IJLAI Transactions on Science and Engineering (ISSN: 3006-3418) 2024, vol. 2, no. 4, pp: 1–8

Research on Image Fusion Based on Deep Adversarial Learning

Yachao Zhang[1], Lijuan Feng[1], and Jiaxiang Wang[1]

School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology
Zhengzhou 450064, China
857003841@qq.com
Corresponding author:Lijuan Feng
Received Oct. 5, 2024; Revised and Accepted Oct. 22, 2024

Abstract. Image fusion, as an important task in computer vision, essentially extracts important features from source images to complement each other and generate fusion images with higher quality and richer information. Infrared and visible images contain different information due to different imaging quantity principles. The key of infrared and visible image fusion algorithm is to integrate the thermal radiation information extracted from infrared images with the captured details and texture information of visible images, so as to obtain a fusion image with complete structure and rich detailed information. Based on the generative adversarial network model, this paper proposes an infrared and visible image fusion method based on dual path dual discriminator generating adversarial network, aiming at the problems existing in the existing research algorithms, such as inadequate extraction of feature information, low efficiency of network model feature transfer, easy loss of shallow information in single-path feature extraction, fewer fusion levels caused by sub-path feature extraction and unbalance of discriminator modes. The gradient path and contrast path based on the difference stitching of source images are constructed at the generator side to improve the detail information and contrast of fused images. The feature information of infrared and visible images is extracted by multi-scale decomposition to solve the problem of incomplete feature extraction on a single scale. Then the source image is introduced into each layer of the double-path dense network, which can improve the efficiency of feature transmission and obtain more source image information. At the end of the discriminator, a double discriminator is used to estimate the region distribution of infrared image and visible image, so as to avoid the mode imbalance problem of the loss of infrared image contrast information in the single discriminator network. Finally, we construct the master-auxiliary gradient and the master-auxiliary strength loss function to improve the information extraction ability of the network model. Compared with other image fusion methods on public data sets, the experimental results show that the proposed method achieves good results on objective evaluation indexes (mean gradient, spatial frequency, structural similarity and peak signal-to-noise ratio).

Keywords: Image fusion, Dual path dual discriminator, Adversarial network, Master-auxiliary gradient.

1. Introduction

Image fusion refers to the extraction of information from multiple images of the same scene with complementary content, the complementary synthesis of image information obtained from different sensors, and the formation of a higher-quality image through integration processing to meet people's demand for more comprehensive information [1]. So image fusion is a very active research direction in recent years. Image fusion technology is often used in remote sensing images, medical images, infrared and visible images with different imaging modes. In the field of remote sensing detection, by combining the information of multi-source remote sensing images, the spatial resolution and spectral resolution of images are improved, and the interpretation and information extraction capabilities of remote sensing images are enhanced, so as to identify ground targets more accurately. In the medical field, image fusion technology is often used to fuse images of different modes, such as CT and MRI, so that doctors can have a more comprehensive understanding of the patient's condition [1-3]. In the field of infrared and visible light, image fusion is widely used in military surveillance, target detection, night navigation and so on. Image fusion technology can be used to improve the visual perception of robots, enabling them to identify and locate targets more accurately. In addition, image fusion technology can also be applied to computer vision, automatic driving, environmental monitoring, disaster detection and prediction and other fields [4,5]. With the continuous development of science and technology, the application scenarios of image fusion technology will continue to expand and deepen.

Infrared and visible image fusion technology is an important research direction of image fusion. Infrared images are obtained by scene thermal radiation imaging, which is less affected by weather, sensitive to the temperature distribution of objects, and significant to thermal targets. However, due to the limitations of sensors, background is blurred and texture is missing. Visible light image mainly reflects the reflected light information of

the object, sensitive to the surface characteristics of the object, rich texture details, edge information is clearer, but will be affected by atmospheric and illumination factors such as image quality. Infrared and visible image fusion is to complement the main thermal target information in infrared image with the texture edge information of visible image, improve the clarity and information of image, and generate fusion images with significant thermal target and rich texture details [6]. Infrared and visible light image fusion technology has risen rapidly in the civilian field, such as in airports, ports and other areas to check passengers carry items, infrared images can directly reflect the body characteristics of pedestrians or their hidden locations, while visible light images can clearly reflect the surrounding background environment. By merging these two images, more detailed background information can be provided to reflect the location of the target and provide strong support for subsequent target tracking, identification and behavior analysis. In the field of autonomous driving, infrared and visible image fusion technology can be used to improve the perception of vehicles in various weather and lighting conditions. Especially at night or in bad weather, infrared images can provide critical information about pedestrians, vehicles, and other obstacles, while visible images provide detailed information about road signs, traffic signals, and more. Through the fusion technology of infrared and visible light, the autonomous driving system can more accurately perceive and understand the surrounding environment, thereby improving the safety and reliability of driving. In the field of fire rescue, infrared and visible image fusion technology can help rescue workers quickly and accurately find the fire source and trapped people. Infrared images can show the location and temperature distribution of the fire source, while visible images provide the specific layout and details of the fire site, and rescue workers can have a more comprehensive understanding of the fire situation, so as to develop a more effective rescue plan [7,8].

At present, with the continuous breakthrough of technology, many tasks have been introduced deep learning, by enriching data sets, neural networks can well solve many problems, while having a certain degree of robustness, but in order to enhance the detailed expression of generated images, The infrared and visible image fusion task is introduced into the Generative Adversarial Network (GAN) [9]. Compared with ordinary neural networks, GAN provides a new training mode, and the parameter updating of the generator depends on the discriminator in a certain sense. The large amount of existing data shows that GAN can generate clearer and more authentic samples. However, the research of infrared and visible image fusion based on GAN still faces many challenges.

1. Most GAN-based image fusion methods use only one discriminator. In the discriminant process, although more contour texture details of visible images can be fitted, less information will be saved in infrared images.
2. In many current end-to-end GAN image fusion methods, most of the generators are simple network structures, resulting in insufficient and unbalanced feature extraction of infrared and visible image, and the problem of ignoring small details.
3. In the common generator network structure, if the infrared image and visible image branches are used for feature extraction, there are fewer levels for fusion features, and it is impossible to obtain better fusion results. If the infrared image and visible image are combined as input, the network model will lose the shallow feature information of the source image.

To solve the problem of partial loss of infrared image information caused by the use of single discriminator, this paper designs a double discriminator to enhance the discriminant ability of infrared image, visible image and fusion image.

2. Infrared and Visible Image Fusion with Dual Path Dual Discriminator

In the fusion stage of traditional infrared and visible image fusion methods, fusion rules need to be designed manually according to specific problems, resulting in poor generalization performance. With the rapid development of deep learning technology, researchers have proposed various image fusion methods based on generative adversarial network to avoid the design of complex fusion rules by using the end-to-end characteristics of generative adversarial network. The fusion result of FusionGAN algorithm can retain the texture information and contrast information well, but because only the deep semantic features of the previous layer are fused in the last layer of the FusionGAN model, the useful information obtained in the middle layer and the shallow layer is lost, and the problems of detail loss and target edge blur appear. Detail-GAN designs the target edge loss function, deepens the layers of the generator and discriminator, and optimizes the target texture, thus making the target appear more clearly in the fusion result, and better solves the edge blur problem of FusionGAN algorithm [10]. However, using only one discriminator for identification still leads to the fusion image approaching a single source image. GANMcC method transforms image fusion into multi-distribution simultaneous estimation problem, which makes the fusion result more balanced with significant contrast and rich texture details. However, the generator network structure in GANMcC algorithm is relatively simple, the source image feature information is not sufficiently extracted, and the feature transmission efficiency in the network structure is low.

Aiming at the problems in image fusion algorithm, such as insufficient source image information preservation, insufficient detail information and unbalance of single authentication modes, this paper proposes a fusion method of Infraredimage (IR) and visible image (VI) generated by dual-path dual-discriminator against the network. By splicing the source image as input, the tiny feature information of the source image is extracted more fully. In the generator, multi-scale decomposition of the source images in the double path is carried out to extract different scale features, dense network blocks are introduced to improve the feature reuse rate, avoid overfitting, and source images are added to each layer of the dense network to retain more source image information. In the discriminator, the dual discriminator and generator are trained against each other to prevent modal imbalance. The main and auxiliary gradient loss and the main and auxiliary intensity loss functions are constructed, and the texture details easily ignored in IR and the contrast information in VI are extracted.

2.1. Structure of Proposed Method

The process of this chapter is shown in Figure 1. Texture information mainly exists in VI, contrast information mainly exists in IR, VI also has contrast information, IR also contains texture information [45]. Inspired by this, contrast and texture information can be extracted unevenly by means of differential matching input, which can fully extract these usually neglected feature information. The input in this paper is divided into gradient path and contrast path. The gradient path is output by combining two VI images and one IR image, which can effectively extract the texture information and a little contrast information in VI. Similarly, contrast path can effectively extract contrast information and a small amount of texture information in IR by combining 2 IR images and 1 VI images. The structure of generator network in most algorithms is a relatively simple convolution layer, and the feature information is not fully extracted and utilized, resulting in the loss of fusion image information. Firstly, multi-scale feature decomposition is applied to the input of two paths to increase the receptive field and obtain the feature information of different scales, so as to avoid the problem of insufficient feature information extraction by using convolution kernel of the same size. Dense blocks are then used so that each layer is connected to all previous layers in the channel dimension as input to the next layer, extracting more valid feature information. IR and VI are introduced into the middle layer of dense blocks respectively, and the operation of source images is added to each layer of the network, which is equivalent to using different depth networks for feature extraction of source images, and more feature information of source images can be retained. In addition, since the single discriminator is only used to distinguish VI from fused images, it is easy to cause modal imbalance in the training process, resulting in fusion results unable to maintain the contrast of IR or texture details in VI. Therefore, this paper classifiers as discriminators to estimate the distribution of two different domains of visible light and infrared light at the same time. The consistency of the probability distribution makes the fusion result have the most significant features in the target distribution, and the generator is more likely to capture the key features and enhance them. Through continuous training between the generator and two discriminators, the final generator can produce a fusion image with significant contrast and rich texture details.

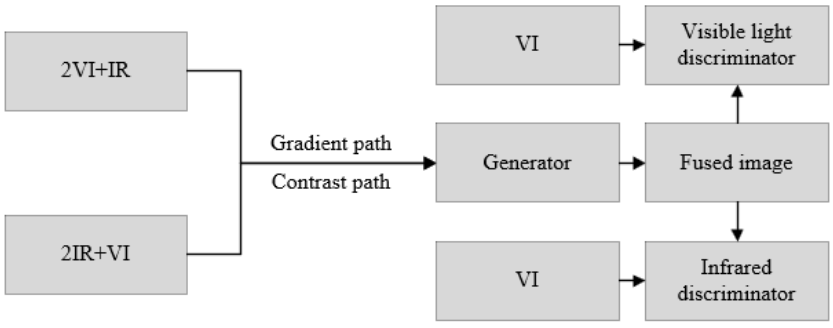

Fig. 1. Proposed method

2.2. Generator Structure

The input of generator structure in this paper is divided into gradient path and contrast path. The gradient path is used to extract high-frequency texture feature information, and two VI images and one IR image are used as input. The contrast path is used to extract significant contrast information in IR [11,12], using two IR and one VI

difference splicing as input. Among the two information extraction paths, the multi-scale decomposition module is firstly adopted, which can solve the problem that only a convolution kernel of a single scale is used to extract features, resulting in the image feature performance at other scales cannot be perceived. The MDB structure is shown in Figure 2. MDB is convolved with 5×5, 3×3 and 1×1 scale convolution kernels, respectively, and activated by batch normalization (BN) and leakyrectified linear unit (LRelu) functions. After that, dense blocks are used for feature extraction, and convolution kernels of 5×5, 3×3 and 3×3 are used for dense blocks respectively. The source image is added to each layer of the dense block for feature extraction. After the convolution operation, BN and Lrelu activation operations are carried out. After the feature map is connected, the BN and LRelu activation operations are carried out by 3×3 convolution kernel. The last layer uses a 1×1 convolution kernel to reduce the dimension of the connected features to a single channel image through Tahn activation function to achieve feature fusion. All steps in the generator structure are set to 1.

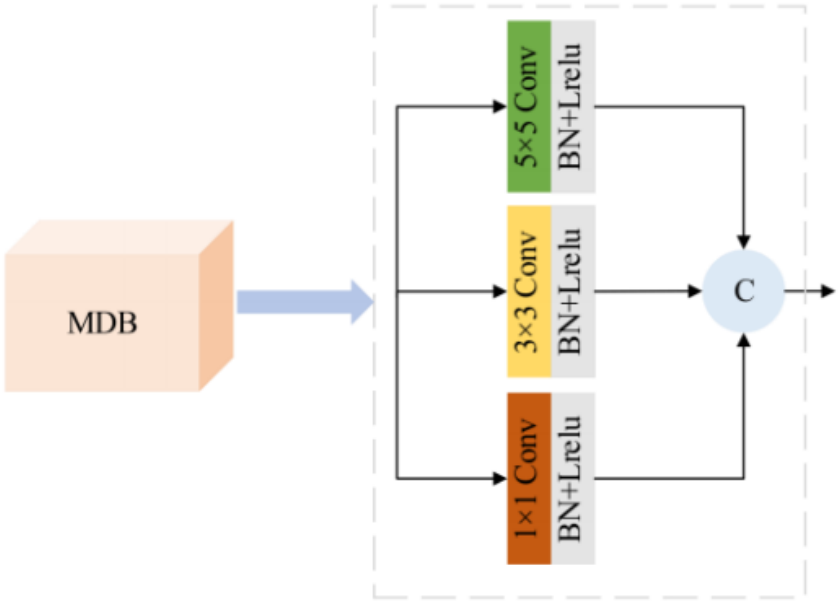

Fig. 2. MDB structure

2.3. Discriminator Structure

The proposed method has two discriminators: infrared light discriminator DIR and visible light discriminator DVI. Among them, DIR is used to distinguish the incoming image from IR or fusion image, and DVI is used to distinguish the incoming image from VI or fusion image, both of which have the same structure, as shown in Figure 3. Input an image, and the discriminator network eventually generates a probability scalar that estimates the probability that the input image is IR or VI, respectively, rather than the image derived from the generator. The discriminator is a 5-layer CNN with step size set to 2 and padding set to VALID. From layer 1 to layer 4, the size of the convolution kernel is set to 3×3. After convolution, BN is processed, and LRelu function is used as the activation function [13-17]. The final layer is the fully connected layer for classification.

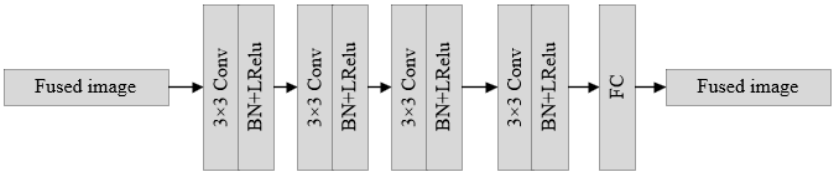

Fig. 3. Discriminator construct

2.4. Loss Function

The loss functions of the model in this paper are generator loss function L_G, IR discriminator loss function Lp and VI discriminator loss function $L_{D_{IR}}$. The IR discriminator DIR and VI discriminator DVI work against the generator G. The two discriminators treat I_{IR} and I_{VI} as real data respectively, and the fused image I_{fused} as false data for D_{IR} and D_{VI}.

The generator loss function L_G consists of two parts: content loss L_{con} and counterloss L_{adv}. L_{con} is used for constraint information extraction and L_{adv} is used for constraint information balance. The formula is as follows:

$$L_G = \gamma L_{con} + L_{adv}. \tag{1}$$

Where γ represents the regularization parameter, which is set to 100 in this paper in order to maintain a balance between content loss and counterloss.

Gradient information is the main information obtained from VI, and the main gradient loss is defined as:

$$L_{Grad-main} = ||\nabla I_{fused} - \nabla I_{VI}||_F^2. \tag{2}$$

Where ∇ represents the second-order gradient operator. $||\cdot||_F$ stands for F-norm. IR has a significant contrast that highlights the target information in the background. As the primary information obtained from IR, the primary strength loss is defined as:

$$L_{intensity-main} = ||I_{fused} - I_{IR}||_F^2. \tag{3}$$

VI also contains some contrast information, and IR also contains some texture details. The auxiliary gradient loss $L_{Grad-aux}$ is constructed between VI and the fused image, and the auxiliary intensity loss $L_{Intensity-aux}$ is constructed between IR and the fused image, defined as:

$$L_{Grad-aux} = ||\nabla I_{fused} - \nabla I_{IR}||_F^2. \tag{4}$$

$$L_{Intensity-aux} = ||I_{fused} - I_{VI}||_F^2. \tag{5}$$

3. Experimental Results and Discussion

3.1. Experiment Settings

The algorithm selects 10 pairs of images from the TNO and RoadScene data sets respectively as test sets. Eight representative methods are selected for comparison, including discraete wavelet transform (DWT) [18], dual-tree complex wavelet transforms (DT-CWT) [19], nonsubsampled contourlet transform (NSCT) [20], Densefuse [21], FusionGAN [22], GANMcC [23], adual-discriminator conditional generative adversarial network for multi-resolution image fusion (DDcGAN) [24], semantic-supervised infrared and visible image fusion via a dual-discriminator generative network adversarial network (SDDGAN) [25]. Ablation experiments are set up, and the original method in this method is compared with each module to prove the effectiveness of the proposed module. All experiments are compared by subjective evaluation and objective evaluation indicators. The experimental environment is Windows10, Python3.6 and Tensorflow1.15. The hardware configuration environment is Intel i7-12700F CPU and NVIDIA GeForce RTX2080 GPU. All experiments are carried out in the same environment.

In order to better analyze the quality of fused images, six objective evaluation indexes, AG, SF, EN, MI, SSIM and PSNR, are selected to analyze the quality of fused images, and they are all positive indicators. AG, SF and EN indexes are calculated by a single fusion image. The index values of MI, SSIM and PSNR are obtained by calculating the values of the fused image and IR, and the fused image and VI. On the TNO data set, the average values of 9 methods and 10 groups of experimental data are shown in Table 1.

It can be seen that among the 10 groups of experimental comparison, most of the objective evaluation indexes of the method in this chapter are better than other comparison methods, among which AG and SF indexes are significantly higher than other methods, indicating that the new method has a clear fusion image texture edge and a strong ability to express small details. The SSIM and PSNR indexes of the new method are optimal, indicating that the fusion result has less distortion and distortion. EN is an indicator to measure the amount of information contained in an image. The larger the EN value, the richer the information in the fused image, but noise will also affect the EN result. It can be seen that DDcGAN has the highest EN value, but it can be seen from the subjective evaluation that the noise generated by this method makes the EN value high. The EN value of the new method is second only to DDcGAN, and there is no obvious noise in the fusion result, indicating that the fusion result of the

Table 1. Comparison of average values of 10 sets of image data on the TNO dataset

Method	AG SF EN MI SSIM PSNR/dB					
DWT	2.8102	5.6864	6.2569	1.7046	0.6126	30.4417
DT-CWT	5.3159	10.9944	6.4396	1.5167	0.8187	31.1602
NSCT	5.5903	11.3393	6.5099	1.5602	0.8754	30.7731
DenseFuse	4.6001	8.8816	6.6893	1.9884	0.8398	31.0345
FusionGAN	2.9768	5.8702	6.3208	2.0220	0.5975	30.3832
GANMcC	3.0603	6.0047	6.5074	2.0402	0.7521	30.5971
DDcGAN	6.8107	12.8200	7.4878	1.5853	0.6564	27.5990
SDDGAN	4.5744	9.1192	7.0434	1.8725	0.8051	29.9164
Proposed	7.2042	13.7953	7.1463	2.0107	0.8847	31.3369

new method contains relatively rich information. The higher the MI index value, the more information contained in each other. It can be seen that compared with most comparison methods, the MI index of the new method has achieved a certain degree of improvement.

The same 6 objective evaluation indicators as those in TNO dataset were selected to conduct generalization experiments on fusion image quality. The average pairs of 10 groups of experimental data in RoadScene dataset were shown in Table 2. It can be seen that the proposed method still achieves the best performance on AG, SF, SSIM and PSNR indexes. The EN value is high because of the noise in DDcGAN image fusion. The experimental results show that the method in this paper achieves the best results in 4 of the 6 indexes, 1 index is the second and 1 index is the third, and the experimental results are excellent. It is proved that the fusion image generated by this method can maintain a good balance of IR and VI information.

Table 2. Comparison of average values of 10 sets of image data on the RoadScene dataset

Method	AG	SF	EN	MI	SSIM	PSNR/dB
DWT	5.2543	11.2149	6.9395	2.6198	0.8013	27.0462
DT-CWT	6.0016	12.9394	6.9455	2.3051	0.7328	26.8691
NSCT	6.6081	14.2062	7.1009	2.4384	0.8511	27.2642
DenseFuse	4.8299	10.4157	7.0365	2.8997	0.8062	27.1210
FusionGAN	3.3602	7.5575	7.0065	2.7752	0.5110	25.4474
GANMcC	3.7967	8.0395	7.1183	2.8589	0.7360	27.1600
DDcGAN	4.9417	11.0247	7.6240	2.4093	0.5008	25.3809
SDDGAN	4.5516	9.8219	7.6196	3.1749	0.7193	27.1744
Proposed	6.7180	14.8152	7.5761	2.9083	0.8544	27.3352

The objective evaluation data of ablation experiments conducted on two data sets in this chapter are shown in Table 3 and Table 4. It can be seen that: (1) On No-Both module, except for PNSR index in MSRS data set, other indicators have the lowest value in each data set, indicating the effectiveness of each module of the proposed method. (2) NO_DC, No_MDB Because the feature information of the source image is not sufficiently extracted by the No_Dense module, the information content and clarity of the fusion results are lower than that of the proposed method. In the TNO and RoadScene data sets, all indexes are lower than that of the proposed method. In MSRS data set, except PSNR, all the indexes of these three modules are lower than that of the method in this paper. (3) The ablation results are consistent with the MSRS dataset test. In the fusion of low-light source images, when the fusion image with small change in gray value and low brightness is generated, the PSNR index value is high, but the visual effect is not good. (4) One-Disc module, which can extract more source image feature information, has lower indicators than the method in this chapter, indicating that the single discriminator will affect the fusion result and reduce the retention of source image feature information.

In summary, the objective evaluation results are consistent with the subjective evaluation results, which proves the validity of each module proposed in this chapter. The method proposed in this paper can retain more information of source images, and extract thermal target information more prominently. The fusion result contains more useful information and has better visual effect.

Table 3. Comparison of ablation experiment results on TNO dataset

Module	AG	SF	EN	MI	SSIM	PSNR/dB
No_DC	4.6428	8.9061	6.6623	1.8637	0.8401	30.7428
No_MDB	6.3835	12.1849	7.0595	1.8381	0.8516	31.1711
No_Dense	6.2091	12.0344	7.0416	1.7810	0.8748	30.9938
One_Disc	4.5723	9.1120	6.7660	1.9167	0.8526	30.8523
No_Both	4.1247	7.9623	6.4008	1.6760	0.6570	29.5430
Proposed	7.2042	13.7953	7.1463	2.0107	0.8847	31.3369

Table 4. Comparison of ablation experiment results on RoadScene dataset

Module	AG	SF	EN	MI	SSIM	PSNR/dB
No_DC	5.0409	10.8995	7.0165	2.7986	0.8418	26.2806
No_MDB	5.3580	11.7188	7.4733	2.8881	0.8490	26.7826
No_Dense	5.3516	11.8289	6.9638	2.6511	0.8428	26.7723
One_Disc	5.1007	11.2476	7.2919	2.8055	0.7882	26.8661
No_Both	4.8786	10.3996	6.8461	2.5975	0.7648	26.3772
Proposed	6.7180	14.8152	7.5761	2.9083	0.8544	27.3352

4. Conclusion

In this chapter, based on the FusionGAN method, an infrared and visible image fusion method for generating adversarial networks with dual paths and dual discriminators is proposed. The neglected information in the source image is obtained by splicing the source image with double-path difference, and the features extracted from the source image under the convolution kernel of different sizes are obtained by using multi-scale decomposition. The dense block added to the source image is used to improve the feature transfer efficiency of the network model. Two discriminators are used to identify the images generated by the generator respectively to avoid the problem that only VI texture is retained and IR information is lost in the single discriminator network. The fusion results of the methods in this chapter retain more characteristic information in IR and VI. Experiments in TNO, RoadScene and MSRS datasets show that, compared with other methods, the subjective evaluation of the method in this chapter is significantly better than that of the comparison method, which can effectively reduce artifacts, make edge information clearer, have higher contrast and better texture details. Most of the values of the objective evaluation index get the optimal or sub-optimal effect, which proves the effectiveness of the method in this chapter. Through the ablation experiment, it is proved that each module of the method in this chapter can improve the fusion result. Compared with the mainstream methods, although the method in this chapter has improved on subjective vision and most evaluation indicators, it still needs to be improved on the processing of low-light night images and the generalization ability in multiple data sets.

5. Conflict of Interest

The authors declare that there are no conflict of interests, we do not have any possible conflicts of interest.

Acknowledgments. This paper was supported by the Science and technology research projects, name: Research on key technologies of image fusion based on deep learning, project number: 242102210187.

References

1. Ma W, Wang K, Li J, et al. Infrared and visible image fusion technology and application: A review[J]. Sensors, 2023, 23(2): 599.
2. Zhang X, Demiris Y. Visible and infrared image fusion using deep learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(8): 10535-10554.
3. Yin S, Wang L, Teng L. Threshold segmentation based on information fusion for object shadow detection in remote sensing images[J]. Computer Science and Information Systems, 2024. doi: 10.2298/CSIS231230023Y.
4. S. Yin, H. Li, Y. Sun, M. Ibrar, and L. Teng. Data Visualization Analysis Based on Explainable Artificial Intelligence: A Survey[J].?IJLAI Transactions on Science and Engineering, vol. 2, no. 2, pp. 13-20, 2024.
5. Tang L, Xiang X, Zhang H, et al. DIVFusion: Darkness-free infrared and visible image fusion[J]. Information Fusion, 2023, 91: 477-493.

6. Zhao Z, Bai H, Zhu Y, et al. DDFM: denoising diffusion model for multi-modality image fusion[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023: 8082-8093.
7. Zhou T, Li Q, Lu H, et al. GAN review: Models and medical image fusion applications[J]. Information Fusion, 2023, 91: 134-148.
8. Yin S, Li H, Laghari A A, et al. An anomaly detection model based on deep auto-encoder and capsule graph convolution via sparrow search algorithm in 6G internet-of-everything[J]. IEEE Internet of Things Journal, vol. 11, no. 18, pp. 29402-29411, 2024.
9. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
10. Tang W, He F, Liu Y, et al. DATFuse: Infrared and visible image fusion via dual attention transformer[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(7): 3159-3172.
11. Vivone G. Multispectral and hyperspectral image fusion in remote sensing: A survey[J]. Information Fusion, 2023, 89: 405-417.
12. Chauhan R B, Shah T V, Shah D H, et al. An overview of image processing for dental diagnosis[J]. Innovation and Emerging Technologies, 2023, 10: 2330001.
13. Wang L, Shoulin Y, Alyami H, et al. A novel deep learning-based single shot multibox detector model for object detection in optical remote sensing images [J]. Geoscience Data Journal, vol. 11, no. 3, pp. 237-251, 2024. https://doi.org/10.1002/gdj3.162.
14. Jiang Y, Yin S. Heterogenous-view occluded expression data recognition based on cycle-consistent adversarial network and K-SVD dictionary learning under intelligent cooperative robot environment[J]. Computer Science and Information Systems, 2023, 20(4): 1869-1883.
15. Deng S Q, Deng L J, Wu X, et al. PSRT: Pyramid shuffle-and-reshuffle transformer for multispectral and hyperspectral image fusion[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-15.
16. Jisi A?and Shoulin Yin. A New Feature Fusion Network for Student Behavior Recognition in Education [J]. Journal of Applied Science and Engineering. vol. 24, no. 2, pp.133-140, 2021.
17. Mijwil M M, Al-Mistarehi A H, Abotaleb M, et al. From Pixels to Diagnoses: Deep Learning's Impact on Medical Image Processing-A Survey[J]. Wasit Journal of Computer and Mathematics Science, 2023, 2(3): 9-15.
18. Alessio S M, Alessio S M. Discrete wavelet transform (DWT)[J]. Digital signal processing and spectral analysis for scientists: concepts and applications, 2016: 645-714.
19. Kingsbury N. A dual-tree complex wavelet transform with improved orthogonality and symmetry properties[C]//Proceedings 2000 international conference on image processing (Cat. No. 00CH37101). IEEE, 2000, 2: 375-378.
20. Zhou J, Cunha A L, Do M N. Nonsubsampled contourlet transform: construction and application in enhancement[C]//IEEE international conference on image processing 2005. IEEE, 2005, 1: I-469.
21. Li H, Wu X J. DenseFuse: A fusion approach to infrared and visible images[J]. IEEE Transactions on Image Processing, 2018, 28(5): 2614-2623.
22. Ma J, Yu W, Liang P, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion[J]. Information fusion, 2019, 48: 11-26.
23. Ma J, Zhang H, Shao Z, et al. GANMcC: A generative adversarial network with multiclassification constraints for infrared and visible image fusion[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 1-14.
24. Shanmugam P, Amali S A M J. Dual-discriminator conditional generative adversarial network optimized with hybrid manta ray foraging optimization and volcano eruption algorithm for hyperspectral anomaly detection[J]. Expert Systems with Applications, 2024, 238: 122058.
25. Zhou H, Wu W, Zhang Y, et al. Semantic-supervised infrared and visible image fusion via a dual-discriminator generative adversarial network[J]. IEEE Transactions on Multimedia, 2021, 25: 635-648.

Biography

Yachao Zhang is with the School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology. Research direction is computer application and AI.
Lijuan Feng is with the School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology. Research direction is computer application and AI.
Jiaxiang Wang is with the School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology. Research direction is computer application and AI.