022	0.5112	0.5385	0.5840	0.8936	4.6178	201.6128	0.8143	0.6769	0.3784	25.0856
XDoG	2023	0.6186	0.5703	0.7572	0.9044	4.1792	47.7266	0.8127	0.6293	0.4087	31.3432
NSCTST	2023	0.6558	0.6336	0.8227	0.9045	4.6722	25.0239	0.8147	0.7088	0.4141	34.1473
EgeFusion	2024	0.2732	0.3643	0.3210	0.8493	2.4442	78.3106	0.8077	0.3565	0.3259	29.1926
FDFusion	2025	0.6066	0.5264	0.7936	0.9033	4.6779	27.4425	0.8146	0.7136	0.3514	33.7466
CVTFD	2025	0.6363	0.6007	0.7992	0.9019	4.6218	17.0483	0.8144	0.6994	0.3871	35.8140
Proposed		0.7272	0.6549	0.8422	0.9092	5.4123	12.1897	0.8183	0.8201	0.5472	37.2709

Table 4. Quantitative average comparative analysis of different methods on the MFFW dataset.

Year	$Q_{AB/F}$	Q_{CB}	Q_E	Q_{FMI}	Q_{MI}	Q_{MSE}	Q_{NCIE}	Q_{NMI}	Q_P	Q_{PSNR}	
PMGI	2020	0.3807	0.5057	0.4245	0.8675	5.0472	80.9829	0.8178	0.7244	0.3554	36.3612
MFFGAN	2021	0.5905	0.5851	0.7557	0.8742	5.0498	82.9576	0.8179	0.7094	0.4887	29.8959
U2Fusion	2022	0.5537	0.5499	0.7076	0.8690	4.8894	94.3035	0.8171	0.6992	0.4784	29.1532
XDoG	2023	0.6090	0.5886	0.7512	0.8745	5.0119	86.1245	0.8178	0.7038	0.5146	29.2046
NSCTST	2023	0.6408	0.6357	0.7976	0.8787	5.2548	65.3151	0.8189	0.7410	0.5402	30.6343
EgeFusion	2024	0.3517	0.4213	0.4581	0.8380	3.3785	77.3397	0.8115	0.4685	0.4378	29.3955
FDFusion	2025	0.5895	0.5710	0.7053	0.8735	5.3144	65.4062	0.8194	0.7504	0.4445	30.6006
CVTFD	2025	0.6290	0.6276	0.8021	0.8796	5.1029	43.0355	0.8181	0.7199	0.5263	32.5370
Proposed		0.7153	0.6771	0.8327	0.8874	5.8296	40.0503	0.8221	0.8230	0.7029	33.0138

4.3. Results on MFI-WHU Dataset

Figures 10 and 11 display the fused images obtained by each method for the MFI-WHU-01 dataset image, along with the corresponding difference maps. From the results, the PMGI generates blurred output, which is clearly observable in the fused image and the associated difference map. The fused images produced by MFFGAN, U2Fusion, XDoG, EgeFusion, and FDFusion fail to adequately preserve background information, resulting in noticeable information loss or artifacts in the corresponding difference images, which indicates poor fusion performance. The NSCTST and CVTFD methods demonstrate relatively good fusion performance; however, the corresponding difference images reveal artifacts or block effects in certain regions, indicating that the fused images do not fully retain the information from the source images. Compared to other algorithms, our method produces fused images that preserve the complete information from the source images without introducing artifacts or noise.

Table 5 presents a quantitative comparison of various methods on the MFI-WHU-01 dataset illustrated in Figure 10. As shown in the table, our algorithm achieves the best performance across 10 evaluation metrics. Both subjective assessments and objective measurements confirm the effectiveness of our method for image fusion on the MFI-WHU dataset. The average performance metrics for each method are depicted in Figure 12 and summarized in Table 6. Except for the $Q_{AB/F}$ and Q_{PSNR} metrics, our algorithm achieves the best average performance across the other eight metrics.

Table 5. Quantitative comparative analysis of different methods on the MFI-WHU-01 dataset in Figure 8.

Year	$Q_{AB/F}$	Q_{CB}	Q_E	Q_{FMI}	Q_{MI}	Q_{MSE}	Q_{NCIE}	Q_{NMI}	Q_P	Q_{PSNR}	
PMGI	2020	0.3835	0.5448	0.4590	0.8658	5.6370	212.2160	0.8209	0.7638	0.5434	24.8630
MFFGAN	2021	0.6371	0.6336	0.8117	0.8748	5.7101	34.0467	0.8212	0.7556	0.7964	32.8101
U2Fusion	2022	0.5167	0.4904	0.7057	0.8616	5.1104	26.7106	0.8184	0.6936	0.5767	33.8640
XDoG	2023	0.6421	0.6913	0.8033	0.8728	5.7376	61.5020	0.8214	0.7563	0.8165	30.2419
NSCTST	2023	0.7124	0.7945	0.8451	0.8814	7.6676	21.9073	0.8332	1.0137	0.8783	34.7249
EgeFusion	2024	0.4040	0.4246	0.5646	0.8372	3.2958	68.7750	0.8119	0.4416	0.5266	29.7565
FDFusion	2025	0.6715	0.7075	0.8344	0.8770	6.3362	36.5794	0.8246	0.8360	0.8186	32.4984
CVTFD	2025	0.7040	0.7841	0.8442	0.8810	7.4901	20.8554	0.8319	0.9900	0.8708	34.9386
Proposed		0.7143	0.8005	0.8452	0.8815	8.0176	20.3441	0.8358	1.0607	0.8817	35.0464



Figure 10. Results of MFI-WHU-01. (a) PMGI; (b) MFFGAN; (c) U2Fusion; (d) XDoG; (e) NSCTST; (f) EgeFusion; (g) FDFusion; (h) CVTFD; (i) Proposed.

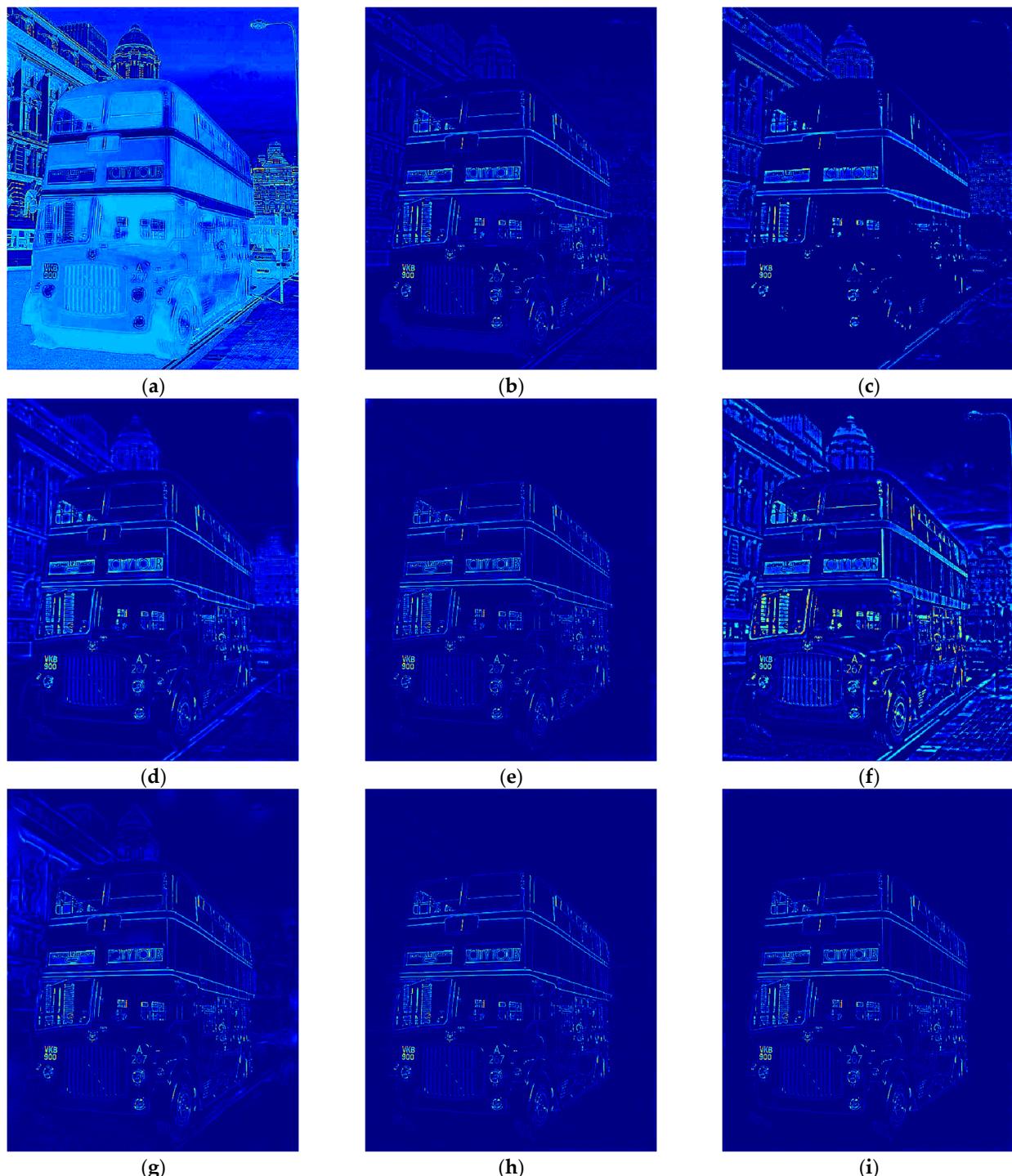
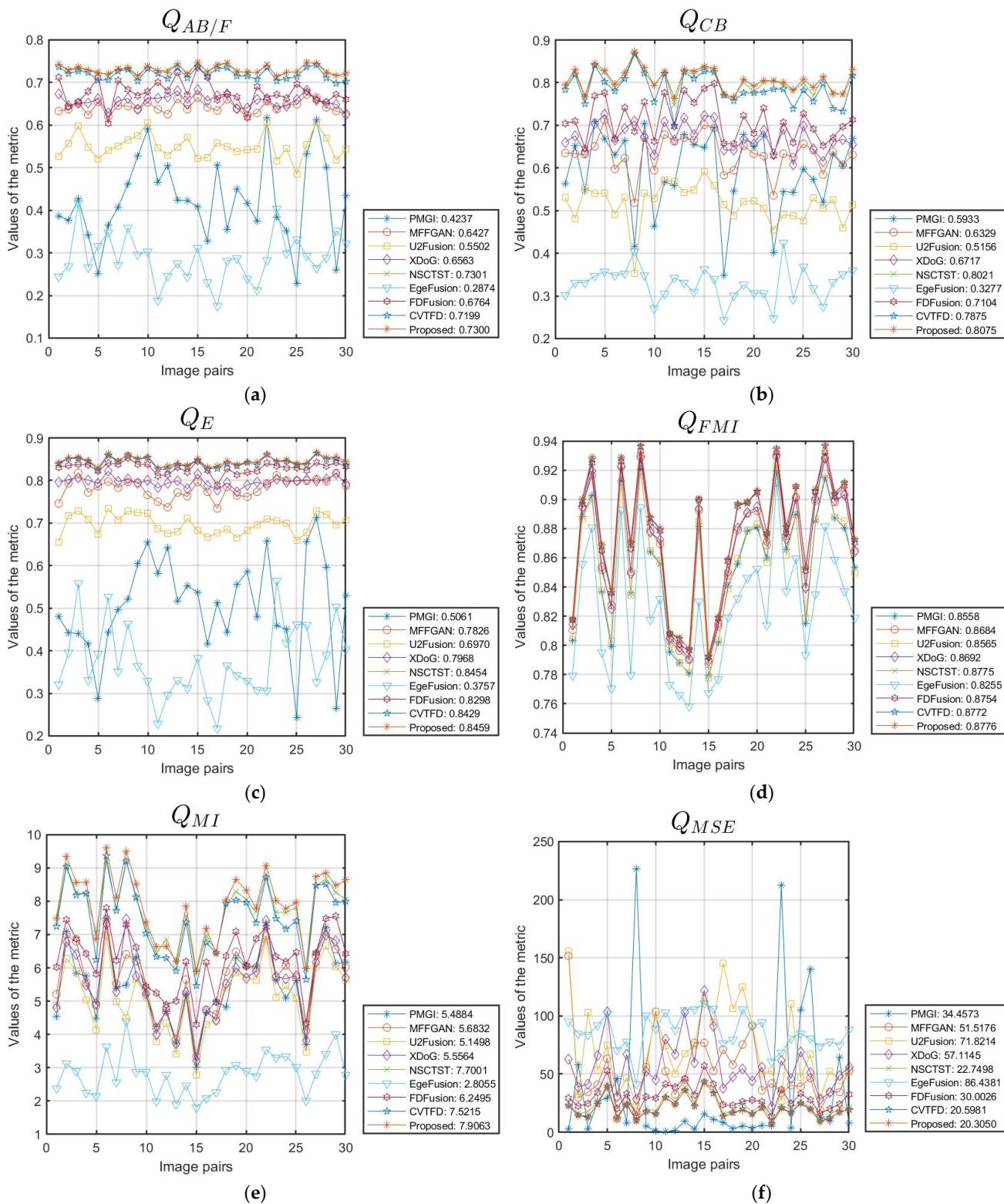


Figure 11. The difference maps of fused images and source image B in Figure 10. (a) PMGI; (b) MFF-GAN; (c) U2Fusion; (d) XDoG; (e) NSCTST; (f) EgeFusion; (g) FDFusion; (h) CVTFD; (i) Proposed.

**Figure 12. Cont.**

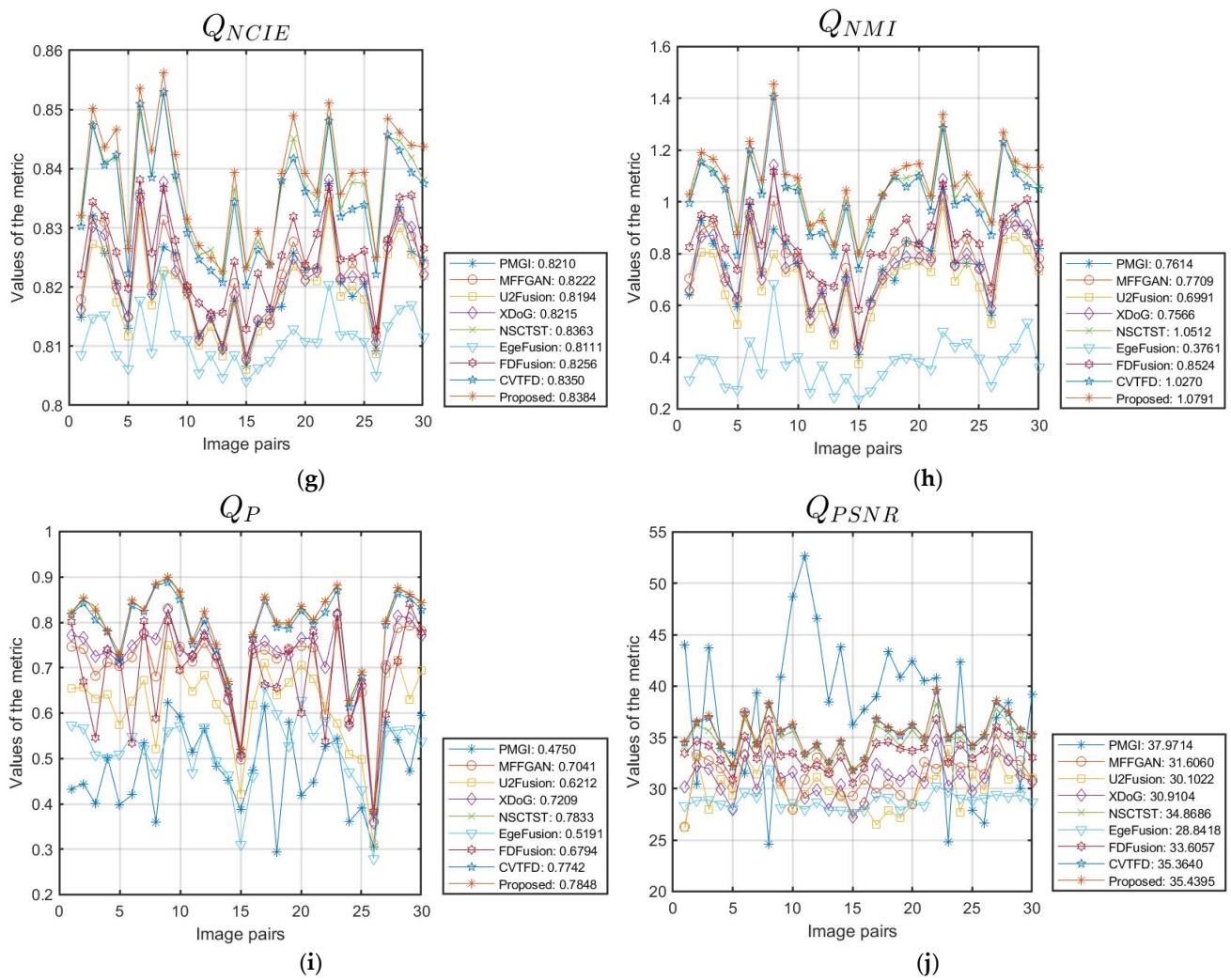


Figure 12. The line chart displays the evaluation metrics across various data from the MFI-WHU dataset. (a) $Q_{AB/F}$; (b) Q_{CB} ; (c) Q_E ; (d) Q_{FMI} ; (e) Q_{MI} ; (f) Q_{MSE} ; (g) Q_{NCIE} ; (h) Q_{NMI} ; (i) Q_P ; (j) Q_{PSNR} .

Table 6. Quantitative average comparative analysis of different methods on the MFI-WHU dataset.

	Year	$Q_{AB/F}$	Q_{CB}	Q_E	Q_{FMI}	Q_{MI}	Q_{MSE}	Q_{NCIE}	Q_{NMI}	Q_P	Q_{PSNR}
PMGI	2020	0.4237	0.5933	0.5061	0.8558	5.4884	34.4573	0.8210	0.7614	0.4750	37.9714
MFFGAN	2021	0.6427	0.6329	0.7826	0.8684	5.6832	51.5176	0.8222	0.7709	0.7041	31.6060
U2Fusion	2022	0.5502	0.5156	0.6970	0.8565	5.1498	71.8214	0.8194	0.6991	0.6212	30.1022
XDoG	2023	0.6563	0.6717	0.7968	0.8692	5.5564	57.1145	0.8215	0.7566	0.7209	30.9104
NSCTST	2023	0.7301	0.8021	0.8454	0.8775	7.7001	22.7498	0.8363	1.0512	0.7833	34.8686
EgeFusion	2024	0.2874	0.3277	0.3757	0.8255	2.8055	86.4381	0.8111	0.3761	0.5191	28.8418
FDFusion	2025	0.6764	0.7104	0.8298	0.8754	6.2495	30.0026	0.8256	0.8524	0.6794	33.6057
CVTFD	2025	0.7199	0.7875	0.8429	0.8772	7.5215	20.5981	0.8350	1.0270	0.7742	35.3640
Proposed		0.7300	0.8075	0.8459	0.8776	7.9063	20.3050	0.8384	1.0791	0.7848	35.4395

4.4. Ablation Experiment

In the ablation study section, we conducted a detailed analysis of the role of consistency verification (CV) in our method. We evaluated the effectiveness of the algorithm by

comparing two models—one without and one with consistency verification—using the average values of 10 evaluation metrics across three datasets: Lytro, MFFW, and MFI-WHU. The corresponding experimental results are shown in Table 7. The data clearly indicate that incorporating consistency verification enhances the performance of multi-focus image fusion, achieving the best results across all evaluation metrics.

Table 7. Ablation study results on consistency verification.

		$Q_{AB/F}$	Q_{CB}	Q_E	Q_{FMI}	Q_{MI}	Q_{MSE}	Q_{NCIE}	Q_{NMI}	Q_P	Q_{PSNR}
Lytro	W/o CV	0.7390	0.7447	0.8815	0.8989	6.9743	21.7820	0.8296	0.9291	0.8173	35.1463
	W/CV	0.7479	0.7515	0.8820	0.9004	7.0950	21.4427	0.8304	0.9455	0.8338	35.2352
MFFW	W/o CV	0.6423	0.6541	0.8141	0.8796	5.2687	42.4645	0.8189	0.7436	0.5467	32.6365
	W/CV	0.7153	0.6771	0.8327	0.8874	5.8296	40.0503	0.8221	0.8230	0.7029	33.0138
MFI-WHU	W/o CV	0.7275	0.8046	0.8459	0.8771	7.7314	20.4173	0.8366	1.0555	0.7834	35.4132
	W/CV	0.7300	0.8075	0.8459	0.8776	7.9063	20.3050	0.8384	1.0791	0.7848	35.4395

4.5. Extended Experiments

In this section, the proposed algorithm is further applied to a range of image fusion tasks, such as infrared and visible fusion [63–65], multi-exposure fusion [66,67], medical image fusion (CT/MRI/PET) [68–70], pan-sharpening (PAN/MS) [71–75], and SAR-optical image fusion [76]. For MRI and PET fusion as well as SAR-optical fusion, we apply RGB and YUV [68] color space transformations to the PET and optical images. For pan-sharpening, we apply RGB and IHS [71] color space transformations to the multispectral images. The fusion results of these extended applications are shown in Figure 13. As illustrated, the proposed algorithm also achieves a favorable performance in these multi-source image fusion tasks, demonstrating its good generalization ability and scalability.

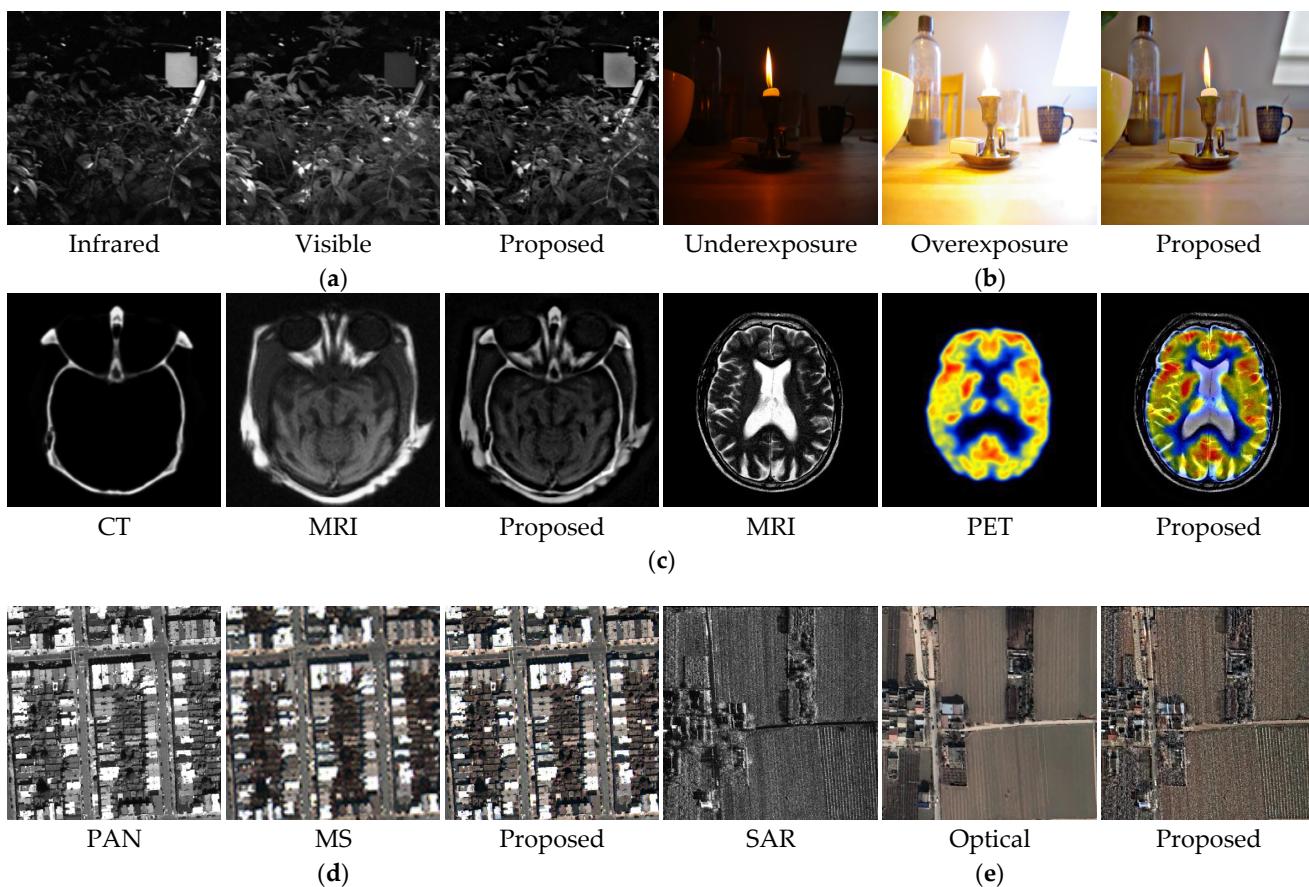


Figure 13. Application of the proposed algorithm in other multi-source image fusion tasks. **(a)** Infrared and visible fusion; **(b)** multi-exposure fusion; **(c)** medical image fusion; **(d)** pan-sharpening; **(e)** SAR and optical fusion.

5. Conclusions

In this study, we introduce a novel multi-focus image fusion algorithm that integrates the dual-channel Rybak neural network (DRYNN) with a consistency verification strategy in the NSCT domain. The primary objective of this method is to enhance the clarity and structural integrity of the fused images while effectively mitigating the impact of noise. To achieve this, the high-frequency sub-bands derived from NSCT decomposition are refined using the DRYNN model in conjunction with consistency verification, allowing for the precise retention of edge and texture details. Meanwhile, the low-frequency sub-bands, which primarily contain coarse information, are fused using a straightforward averaging technique to preserve overall image content. Consistency verification plays a significant role in our algorithm and serves as an innovative aspect of our approach. Ablation experiments have also confirmed the contribution of consistency verification in image fusion.

The proposed fusion framework is rigorously evaluated using three widely recognized public datasets, with performance assessed through both qualitative visual comparisons and quantitative metrics. To gain further insight into the fidelity of the fusion process, difference maps between the fused results and the original input images are generated, highlighting regions of information gain or loss. Additionally, we employ ten standard evaluation metrics to objectively compare our method against existing state-of-the-art fusion algorithms.

Our experimental results clearly demonstrate the superior performance of the proposed algorithm, both in terms of visual perception and numerical accuracy. Looking ahead, we intend to extend this fusion framework to broader application domains such

as change detection [77–81], hyperspectral and multispectral fusion [82–84], and panchromatic and hyperspectral image fusion [85,86], where the need for robust, high-quality fusion techniques is critical for decision-making and analysis. In addition, the efficiency of our algorithm is relatively low and does not meet real-time requirements, which is also a problem we aim to address in our future work.

Author Contributions: The experiments and data collection were conducted by M.L., S.S., Z.J., L.L. and H.M. The manuscript was drafted by M.L., with contributions from the co-authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China Postdoctoral Science Foundation under Grant No. 2024M752692; the Natural Science Foundation of Xinjiang Uygur Autonomous Region under Grant No. 2024D01C240; the National Natural Science Foundation of China under Grant No. 62261053; the Tianshan Talent Training Project-Xinjiang Science and Technology Innovation Team Program (2023TSYCTD0012).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Wang, W.; Deng, L.; Vivone, G. A general image fusion framework using multi-task semi-supervised learning. *Inf. Fusion* **2024**, *108*, 102414. [[CrossRef](#)]
- Wu, X.; Cao, Z.; Huang, T.; Deng, L.; Chanussot, J.; Vivone, G. Fully-connected transformer for multi-source image fusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, *47*, 2071–2088. [[CrossRef](#)] [[PubMed](#)]
- Vivone, G.; Deng, L. Deep learning in remote sensing image fusion: Methods, protocols, data, and future perspectives. *IEEE Geosci. Remote Sens. Mag.* **2025**, *13*, 269–310. [[CrossRef](#)]
- Matteo, C.; Giuseppe, G.; Gemine, V. Hyperspectral pansharpening: Critical review, tools, and future perspectives. *IEEE Geosci. Remote Sens. Mag.* **2025**, *13*, 311–338.
- Jie, Y.; Xu, Y.; Li, X.; Zhou, F.; Lv, J.; Li, H. FS-Diff: Semantic guidance and clarity-aware simultaneous multimodal image fusion and super-resolution. *Inf. Fusion* **2025**, *121*, 103146. [[CrossRef](#)]
- Zhang, X.; Yan, H. Medical image fusion and noise suppression with fractional-order total variation and multi-scale decomposition. *IET Image Process.* **2021**, *15*, 1688–1701. [[CrossRef](#)]
- Jie, Y.; Li, X.; Wang, M.; Zhou, F.; Tan, H. Medical image fusion based on extended difference-of-Gaussians and edge-preserving. *Expert Syst. Appl.* **2023**, *227*, 120301. [[CrossRef](#)]
- Zheng, K.; Cheng, J.; Liu, Y. Unfolding coupled convolutional sparse representation for multi-focus image fusion. *Inf. Fusion* **2025**, *118*, 102974. [[CrossRef](#)]
- Li, B.; Zhang, L.; Liu, J.; Peng, H. Multi-focus image fusion with parameter adaptive dual channel dynamic threshold neural P systems. *Neural Netw.* **2024**, *179*, 106603. [[CrossRef](#)]
- Zhang, X.; He, H.; Zhang, J. Multi-focus image fusion based on fractional order differentiation and closed image matting. *ISA Trans.* **2022**, *129*, 703–714. [[CrossRef](#)]
- Liu, J.; Wu, G.; Liu, Z.; Wang, D.; Jiang, Z.; Ma, L.; Zhong, W.; Fan, X.; Liu, R. Infrared and visible image fusion: From data compatibility to task adaption. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, *47*, 2349–2369. [[CrossRef](#)] [[PubMed](#)]
- Yan, H.; Zhang, J.; Zhang, X. Injected infrared and visible image fusion via L_1 decomposition model and guided filtering. *IEEE Trans. Comput. Imaging* **2022**, *8*, 162–173. [[CrossRef](#)]
- Yan, H.; Zhang, X. Adaptive fractional multi-scale edge-preserving decomposition and saliency detection fusion algorithm. *ISA Trans.* **2020**, *107*, 160–172. [[CrossRef](#)] [[PubMed](#)]
- Zhang, X.; Yan, H.; He, H. Multi-focus image fusion based on fractional-order derivative and intuitionistic fuzzy sets. *Front. Inf. Technol. Electron. Eng.* **2020**, *21*, 834–843. [[CrossRef](#)]
- Li, L.; Zhao, X.; Hou, H.; Zhang, X.; Lv, M.; Jia, Z.; Ma, H. Fractal dimension-based multi-focus image fusion via coupled neural P systems in NSCT domain. *Fractal Fract.* **2024**, *8*, 554. [[CrossRef](#)]
- Zhang, X. Deep learning-based multi-focus image fusion: A survey and a comparative study. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 4819–4838. [[CrossRef](#)]
- Luo, F.; Zhao, B. A review on multi-focus image fusion using deep learning. *Neurocomputing* **2025**, *618*, 129125. [[CrossRef](#)]
- Zhai, H.; Zhang, G.; Zeng, Z.; Xu, Z.; Fang, A. LSKN-MFIF: Large selective kernel network for multi-focus image fusion. *Neurocomputing* **2025**, *635*, 129984. [[CrossRef](#)]

19. Quan, Y.; Wan, X.; Tang, Z.; Liang, J.; Ji, H. Multi-focus image fusion via explicit defocus blur modelling. *Proc. AAAI Conf. Artif. Intell.* **2025**, *39*, 6657–6665. [[CrossRef](#)]
20. Tang, L.; Zhang, H.; Xu, H.; Ma, J. Deep learning-based image fusion: A survey. *J. Image Graph.* **2023**, *28*, 3–36. [[CrossRef](#)]
21. Li, L.; Ma, H.; Jia, Z.; Si, Y. A novel multiscale transform decomposition based multi-focus image fusion framework. *Multimed. Tools Appl.* **2021**, *80*, 12389–12409. [[CrossRef](#)]
22. Li, L.; Si, Y.; Wang, L.; Jia, Z.; Ma, H. A novel approach for multi-focus image fusion based on SF-PAPCNN and ISML in NSST domain. *Multimed. Tools Appl.* **2020**, *79*, 24303–24328. [[CrossRef](#)]
23. Li, L.; Shi, Y.; Lv, M.; Jia, Z.; Liu, M.; Zhao, X.; Zhang, X.; Ma, H. Infrared and visible image fusion via sparse representation and guided filtering in Laplacian pyramid domain. *Remote Sens.* **2024**, *16*, 3804. [[CrossRef](#)]
24. Li, L.; Song, S.; Lv, M.; Jia, Z.; Ma, H. Multi-focus image fusion based on fractal dimension and parameter adaptive unit-linking dual-channel PCNN in curvelet transform domain. *Fractal Fract.* **2025**, *9*, 157. [[CrossRef](#)]
25. Lv, M.; Li, L.; Jin, Q.; Jia, Z.; Chen, L.; Ma, H. Multi-focus image fusion via distance-weighted regional energy and structure tensor in NSCT domain. *Sensors* **2023**, *23*, 6135. [[CrossRef](#)]
26. Li, L.; Lv, M.; Jia, Z.; Ma, H. Sparse representation-based multi-focus image fusion method via local energy in shearlet domain. *Sensors* **2023**, *23*, 2888. [[CrossRef](#)]
27. Zhang, Z.; Li, H.; Xu, T.; Wu, X.; Kittler, J. DDBFusion: An unified image decomposition and fusion framework based on dual decomposition and Bézier curves. *Inf. Fusion* **2025**, *114*, 102655. [[CrossRef](#)]
28. Quan, Y.; Wan, X.; Zheng, T.; Huang, Y.; Ji, H. Dual-path deep unsupervised learning for multi-focus image fusion. *IEEE Trans. Multimed.* **2025**, *27*, 1165–1176. [[CrossRef](#)]
29. Xie, X.; Jiang, Q.; Chen, D.; Guo, B.; Li, P.; Zhou, S. StackMFF: End-to-end multi-focus image stack fusion network. *Appl. Intell.* **2025**, *55*, 503. [[CrossRef](#)]
30. Panigrahy, C.; Seal, A.; Mahato, N.K. Fractal dimension based parameter adaptive dual channel PCNN for multi-focus image fusion. *Opt. Lasers Eng.* **2020**, *133*, 106141. [[CrossRef](#)]
31. Joshua, A.; Balasubramaniam, P. An adaptive low-light image enhancement method via fusion of a new intuitionistic fuzzy generator and fractal-fractional derivative. *Signal Image Video Process.* **2025**, *19*, 233. [[CrossRef](#)]
32. Xian, Y.; Zhao, G. Multi-focus image fusion based on visual depth and fractional-order differentiation operators embedding convolution norm. *Signal Process.* **2025**, *233*, 109955. [[CrossRef](#)]
33. Lu, J.; Tan, K. Multi-focus image fusion using residual removal and fractional order differentiation focus measure. *Signal Image Video Process.* **2024**, *18*, 3395–3410. [[CrossRef](#)]
34. Li, X.; Chen, H. Multi-focus image fusion via adaptive fractional differential and guided filtering. *Multimed. Tools Appl.* **2024**, *83*, 32923–32943. [[CrossRef](#)]
35. Yu, L.; Zeng, Z. Fractional-order differentiation based sparse representation for multi-focus image fusion. *Multimed. Tools Appl.* **2022**, *81*, 4387–4411. [[CrossRef](#)]
36. Ouyang, Y.; Zhai, H.; Hu, H. FusionGCN: Multi-focus image fusion using superpixel features generation GCN and pixel-level feature reconstruction CNN. *Expert Syst. Appl.* **2025**, *262*, 125665. [[CrossRef](#)]
37. Li, S.; Huang, S. AFA-Mamba: Adaptive feature alignment with global-local Mamba for hyperspectral and LiDAR data classification. *Remote Sens.* **2024**, *16*, 4050. [[CrossRef](#)]
38. Li, H.; Shen, T.; Zhang, Z.; Zhu, X.; Song, X. EDMF: A new benchmark for multi-focus images with the challenge of exposure difference. *Sensors* **2024**, *24*, 7287. [[CrossRef](#)]
39. Zhang, H.; Le, Z. MFF-GAN: An unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion. *Inf. Fusion* **2021**, *66*, 40–53. [[CrossRef](#)]
40. Avci, D.; Sert, E.; Özyurt, F.; Avci, E. MFIF-DWT-CNN: Multi-focus image fusion based on discrete wavelet transform with deep convolutional neural network. *Multimed. Tools Appl.* **2024**, *83*, 10951–10968. [[CrossRef](#)]
41. Goyal, N.; Goyal, N. Dual-channel Rybak neural network based medical image fusion. *Opt. Laser Technol.* **2025**, *181*, 112018. [[CrossRef](#)]
42. Qi, Y.; Yang, Z.; Lian, J.; Guo, Y.; Sun, W.; Liu, J.; Wang, R.; Ma, Y. A new heterogeneous neural network model and its application in image enhancement. *Neurocomputing* **2021**, *440*, 336–350. [[CrossRef](#)]
43. Sinha, A.; Agarwal, R.; Kumar, V.; Garg, N.; Pundir, D.S.; Singh, H.; Rani, R.; Panigrahy, C. Multi-modal medical image fusion using improved dual-channel PCNN. *Med. Biol. Eng. Comput.* **2024**, *62*, 2629–2651. [[CrossRef](#)] [[PubMed](#)]
44. Li, X.; Zhou, F.; Tan, H. Multi-focus image fusion based on nonsubsampled contourlet transform and residual removal. *Signal Process.* **2021**, *184*, 108062. [[CrossRef](#)]
45. Zafar, R.; Farid, M.S.; Khan, M.H. Multi-focus image fusion: Algorithms, evaluation, and a library. *J. Imaging* **2020**, *6*, 60. [[CrossRef](#)]
46. Nejati, M.; Samavi, S.; Shirani, S. Multi-focus image fusion using dictionary-based sparse representation. *Inf. Fusion* **2015**, *25*, 72–84. [[CrossRef](#)]

47. Xu, S.; Wei, X.; Zhang, C. MFFW: A newdataset for multi-focus image fusion. *arXiv* **2020**, arXiv:2002.04780.
48. Zhang, H.; Xu, H.; Xiao, Y. Rethinking the image fusion: A fast unified image fusion network based on proportional maintenance of gradient and intensity. In Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI), New York, NY, USA, 7–12 February 2020; Volume 34, pp. 12797–12804.
49. Xu, H.; Ma, J.; Jiang, J. U2Fusion: A unified unsupervised image fusion network. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 502–518. [[CrossRef](#)]
50. Tang, H.; Liu, G.; Qian, Y. EgeFusion: Towards edge gradient enhancement in infrared and visible image fusion with multi-scale transform. *IEEE Trans. Comput. Imaging* **2024**, *10*, 385–398. [[CrossRef](#)]
51. Jie, Y.; Li, X.; Tan, T.; Yang, L.; Wang, M. Multi-modality image fusion using fuzzy set theory and compensation dictionary learning. *Opt. Laser Technol.* **2025**, *181*, 112001. [[CrossRef](#)]
52. Qu, X.; Yan, J.; Xiao, H. Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain. *Acta Autom. Sin.* **2008**, *34*, 1508–1514. [[CrossRef](#)]
53. Liu, Z.; Blasch, E.; Xue, Z. Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 94–109. [[CrossRef](#)]
54. Haghigat, M.; Razian, M. Fast-FMI: Non-reference image fusion metric. In Proceedings of the IEEE 8th International Conference on Application of Information and Communication Technologies, Astana, Kazakhstan, 15–17 October 2014; pp. 424–426.
55. Chen, F.; Sha, Y.; Ji, H.; Peng, K.; Liang, X. Integrating multifractal features into machine learning for improved prediction. *Fractal Fract.* **2025**, *9*, 205. [[CrossRef](#)]
56. Li, J.; Wang, C.; Su, W.; Ye, D.; Wang, Z. Uncertainty-aware self-attention model for time series prediction with missing values. *Fractal Fract.* **2025**, *9*, 181. [[CrossRef](#)]
57. Bi, X.; Qie, R.; Tao, C.; Zhang, Z.; Xu, Y. Unsupervised multimodal UAV image registration via style transfer and cascade network. *Remote Sens.* **2025**, *17*, 2160. [[CrossRef](#)]
58. Wadood, A.; Albalawi, H.; Alatwi, A.M.; Anwar, H.; Ali, T. Design of a novel fractional whale optimization-enhanced support vector regression (FWOA-SVR) model for accurate solar energy forecasting. *Fractal Fract.* **2025**, *9*, 35. [[CrossRef](#)]
59. Zhang, X.; Dai, L. Image enhancement based on rough set and fractional order differentiator. *Fractal Fract.* **2022**, *6*, 214. [[CrossRef](#)]
60. Zhang, X.; Liu, R.; Ren, J.; Gui, Q. Adaptive fractional image enhancement algorithm based on rough set and particle swarm optimization. *Fractal Fract.* **2022**, *6*, 100. [[CrossRef](#)]
61. Zhang, X.; Boutat, D.; Liu, D. Applications of fractional operator in image processing and stability of control systems. *Fractal Fract.* **2023**, *7*, 359. [[CrossRef](#)]
62. Tabirca, A.I.; Dumitrescu, C.; Radu, V. Enhancing banking transaction security with fractal-based image steganography using fibonacci sequences and discrete wavelet transform. *Fractal Fract.* **2025**, *9*, 95. [[CrossRef](#)]
63. Li, L.; Lv, M.; Jia, Z.; Jin, Q.; Liu, M.; Chen, L.; Ma, H. An effective infrared and visible image fusion approach via rolling guidance filtering and gradient saliency map. *Remote Sens.* **2023**, *15*, 2486. [[CrossRef](#)]
64. Heredia-Aguado, E.; Cabrera, J.J.; Jiménez, L.M.; Valiente, D.; Gil, A. Static early fusion techniques for visible and thermal images to enhance convolutional neural network detection: A performance analysis. *Remote Sens.* **2025**, *17*, 1060. [[CrossRef](#)]
65. Li, L.; Ma, H. Saliency-guided nonsubsampled shearlet transform for multisource remote sensing image fusion. *Sensors* **2021**, *21*, 1756. [[CrossRef](#)] [[PubMed](#)]
66. Zhang, X. Benchmarking and comparing multi-exposure image fusion algorithms. *Inf. Fusion* **2021**, *74*, 111–131. [[CrossRef](#)]
67. Zhang, W.; Wang, C.; Zhu, J. MEF-CAAN: Multi-exposure image fusion based on a low-resolution context aggregation attention network. *Sensors* **2025**, *25*, 2500. [[CrossRef](#)]
68. Yin, M.; Liu, X.; Liu, Y.; Chen, X. Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain. *IEEE Trans. Instrum. Meas.* **2019**, *68*, 49–64. [[CrossRef](#)]
69. Zhu, Z.; Wang, Z.; Qi, G.; Mazur, N.; Yang, P.; Liu, Y. Brain tumor segmentation in MRI with multi-modality spatial information enhancement and boundary shape correction. *Pattern Recognit.* **2024**, *153*, 110553. [[CrossRef](#)]
70. Zhu, Z.; He, X.; Qi, G.; Li, Y.; Cong, B.; Liu, Y. Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal MRI. *Inf. Fusion* **2023**, *91*, 376–387. [[CrossRef](#)]
71. Liu, Y.; Wang, Z. A practical pan-sharpening method with wavelet transform and sparse representation. In Proceedings of the 2013 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, China, 22–23 October 2013; pp. 288–293.
72. Vivone, G.; Dalla Mura, M.; Garzelli, A.; Restaino, R.; Scarpa, G. A new benchmark based on recent advances in multispectral pansharpening: Revisiting pansharpening with classical and emerging pansharpening methods. *IEEE Geosci. Remote Sens. Mag.* **2021**, *9*, 53–81. [[CrossRef](#)]
73. Vivone, G. Robust band-dependent spatial-detail approaches for panchromatic sharpening. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6421–6433. [[CrossRef](#)]
74. Wen, X.; Ma, H.; Li, L. A three-branch pansharpening network based on spatial and frequency domain interaction. *Remote Sens.* **2025**, *17*, 13. [[CrossRef](#)]

75. Wen, X.; Ma, H.; Li, L. A multi-stage progressive pansharpening network based on detail injection with redundancy reduction. *Sensors* **2024**, *24*, 6039. [[CrossRef](#)] [[PubMed](#)]
76. Li, J.; Zhang, J.; Yang, C.; Liu, H.; Zhao, Y.; Ye, Y. Comparative analysis of pixel-level fusion algorithms and a new high-resolution dataset for SAR and optical image fusion. *Remote Sens.* **2023**, *15*, 5514. [[CrossRef](#)]
77. Li, L.; Ma, H.; Jia, Z. Multiscale geometric analysis fusion-based unsupervised change detection in remote sensing images via FLICM model. *Entropy* **2022**, *24*, 291. [[CrossRef](#)]
78. Chen, Z.; Chen, H.; Leng, J.; Zhang, X.; Gao, Q.; Dong, W. VMMCD: VMamba-based multi-scale feature guiding fusion network for remote sensing change detection. *Remote Sens.* **2025**, *17*, 1840. [[CrossRef](#)]
79. Li, L.; Ma, H.; Jia, Z. Change detection from SAR images based on convolutional neural networks guided by saliency enhancement. *Remote Sens.* **2021**, *13*, 3697. [[CrossRef](#)]
80. Zhong, H.; Wu, C.; Xiao, Z. LRNet: Change detection in high-resolution remote sensing imagery via a localization-then-refinement strategy. *Remote Sens.* **2025**, *17*, 1849. [[CrossRef](#)]
81. Li, L.; Ma, H.; Zhang, X.; Zhao, X.; Lv, M.; Jia, Z. Synthetic aperture radar image change detection based on principal component analysis and two-level clustering. *Remote Sens.* **2024**, *16*, 1861. [[CrossRef](#)]
82. He, Y.; Li, H.; Zhang, M.; Liu, S.; Zhu, C.; Xin, B.; Wang, J.; Wu, Q. Hyperspectral and multispectral remote sensing image fusion based on a retractable spatial–spectral transformer network. *Remote Sens.* **2025**, *17*, 1973. [[CrossRef](#)]
83. Li, J.; Zheng, K.; Gao, L.; Han, Z.; Li, Z.; Chanussot, J. Enhanced deep image prior for unsupervised hyperspectral image super-resolution. *IEEE Trans. Geosci. Remote Sens.* **2025**, *63*, 5504218. [[CrossRef](#)]
84. Vivone, G. Multispectral and hyperspectral image fusion in remote sensing: A survey. *Inf. Fusion* **2023**, *89*, 405–417. [[CrossRef](#)]
85. Vivone, G.; Garzelli, A.; Xu, Y.; Liao, W.; Chanussot, J. Panchromatic and hyperspectral image fusion: Outcome of the 2022 WHISPERS hyperspectral pansharpening challenge. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2023**, *16*, 166–179. [[CrossRef](#)]
86. Shen, X.; Chen, L.; Liu, H.; Zhou, X.; Vivone, G.; Chanussot, J. Iteratively regularizing hyperspectral and multispectral image fusion with framelets. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *18*, 5331–5346. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.