

Research Article

The Fusion of Multi-Focus Images Based on the Complex Shearlet Features-Motivated Generative Adversarial Network

Lei Wang ¹, ZhouQi Liu ¹, Jin Huang ¹, Cong Liu ¹, LongBo Zhang ¹
and ChunXiang Liu ²

¹School of Computer Science and Technology, Shandong University of Technology, Zibo 255000, China

²Anhui Key Laboratory of Plant Resources and Plant Biology, Huaibei Normal University, Huaibei 235000, China

Correspondence should be addressed to Lei Wang; wanglei0511@sdut.edu.cn

Received 29 June 2021; Accepted 15 July 2021; Published 30 July 2021

Academic Editor: Chi-Hua Chen

Copyright © 2021 Lei Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The traditional methods for multi-focus image fusion, such as the typical multi-scale geometric analysis theory-based methods, are usually restricted by sparse representation ability and the transferring efficiency of the fusion rules for the captured features. Aiming to integrate the partially focused images into the fully focused image with high quality, the complex shearlet features-motivated generative adversarial network is constructed for multi-focus image fusion in this paper. Different from the popularly used wavelet, contourlet, and shearlet, the complex shearlet provides more flexible multiple scales, anisotropy, and directional sub-bands with the approximate shift invariance. Therefore, the features in complex shearlet domain are more effective. With the help of the generative adversarial network, the whole procedure of multi-focus fusion is modeled to be the process of adversarial learning. Finally, several experiments are implemented and the results prove that the proposed method outperforms the popularly used fusion algorithms in terms of four typical objective metrics and the comparison of visual appearance.

1. Introduction

The target information may display the differentiation for the lengths of the focus during the imaging procedure, that is, the closer the object to the focus is, the clearer the image is. On the other hand, it is difficult to synchronously get the full-focus image by only one imaging device [1]. A common method to deal with this problem is to fuse multiple images of the same scene into images of different focal lengths, which is called the multi-focus image fusion and has been widely used in military monitoring, image analysis, and transportation [2]. For example, in modern wars, the multi-focus images can be used to monitor important targets and facilities of the enemy, and in the transportation domain, the multi-focus images can be used to track logistics and vehicle information and even penalize violations.

Nowadays, there are mainly four kinds of strategies for the fusion of multi-focus images: the spatial domain methods, the early transform domain methods, the multi-scale geometric analysis theory-based methods, and the deep

learning theory-based methods. The spatial domain methods usually directly implement the linear computation on the image pixel, for example, the averaging method, maxing method, and weighted method. The early transform domain methods include the Laplace pyramid-based method, wavelet-based method, etc.

In these methods, the multi-focus images are decomposed into different scales and each scale is with a limited number of sub-bands. Then, the features in different levels can be obtained for fusion. For example, in reference [3], the authors proposed a fusion method by using the extremum of the wavelet coefficients in different sub-bands. Dou et al. [4] proposed a fusion method by using the region energy in different high-pass sub-band coefficients by considering their distributions. Due to the limited number of high sub-bands, some block artifacts of edges may appear in these methods. To deal with these problems, multi-scale geometric analysis theory-based methods have been popularly reported in recent years. The curvlet transform, contourlet transform, non-subsampled contourlet transform (NSCT), and the

shearlet are the typical decomposition tools in this period. For example, Li and Yang [5] proposed a fusion method by combining the wavelet and curvelet to overcome the disadvantages of the wavelet. He et al. [6] proposed a multi-focus image fusion method based on the improved contourlet package. Qu et al. [7] proposed the spatial frequency-motivated PCNN model in NSCT domain, and the spatial frequency was used to implement the firing mapping in the method. Liao et al. [8] proposed a shearlet-based fusion method by employing the statistical information of the shearlet coefficients. Considering the fusion procedure of the aforementioned methods, it is obvious that the fusion results are highly determined by the performance of the decomposition abilities.

From the viewpoint of fusion rules, the fusion procedure of the multi-scale geometric analysis theory-based methods can be modeled to be the classification problem of the multi-scale transformation coefficients. There are also three typical categories for the fusion rules: the active level metric-based rule, the kernel learning-based fusion rule, and the neural network-based fusion rule. For the former, the algebraic operations, such as averaging and maxing, are popularly used. The second one includes the ICA, SVM, and PCA. In literature [9], the principal component analysis (PCA) is implemented in dictionary training to reduce the dimension of transform coefficients. In literature [10], the cartoon components and the texture components are combined by ICA. The artifacts are easy to produce for the classification determined by the simple computations for the single coefficient. The neural network-based rule has been popularly reported in recent years, and some good results have been obtained. For example, in literature [11], the PCNN model is used to be the fusion rule by combining the NSCT together. Though good results have been obtained, these models are not abstract enough, which means the features are all in low levels. So, advanced neural network models should be further developed.

Compared with the traditional neural networks, the deeply coupled neural network is the breakthrough in this domain and has been popularly applied in image denoising, image recognition and classification, and in image fusion [12]. For example, a novel image fusion method for the multi-focus image was proposed based on the support value-motivated deep convolutional neural network model in literature [13]; a general multi-focus image fusion framework, called IF-CNN, was developed based on the deeply convolutional neural network in literature [14]. The MFF-GAN and Pan-GAN, under the mechanism of the unsupervised generative adversarial network, are proposed in the literature [15] and [16], respectively, and detail preserving adversarial learning model is proposed in literature [17, 18]. Specially, according to some recent references, good results are always obtained by the GAN-based methods. The main reason lies in its unique characteristics: firstly, the GAN model has more complex and deep network structure than the commonly used neural network; secondly, modeling the fusion process into the adversarial learning is more in line with the general principle of human understanding of the

world. However, the common characteristics of the well-known methods are that they are directly developed in the pixel level, the important image features are not carefully used.

In order to overcome the shortcomings of the above methods, a multi-focus image fusion method based on the GAN in the complex shearlet domain is developed. Different from the traditional transformation methods, such as the curvelet and the contourlet, the complex shearlet can divide the source images into high-pass and low-pass sub-bands to provide more useful features. Besides, the computational efficiency of complex shearlet is higher than that of the NSCT to get the same shift invariance. With the help of the GAN, the whole fusion can be modeled by the adversarial learning of the features in the complex shearlet. Therefore, better fusion results can be obtained in the feature level.

The rest of this paper is organized as follows. The details of the whole method are given in Section 2. Experimental results and some important discussions are given in Section 3. Finally, the paper is concluded in Section 4.

2. Methodology

Figure 1 shows the structure of the proposed method. Firstly, the images to be fused are input into the GAN, and at the same time, the complex shearlet is implemented to get the high-pass subbands for them. Then, the features in the complex shearlet domain are computed to produce the new form of the loss function. The loss function is updated to drive the training of the GAN, and the final fusion results can be obtained after the training is finished.

2.1. The Complex Shearlet Transform. As one of the most famous multi-scale geometric transformation tools, the complex shearlet transform can extract directional information of different scales and deliver highly sparse approximations of the 2D signals. Generally speaking, it divides the source images into low-pass and high-pass sub-band images in different levels, i.e., the approximately sparse representation of the source images and the obvious feature information of the images. Different from the discrete wavelet, contourlet, and shearlet, the complex shearlet is realized based on the multi-scale pyramid filters and the Hilbert transform [19, 20]. The former gives the multiple partitions of the image, and the latter provides the directional sub-bands in the complex space. Figure 2 gives an example of the complex shearlet transform on a “clock with the left focus.”

2.2. The Feature in the High-Pass Sub-Bands. After the complex shearlet transform is done, the shearlet coefficients with large absolute values are considered to be the sharp brightness or salient features, meaning that they are the focused regions in the source images. Considering the aim of bringing the focus to the fused image, it needs to extract the focus firstly by using the complex shearlet coefficients.

On the other hand, the features in the multi-focus images can be uniformly described by the activity-level

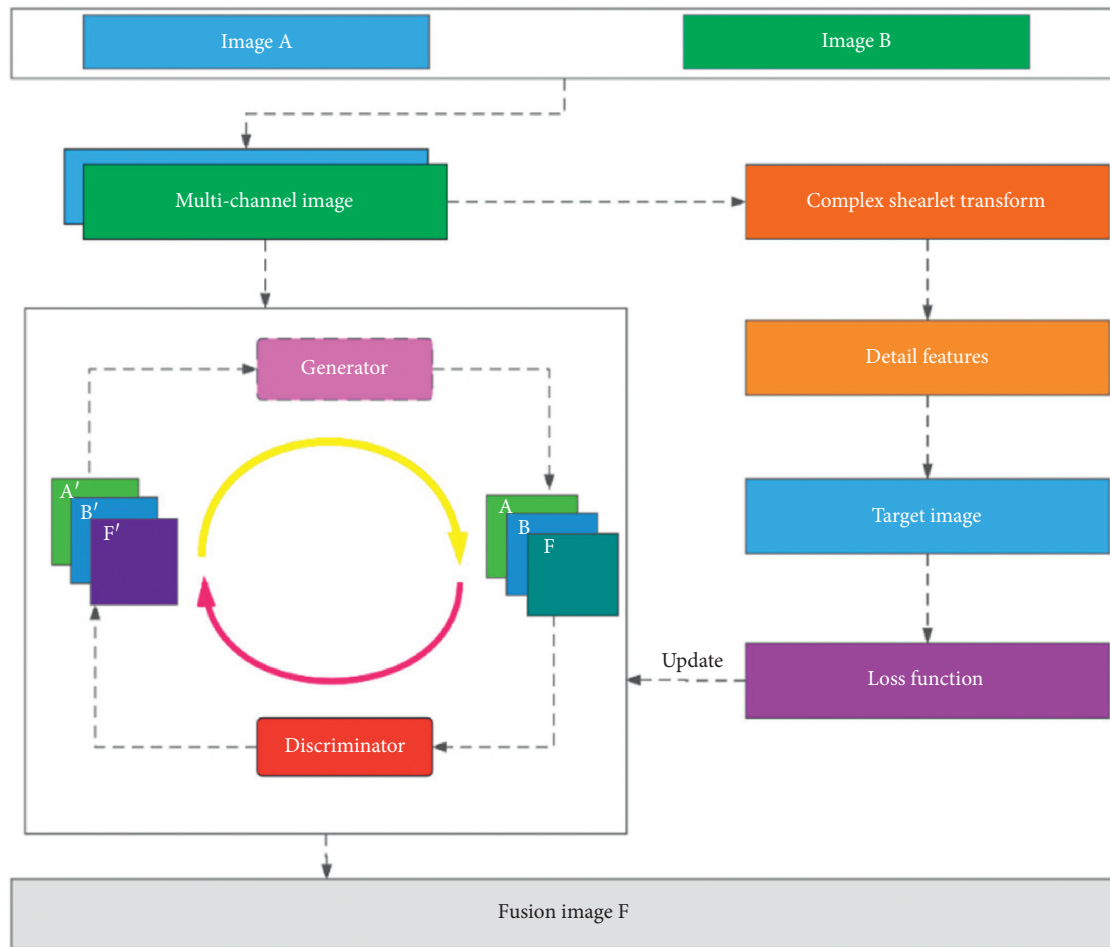


FIGURE 1: The architecture of the proposed method.

measurements, such as the local energy, standard deviation, and spatial frequency [21, 22]. For the above reasons, local energy and spatial frequency are used to represent the important features in the high-pass coefficients. Furthermore, different from their common form used in other literatures, they are computed in multiple scales and directions.

2.3. The GAN Model

2.3.1. The Structure of the GAN. Usually, the complete structure of the GAN network consists of two parts: the generator and the discriminator [23, 24]. The detailed structure of the GAN model used in this paper is shown in Figures 3 and 4.

For the generator, five convolutional layers are used to extract features. A 5×5 convolution kernel is used in the first convolutional layer, and a 3×3 convolution kernel is used for the other four layers. The inputs of each layer are connected by the outputs of all the previous layers, with the aim of speeding up the convergence and improving the stability of the model [25]. All the activation functions are set to be "ReLU," i.e., rectified linear unit. Furthermore, layer normalization (BN) is also employed to preserve the contrast

information of the source images. It calculates the average value of all the dimensional inputs in each layer and finally implements the normalization operation. The advantages are to reduce the sensitivity of initializing data and effectively avoid the gradient disappearance problem [26].

Different from the generator, the main propose of the discriminator is to give the decision by classification. As shown in Figure 4, the discriminator has the same structure with the convolutional neural network which has two inputs, i.e., the Laplacian joint enhanced image from the source images and the fused image from the generator. Four layers of the 3×3 filters are designed to implement the convolution to capture the feature information. Meanwhile, in order to reduce the loss of important information caused by using the downsampling scheme, the activation function is set to be "ReLU." Finally, the fully connected layer is used to classify, and the sigmoid function is employed to output the final results.

2.3.2. The Loss Function. The loss function plays the role of minimizing the loss of the training to get the ideal model, and it usually consists of the generator loss function and discriminator loss function, as shown in the following formula:

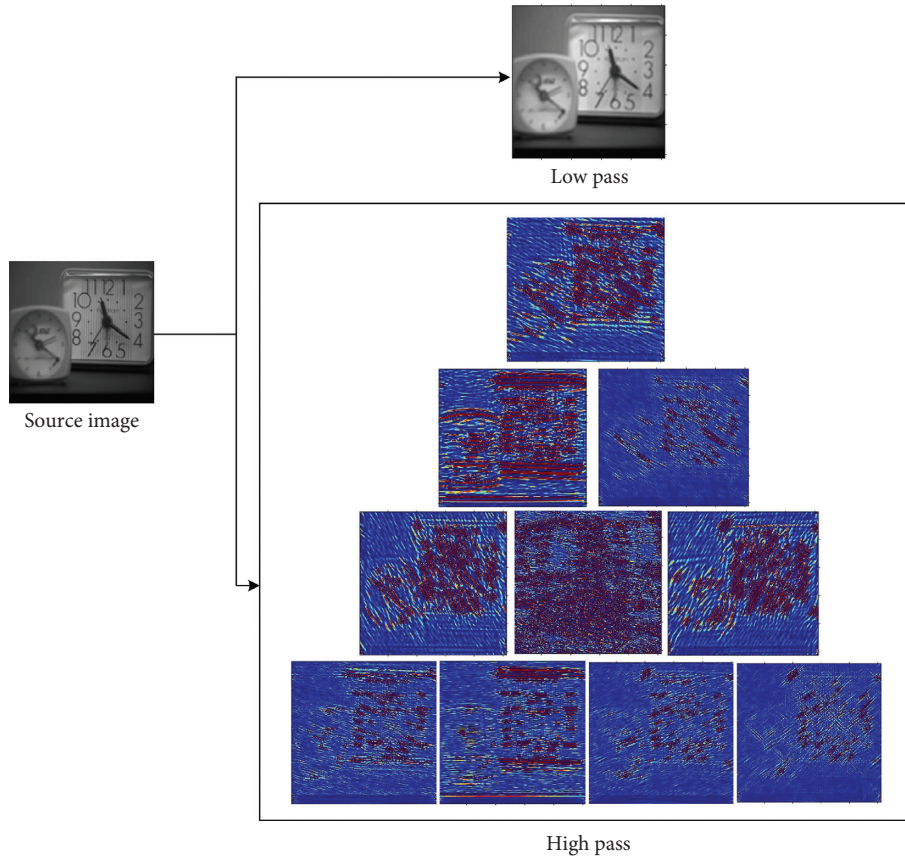


FIGURE 2: An example of the complex shearlet transformation.

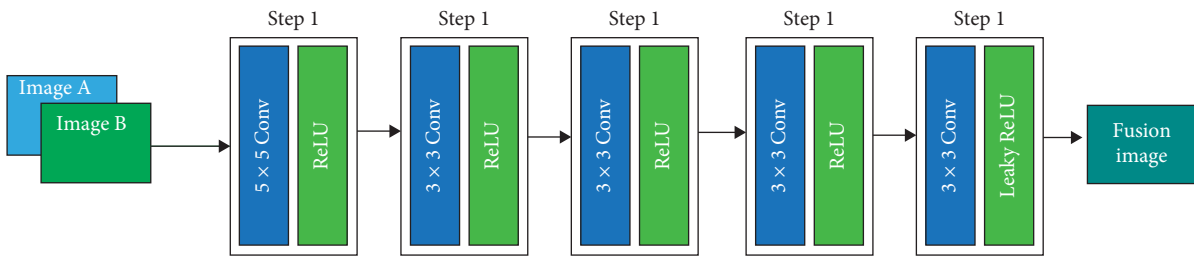


FIGURE 3: The structure of the generator in GAN.

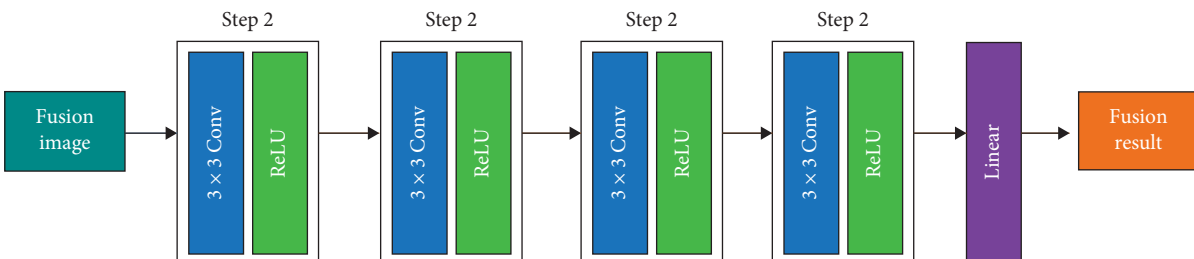


FIGURE 4: The structure of the discriminator in GAN.

$$L_{\text{GAN}} = \{\min(L_G), \min(L_D)\}, \quad (1)$$

where L_G , L_D are the generator loss function and the discriminator loss function, respectively.

According to the original model, L_G is defined by formula (2). It is computed by the summarizing the confrontation loss R and the content loss C from the procedure of the image generation.

$$L_G = R + \lambda C, \quad (2)$$

$$R = \frac{1}{N} \sum_{n=1}^N (D(I_f^n - \alpha))^2, \quad (3)$$

$$C = \mu_1 f_{\text{in}} + \mu_2 f_{\text{grad}}, \quad (4)$$

where λ is a balanced weight between R and C , I_f^n is the target image to be fused, N is the number of fused images, $D(I_f^n)$ is the result of classification, α means that the false data is recognized to be true by the discriminator, f_{in} is the intensity loss, f_{grad} is the gradient loss, and μ is the balanced weight.

L_D can be expressed as

$$L_D = \frac{1}{N} \sum_{n=1}^N [D(G_{I_f}) - b]^2 + [D(G_{\text{joint}}) - c]^2, \quad (5)$$

where G_{joint} is the joint Laplace enhanced gradient map, G_{I_f} is the gradient map of the fused image, and b and c are their labels, respectively.

From all the formulas above, it can be seen that target image to be fused is very important in confrontation learning. The common way to get it is to average the images to be fused or let it be initialized by one of the images to be fused. The drawback is that it is far from the final results and should spend much time and resource to get the optimal decision during the confrontation.

Therefore, a new form of the target image is proposed. Let $C_\lambda(a, b)$ be the low-pass sub-band coefficient at position (a, b) , $\lambda = A, B$. The low-pass coefficient of the fused image can be obtained by

$$C_F(a, b) = \begin{cases} C_A(a, b), & E_A(a, b) \geq E_B(a, b), \\ C_B(a, b), & E_A(a, b) < E_B(a, b), \end{cases} \quad (6)$$

where E_λ is the local energy computed in the 3×3 neighborhood.

Let $C_\lambda^{l,k}(i, j)$ be the high-pass coefficient at (i, j) in the l -th sub-band and the k -th level after implementing the complex shearlet transformation, $\lambda = A, B$; then, the high-pass coefficient of the fused image can be obtained by

$$C_F^{l,k}(i, j) = \begin{cases} C_A^{l,k}(i, j), & S_A^{l,k}(i, j) \geq S_B^{l,k}(i, j), \\ C_B^{l,k}(i, j), & S_A^{l,k}(i, j) < S_B^{l,k}(i, j), \end{cases} \quad (7)$$

where s_λ is the spatial frequency which can be computed by the following formula:

$$s_\lambda = \sum_{q=1, q \neq k}^K \sum_{p=1, l \neq p}^L s(C_\lambda^{l,k}, C_\lambda^{p,q}). \quad (8)$$

Then, the target image I_f^n can be obtained by applying the inversion of the complex shearlet transform on the fused low-pass and high-pass sub-bands.

3. Experimental Results and Analysis

The experiments are implemented to show the performance of the proposed method. The platform used is Inspur big data server NF5280M4 with Intel Xeon CPU and 256 GB RAM. 100 pairs of multi-focus images are used for the training. All the data can be downloaded from the web [27–30].

Seven typical methods, i.e., PCNN-based method (PCNN for short) [31], the contourlet-based method (contourlet for short) [32], the GAN-based method (GAN for short) [17], the DCNN-based method (DCNN for short) [33], the discrete shearlet-based method (shearlet for short) [34], the convolutional sparse representation-based method (CSR for short) [35], and sparse representation and sum modified-Laplacian-based method (SR-SML for short) [36], are implemented. The level of the complex shearlet is four.

So far, how to evaluate the quality of the fusion results is still a confusing question. Subjectively visual and objectively quantitative comparison is the mainstream practice in this domain. Without loss of generality, mutual information, entropy, standard deviation (MI, En, and SD for short, respectively), and $Q^{AB/F}$ are selected to be the metrics. The greater their value, the better the fusion images [37–39].

To save space, only “Pepsi-Cola,” “Plane,” “Clocks,” “Flower,” “Cup,” and “Calendar” are shown in Figure 5. All the fusion results are shown in Figures 6–11. In Figure 7, the middle part of the “Plane” is partially enlarged to compare the local detail features.

From the above methods, we can see that though the focus regions are expressed better than the source image, the fusion results are different from each other. For the PCNN-based method, blurred edges obviously occur, and so the details are not clear enough. For the contourlet method, shearlet method, CSR method, and SR-SML method, though the results are improved, the contours are sharpened and the phenomenon of ghosting occurs. This can be explained by comparing the ability of the sparse representations for the important image features.

As for the GAN based the DCNN-based method, the results are much clear, but the texture information is not good enough by comparing the results obtained by the proposed method. This is because these two models are directly learned based on the pixel of the images to be fused. The importance of the feature in the procedure of learning is not fully considered. On the other hand, the texture information in the proposed method is highly improved and the ghosting phenomenon is suppressed to the greatest extent. Furthermore, this can also be proved by the enlarged images in Figures 7 and 11. In addition, from the objective comparison in Tables 1 and 2, the best value of the four



FIGURE 5: Parts of the source images in experiments.

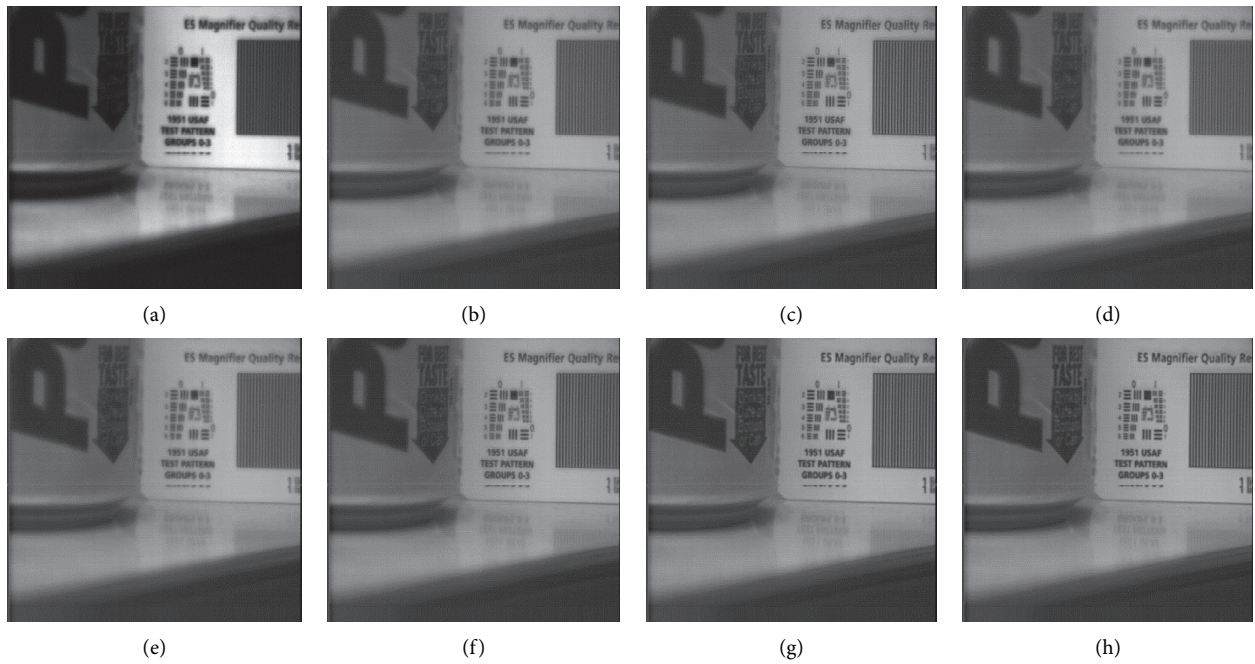


FIGURE 6: The fusion results of “Pepsi-Cola.” (a) PCNN. (b) Contourlet. (c) GAN. (d) Shearlet. (e) CSR. (f) SR-SML. (g) DCNN. (h) Proposed.

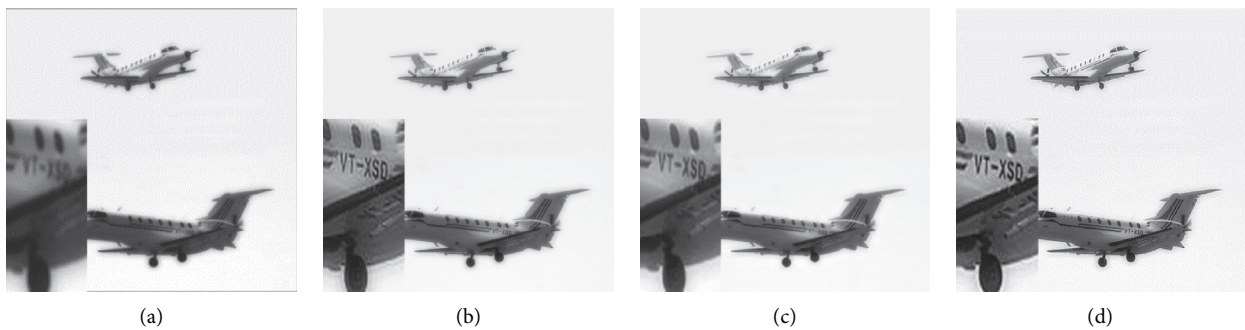


FIGURE 7: Continued.

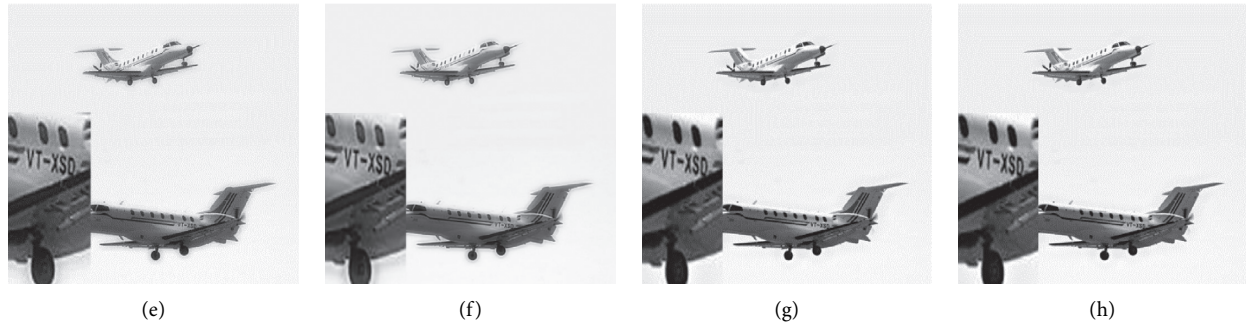


FIGURE 7: The fusion results of “Plane.” (a) PCNN. (b) Contourlet. (c) GAN. (d) Shearlet. (e) CSR. (f) SR-SML. (g) DCNN. (h) Proposed.

TABLE 1: The fusion performance of the seven methods and the proposed method from Figures 6–8.

Method	Pepsi-Cola				Clock				Plane			
	SD	$Q^{AB/F}$	En	MI	SD	$Q^{AB/F}$	En	MI	SD	$Q^{AB/F}$	En	MI
PCNN	43.71	0.36	6.58	4.46	39.27	0.50	6.98	5.01	44.97	0.40	3.91	3.39
Contourlet	44.11	0.59	7.00	5.39	39.41	0.57	7.00	5.19	46.24	0.48	3.98	3.35
Shearlet	44.07	0.61	7.10	5.34	39.44	0.52	7.00	5.21	46.70	0.49	4.06	3.41
GAN	44.25	0.70	7.16	5.38	39.95	0.62	7.06	5.20	45.842	0.69	4.04	3.72
CSR	45.23	0.76	7.10	5.50	40.50	0.68	7.03	5.42	48.10	0.73	4.08	3.60
SR-SML	45.40	0.78	7.11	5.55	40.88	0.69	7.05	5.44	48.90	0.76	4.19	3.62
DCNN	44.80	0.74	7.06	5.41	39.66	0.65	7.00	5.32	46.85	0.68	4.03	3.58
Proposed	45.25	0.78	7.20	5.60	40.88	0.69	7.10	5.53	50.15	0.76	4.28	3.72

TABLE 2: The fusion performance of the seven methods and the proposed method from Figures 9–11.

Method	Flower				Cup				Calendar			
	SD	$Q^{AB/F}$	En	MI	SD	$Q^{AB/F}$	En	MI	SD	$Q^{AB/F}$	En	MI
PCNN	40.61	0.41	3.90	4.72	38.22	0.43	4.01	4.21	37.42	0.40	4.88	4.31
Contourlet	40.62	0.40	4.02	4.73	38.39	0.42	4.03	4.38	37.65	0.52	5.35	4.47
Shearlet	40.81	0.42	4.06	4.89	38.44	0.45	4.08	4.41	38.15	0.56	5.55	4.55
GAN	40.81	0.57	4.80	4.99	39.10	0.51	4.10	4.63	38.59	0.60	5.76	4.91
CSR	40.89	0.59	4.65	4.77	38.55	0.54	4.10	4.60	39.03	0.66	5.90	4.89
SR-SML	41.19	0.57	4.66	5.12	38.71	0.55	4.14	4.72	39.11	0.67	5.93	4.91
DCNN	40.16	0.60	4.81	5.16	40.01	0.58	4.09	4.81	39.98	0.65	5.98	5.24
Proposed	41.26	0.68	4.96	5.27	40.21	0.63	4.18	4.93	39.91	0.67	6.00	5.29

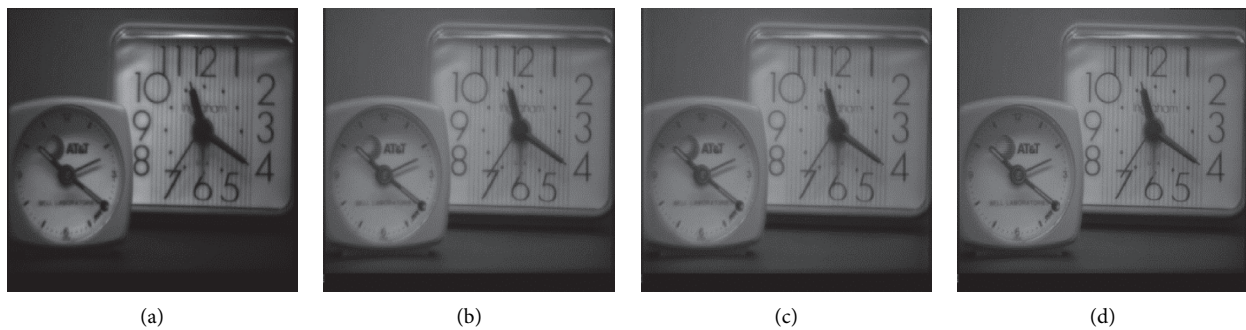


FIGURE 8: Continued.

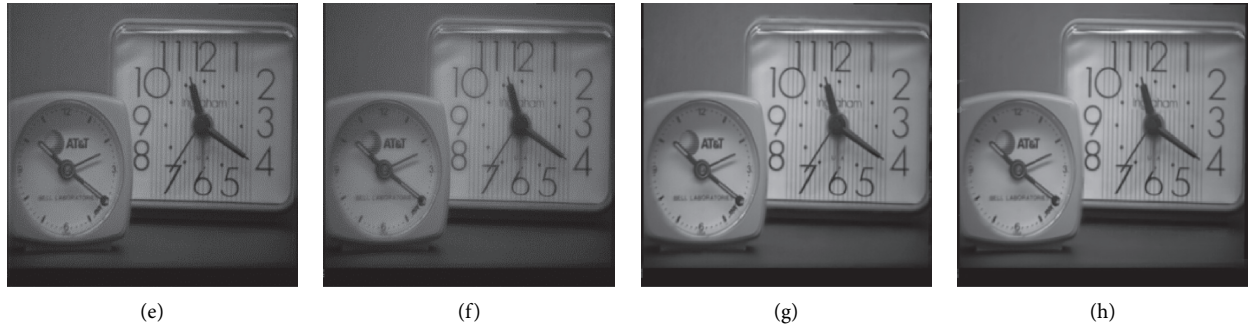


FIGURE 8: The fusion results of “Clocks.” (a) PCNN. (b) Contourlet. (c) GAN. (d) Shearlet. (e) CSR. (f) SR-SML. (g) DCNN. (h) Proposed.

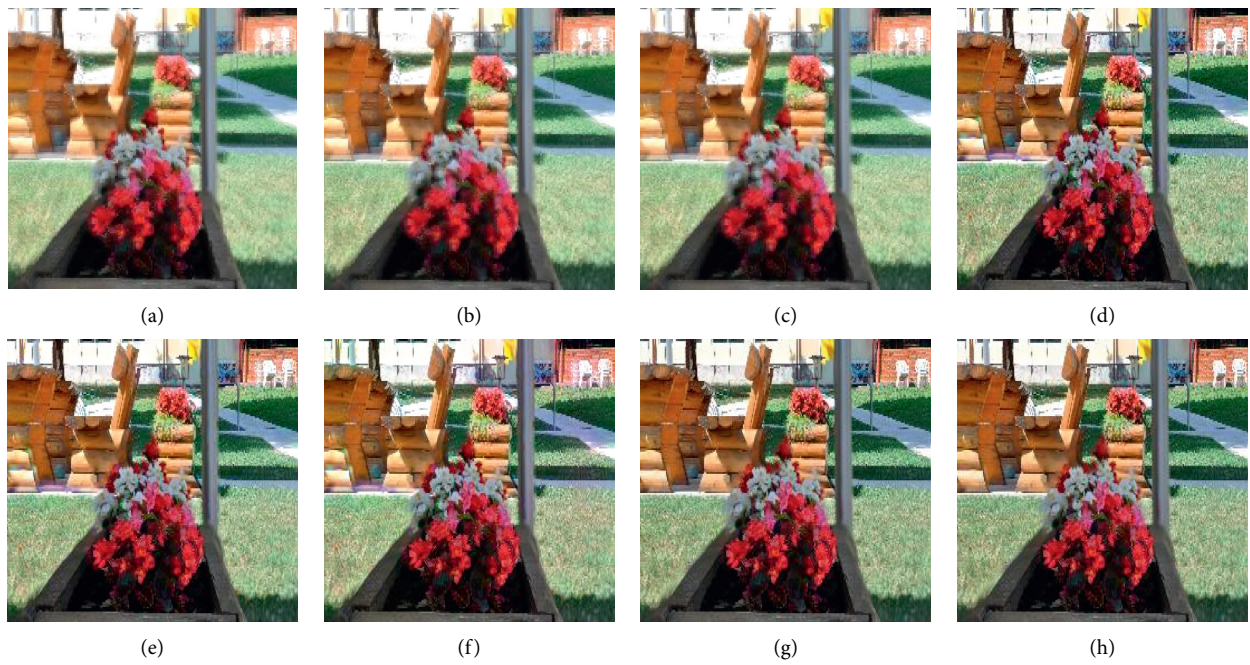


FIGURE 9: The fusion results of “Flower.” (a) PCNN. (b) Contourlet. (c) GAN. (d) Shearlet. (e) CSR. (f) SR-SML. (g) DCNN. (h) Proposed.

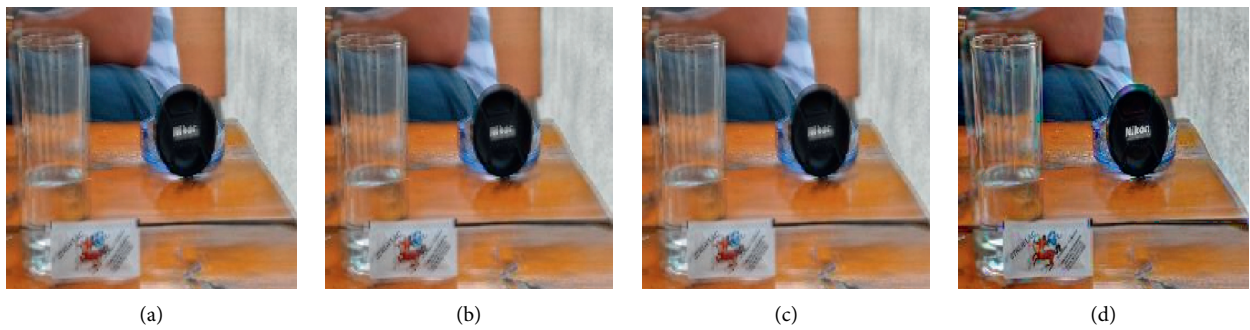


FIGURE 10: Continued.

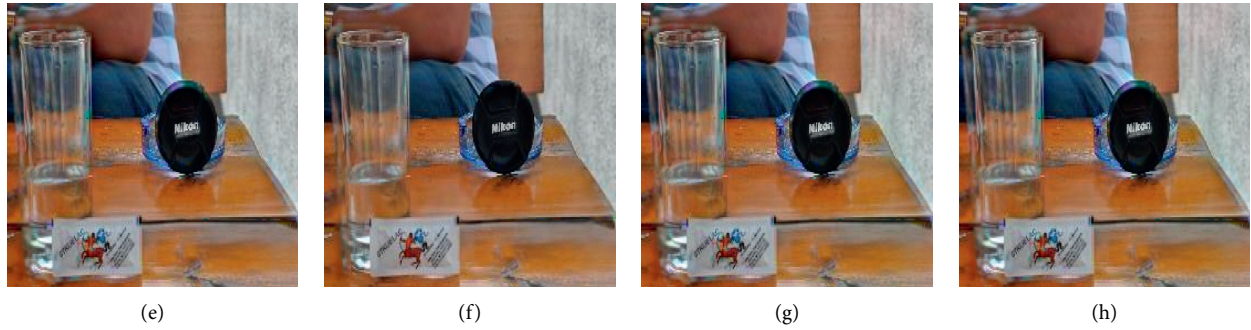


FIGURE 10: The fusion results of “Cup.” (a) PCNN. (b) Contourlet. (c) GAN. (d) Shearlet. (e) CSR. (f) SR-SML. (g) DCNN. (h) Proposed.

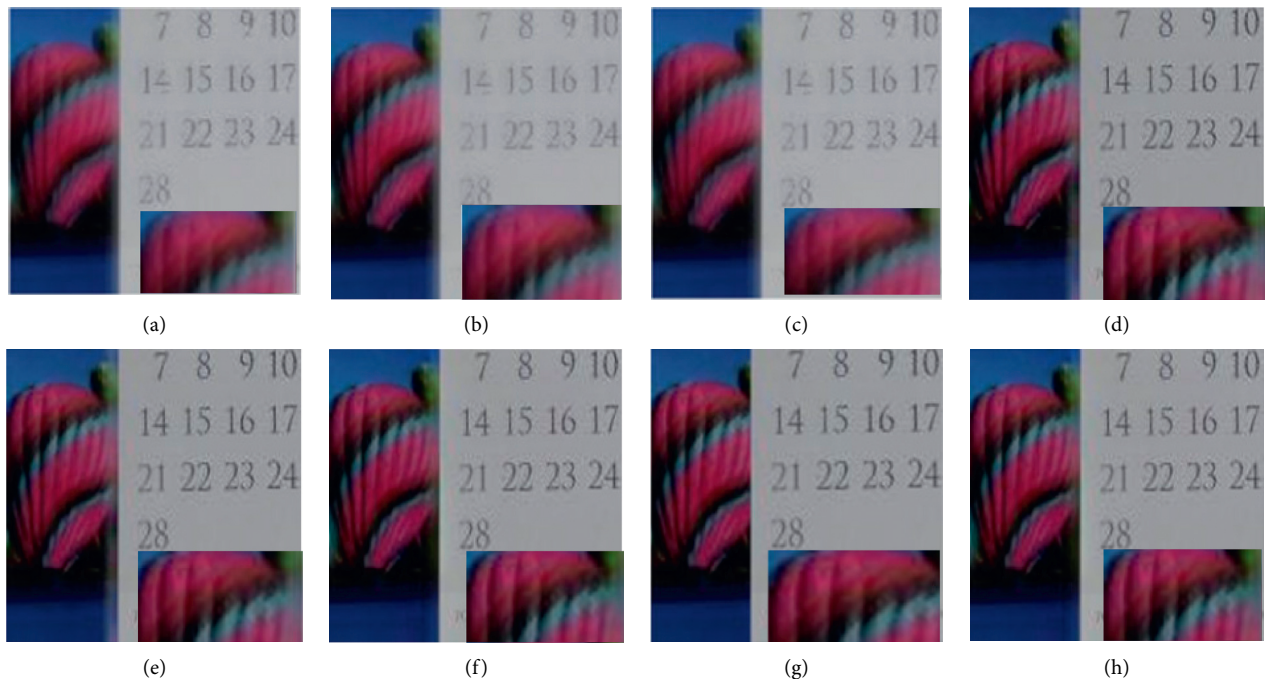


FIGURE 11: The fusion results of “Calendar.” (a) PCNN. (b) Contourlet. (c) GAN. (d) Shearlet. (e) CSR. (f) SR-SML. (g) DCNN. (h) Proposed.

metrics can be almost obtained by the proposed method. All the above facts fully demonstrate the effectiveness and accuracy of the proposed method.

4. Conclusion

To get better fusion results for the multi-focus images, the features-motivated generative adversarial network is constructed with the help of the complex shearlet transform. Six typical experiments have been carefully implemented to show the full evidence of the effectiveness and accuracy. In the future, more complex models will be built to further improve the fusion performance.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (61502282 and 61902222), the Natural Science Foundation of Shandong Province (ZR2015FQ005), and the Taishan Scholars Program of Shandong Province (tsqn201909109).

References

- [1] S. C. Kulkarni and P. P. Rege, “Pixel level fusion techniques for SAR and optical images: a review,” *Information Fusion*, vol. 59, pp. 13–29, 2020.
- [2] Y. Liu, X. Chen, Z. Wang, Z. J. Wang, R. K. Ward, and X. Wang, “Deep learning for pixel-level image fusion: recent advances and future prospects,” *Information Fusion*, vol. 42, pp. 158–173, 2018.

- [3] J. Tian, G. Liu, and J. Liu, "Multi-focus image fusion based on edges and focused region extraction," *Optik*, vol. 171, pp. 611–624, 2018.
- [4] J. Dou, Q. Qin, and Z. Tu, "Image fusion based on wavelet transform with genetic algorithms and human visual system," *Multimedia Tools and Applications*, vol. 78, no. 9, pp. 12491–12517, 2019.
- [5] S. Li and B. Yang, "Multifocus image fusion by combining curvelet and wavelet transform," *Pattern Recognition Letters*, vol. 29, no. 9, pp. 1295–1301, 2008.
- [6] K. He, D. Zhou, X. Zhang, and R. Nie, "Multi-focus: focused region finding and multi-scale transform for image fusion," *Neurocomputing*, vol. 320, pp. 157–170, 2018.
- [7] X.-B. Qu, J.-W. Yan, H.-Z. Xiao, and Z.-Q. Zhu, "Image fusion algorithm based on spatia frequency-motivated pulse coupled neural networks in non-subsampled contourlet transform domain," *Acta Automatica Sinica*, vol. 34, no. 12, pp. 1508–1514, 2009.
- [8] Y. Liao, W. Huang, L. Shang et al., "Image fusion based on shearlet and improved PCNN," *Computer Engineering and Application*, vol. 50, pp. 142–146, 2014.
- [9] V. P. S. Naidu, "Hybrid DDCT-PCA based multi sensor image fusion," *Journal of Optics*, vol. 43, no. 1, pp. 48–61, 2014.
- [10] D. Carone, G. W. J. Harston, J. Garrard et al., "ICA-based denoising for ASL perfusion imaging," *NeuroImage*, vol. 200, pp. 363–372, 2019.
- [11] T. Li and Y. Wang, "Biological image fusion using a NSCT based variable weight method," *Information Fusion*, vol. 85, 2011.
- [12] Y. Yang, Z. Nie, S. Huang, P. Lin, and J. Wu, "Multilevel features convolutional neural network for multifocus image fusion," *IEEE Transactions on Computational Imaging*, vol. 5, no. 2, pp. 262–273, 2019.
- [13] C. Du, S. Gao, Y. Liu, and B. Gao, "Multi-focus image fusion using deep support value convolutional neural network," *Optik*, vol. 176, pp. 567–578, 2019.
- [14] Y. Zhang, Y. Liu, P. Sun, H. Yan, X. Zhao, and L. Zhang, "IFCNN: a general image fusion framework based on convolutional neural network," *Information Fusion*, vol. 54, pp. 99–118, 2020.
- [15] H. Zhang, Z. Le, Z. Shao, H. Xu, and J. Ma, "MFF-GAN: an unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion," *Information Fusion*, vol. 66, pp. 40–53, 2021.
- [16] J. Ma, W. Yu, C. Chen, P. Liang, X. Guo, and J. Jiang, "Pan-GAN: an unsupervised pan-sharpening method for remote sensing image fusion," *Information Fusion*, vol. 62, pp. 110–120, 2020.
- [17] J. Ma, W. Yu, P. Liang, C. Li, and J. Jiang, "FusionGAN: a generative adversarial network for infrared and visible image fusion," *Information Fusion*, vol. 48, pp. 11–26, 2019.
- [18] B. Lu, Y. Hu, L. Lin et al., "Using ensemble deep learning method to integrate multi-source data to develop national visibility grid data," *Advances in Meteorological Science and Technology*, vol. 8, pp. 77–82, 2018.
- [19] X. Liu, Ke Xu, P. Zhou, and J. Chi, "Edge detection of retinal OCT image based on complex shearlet transform," *IET Image Processing*, vol. 13, pp. 1686–1693, 2019.
- [20] H. Karbalaali, A. Javaherian, S. Dahlke, and S. Torabi, "Channel edge detection using 2D complex shearlet transform: a case study from the South Caspian Sea," *Exploration Geophysics*, vol. 49, no. 5, pp. 704–712, 2018.
- [21] X. Jin, D. Zhou, S. Yao et al., "Multi-focus image fusion method using S-PCNN optimized by particle swarm optimization," *Soft Computing*, vol. 22, no. 19, pp. 6395–6407, 2018.
- [22] N. Hayat and M. Imran, "Ghost-free multi exposure image fusion technique using dense SIFT descriptor and guided filter," *Journal of Visual Communication and Image Representation*, vol. 62, pp. 295–308, 2019.
- [23] X. Liu, Y. Wang, and Q. Liu, "PSGAN: a generative adversarial network for remote sensing image pan-sharpening," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, pp. 873–877, Athens, Greece, October 2018.
- [24] X. Guo, R. Nie, J. Cao D, Z. L. Mei, and K. He, "FuseGAN: learning to fuse multi-focus image via conditional generative adversarial network," *IEEE Trans. on Multimedia*, vol. 21, pp. 1982–1996, 2019.
- [25] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2261–2269, Honolulu, HI, USA, July 2017.
- [26] C.-B. Du and S.-S. Gao, "Multi-focus image fusion with the all convolutional neural network," *Optoelectronics Letters*, vol. 14, no. 1, pp. 71–75, 2018.
- [27] S. Slavica, "Multi-focus image fusion based on empirical mode decomposition," in *Twentieth International Electro technical and Computer Science Conference, ERK 2011*, San Francisco, CA, USA, October 2011.
- [28] M. Nejati, S. Samavi, and S. Shirani, "Multi-focus image fusion using dictionary-based sparse representation," *Information Fusion*, vol. 25, pp. 72–84, 2015.
- [29] G. Piella, "A general framework for multiresolution image fusion: from pixels to regions," *Information Fusion*, vol. 4, no. 4, pp. 259–280, 2003.
- [30] "Multi-focus-Image-Fusion-Dataset," <https://github.com/sametaimaz/Multi-focus-Image-Fusion-Dataset>.
- [31] Z. Wang, Y. Ma, and J. Gu, "Multi-focus image fusion using PCNN," *Pattern Recognition*, vol. 43, no. 6, pp. 2003–2016, 2010.
- [32] Z. Zhu, H. Yin, Y. Chai, Y. Li, and G. Qi, "A novel multi-modality image fusion method based on image decomposition and sparse representation," *Information Sciences*, vol. 432, pp. 516–529, 2018.
- [33] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Information Fusion*, vol. 36, pp. 191–207, 2017.
- [34] X. Liu, Y. Zhou, and J. Wang, "Image fusion based on shearlet transform and regional features," *AEU - International Journal of Electronics and Communications*, vol. 68, no. 6, pp. 471–477, 2014.
- [35] Y. Liu, X. Chen, R. K. Ward, and Z. Jane Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Processing Letters*, vol. 23, no. 12, pp. 1882–1886, 2016.
- [36] X. Li, X. Zhang, and M. Ding, "A sum-modified-Laplacian and sparse representation based multimodal medical image fusion in Laplacian pyramid domain," *Medical & Biological Engineering & Computing*, vol. 57, no. 10, pp. 2265–2275, 2019.
- [37] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," *Computers & Electrical Engineering*, vol. 37, no. 5, pp. 744–756, 2011.
- [38] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, 2006.
- [39] V. Aslantans and E. Bendes, "A new image quality metric for image fusion: the sum of the correlations of differences," *AEU - International Journal of Electronics and Communications*, vol. 69, no. 12, pp. 1890–1896, 2015.