International Journal of
Human Computations & Intelligence
Vol. 02, Issue 06, Nov/Dec 2023

REVIEW ARTICLE OPEN ACCESS

A Survey on Deep Learning techniques in Image fusion

Vasudha G S . Kusuma Kumari B M

Department of Studies and Research in Computer Applications,
Tumkur University, Tumakuru, Karnataka, India.

DOI: **10.5281/zenodo.10444476**
Received: 18 October 2023 / Revised: 16 November 2023 / Accepted: 01 December 2023

Abstract – In the ever-evolving field of image fusion, the integration of deep learning techniques has led to remarkable advancements in the quality and applicability of fused images. This review work provides a comprehensive overview of state of art deep learning based image fusion techniques. We delve into the fundamental concepts, methodologies and challenges that have emerged in this domain. This work covers various aspects of deep learning-based image fusion, including multi-modal, multi-scale fusion, and cross modality fusion. This work offers insights into the practical applications of deep learning based image fusion across various domains. We highlight the potential benefits and limitations in this dynamic field.

Index Terms – Image fusion, deep learning, CNN, GAN, AE

I. INTRODUCTION

Image fusion is a process of extracting and integrating significant information from various images. The advancement of deep learning has greatly accelerated the progress in the image fusion domain. Before the rise of deep learning, significant research had been conducted in the realm of image fusion. In traditional image fusion techniques like, spatial and frequency domain, image fusion was accomplished by applying mathematical transformations to manually asses the activity level and fusion rules [1][2][3][4]. Traditional fusion methods encompass a range of techniques, including multi-scale transform-based [5], sparse representation-based, subspace-based [6], saliency-based [7], and total variation-based methods [8]. These methods are employed to combine information from various sources or scales effectively. However, the limitations of these methods have become increasingly noticeable. On one hand, to ensure that subsequent feature fusion is feasible, traditional methods are compelled to employ identical transformations for various source images when extracting features. Regrettably, this approach fails to account for the unique characteristics of the source images, potentially resulting in less expressive feature extraction.

MILESTONE
RESEARCH.IN
OPEN ACCESS
Electronic ISSN: **2583-5696**

On the other hand, the conventional feature fusion strategies are so rudimentary that they yield severely restricted fusion performance. The reason for incorporating deep learning into image fusion is to address the constraints associated with traditional methods [1][9] 10]. To begin with, deep learning-based approaches have the capability to employ distinct network branches for specialized feature extraction, allowing for the acquisition of more tailored features.

Furthermore, deep learning-based techniques have the capacity to acquire a more rational feature fusion strategy by leveraging thoughtfully designed loss functions, thereby achieving adaptive feature fusion. There is a dearth of a thorough study and analysis of the most recent deep-learning techniques in various fusion domains. Rest of the paper is organized in different sections. Deep learning techniques in image fusion are covered in Section 2. The real time applications of these techniques in various domains are discussed in section 3. Section 4 of the document discusses various challenges.. Finally, the papers with an outline of major ideas are concluded.

II. DEEP LEARNING IMAGE FUSION TECHNIQUES

Here, in this work, we review current developments on the usage of deep learning in several types of image fusion scenarios, including the sharpening, multi-modal, and image fusion. Most Commonly used deep learning models in image fusion are Convolution Neural Network (CNN), Convolution Sparse Representation (CSR) and Stacked Auto Encoder (SAE).

- **Auto Encoders**

The auto-encoders (AE) model is a variant of feed forward neural networks. This model's two stages—encoding and decoding—are the same as those in other feed forward neural network models [11]. Initially, auto encoder (AE) methods involve training an auto encoder on a publicly available dataset like MS-COCO. In this setup, the encoder focuses on extracting valuable features from input images, while the decoder's role is to reconstruct the original image from these encoded features [12][13]. Subsequently, it becomes evident that the trained auto encoder can effectively tackle two sub-problems within image fusion: feature extraction and image reconstruction. Consequently, the crux of image fusion hinges on crafting effective feature fusion strategies.

- **Convolution Neural Network**

The convolution neural network is typically introduced into the image fusion process in two distinct stages. The initial approach involves the complete implementation of feature extraction, feature fusion, and image reconstruction by employing carefully crafted loss functions and network architectures. [14].

- **Generative Neural network**

In its initial iteration, the GAN was intended to produce more convincing images with unsupervised approach. Subsequently, several computer vision tasks have had considerable success with it [15]. GANs operate on the foundational concept of setting up a competitive game between a generator and a discriminator, wherein one player's objective is to generate data while the other's is to differentiate between real and generated data. The discriminator seeks to discern whether a sample is from the model distribution or the data distribution, while the generator attempts to trick it by creating various samples out of noise as input [16].

III. APPLICATIONS

- **In Multi exposure Image fusion**

The most popular techniques for multi-exposure picture fusion are CNN and GAN. To create a fusion map, which is then utilised to create the final fused image, CNN uses trained networks to extract features and pixel positions from source images with varying exposures [17]. Finding a high performance non-reference metric to assess fused results is the challenging aspect with CNN. To achieve a high-quality multi-exposure fusion, the GAN approach relies on all of the original image's information, including the scene's structure and exposure condition.

- **Multi Focus Image Fusions**

CNN and GAN are most widely employed in the fusing multi-focus images. CNN in this instance is based on a decision map that can identify clear and blurred pixels in order to build a decision map. The decision map is used to obtain final fused image by selecting and combining pixels [18] [19]. The decision map-based GAN approach often uses the generator's decision map to obtain the fused result before using adversarial learning to bring the fused result as close to the reference full-clear image as possible [20]. To ensure richer texture and visual fidelity, the GAN technique optimises the reconstructed fused image [21].

- **Infrared and visible image Fusion**

High-contrast, texture-rich infrared and visible image fusion employs AE, CNN, and GAN techniques [22]. In AE, the encoder is utilised to extract useful features while the decoder reconstructs the input image. The improvement of visible and infrared fused images is not possible with AE since the procedures are hand calculated and not learnable [23][24]. CNN constraints the performance of image fusion because of the usage of pre-trained networks [25]. High-performance fused images are produced using the GAN approach, however maintaining the correct balance of generator to discriminator during training is difficult [22].

- **In medical Imaging**

The CNN and GAN fusion techniques are frequently employed in medical image fusion [26] [27] [28]. In the merging of medical images, GAN guarantees outstanding performance. There's a good chance that function information will cover over texture information. The GAN in medical image fusion must overcome this obstacle [22].

IV. CHALLENGES

In recent years, image fusion has been widely used in a variety of applications, including remote sensing, photography, surveillance, and medical diagnosis. Here, several significant challenges pertaining to various fields are explored.

- Finding an authentic and high-quality dataset is a challenging aspect of research in the field of image fusion.
- The majority of deep learning techniques for image fusion make the assumption that the images are already registered. Digital and multimodal photographs, however, are not registered because the principle sensors differ, the resolution of the source images varies [29].
- Extracting all the data necessary to create an efficient fused image is another challenge [30].
- The fused images frequently serve as useful input for later applications. However, the majority of fusion algorithms don't take the application of the fused image into account when fusing them.
- Real-time picture fusion with high performance is required for several practical applications. Therefore developing a real time image fusion technique is need of the day.

V. CONCLUSION

In summary, deep learning techniques have revolutionized image fusion, enabling applications across diverse domains such as medicine, remote sensing, and surveillance. Despite remarkable progress, challenges like computational complexity and interpretability persist. Nevertheless, the potential benefits of deep learning in image fusion are undeniable, promising improved decision-making and data analysis. As technology evolves, image fusion's role in enhancing our understanding of complex data and driving advancements in numerous fields remains pivotal. With ongoing research and innovation, the future of deep learning-based image fusion appears bright, poised to continue shaping our world in meaningful ways. Looking ahead, the future scope of image fusion through deep learning holds exciting possibilities. Anticipated developments include more efficient models, enhanced interpretability, and adaptation to emerging imaging technologies like hyper spectral and 3D imaging. These advancements will further solidify image fusion's position as a transformative tool in various applications. In conclusion, while challenges persist, the ever-evolving landscape of deep learning-based image fusion promises a future filled with innovative solutions and continued contributions to a wide array of domains.

REFsERENCE

1. Pajares, G., & De La Cruz, J. M. (2004). A wavelet-based image fusion tutorial. *Pattern recognition*, *37*(9), 1855-1872.
2. Li, S., Yang, B., & Hu, J. (2011). Performance comparison of different multi-resolution transforms for image fusion. *Information Fusion*, *12*(2), 74-84.
3. Mo, Y., Kang, X., Duan, P., Sun, B., & Li, S. (2021). Attribute filter based infrared and visible image fusion. *Information Fusion*, *75*, 41-54.
4. Li, S., Kang, X., & Hu, J. (2013). Image fusion with guided filtering. *IEEE Transactions on Image processing*, *22*(7), 2864-2875.
5. Liu, Y., Liu, S., & Wang, Z. (2015). A general framework for image fusion based on multi-scale transform and sparse representation. *Information fusion*, *24*, 147-164.

6. Harsanyi, J. C., & Chang, C. I. (1994). Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. *IEEE Transactions on geoscience and remote sensing*, *32*(4), 779-785.
7. Han, J., Pauwels, E. J., & De Zeeuw, P. (2013). Fast saliency-aware multi-modality image fusion. *Neurocomputing*, *111*, 70-80.
8. Ma, J., Chen, C., Li, C., & Huang, J. (2016). Infrared and visible image fusion via gradient transfer and total variation minimization. *Information Fusion*, *31*, 100-109.
9. Zhang, H., Xu, H., Xiao, Y., Guo, X., & Ma, J. (2020, April). Rethinking the image fusion: A fast unified image fusion network based on proportional maintenance of gradient and intensity. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 07, pp. 12797-12804).
10. Zhang, Y., Liu, Y., Sun, P., Yan, H., Zhao, X., & Zhang, L. (2020). IFCNN: A general image fusion framework based on convolutional neural network. *Information Fusion*, *54*, 99-118.
11. Chen, X., & Konukoglu, E. (2018). Unsupervised detection of lesions in brain MRI using constrained adversarial auto-encoders. *arXiv preprint arXiv:1806.04972*.
12. Jian, L., Yang, X., Liu, Z., Jeon, G., Gao, M., & Chisholm, D. (2020). SEDRFuse: A symmetric encoder–decoder with residual block network for infrared and visible image fusion. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-15.
13. Long, Y., Jia, H., Zhong, Y., Jiang, Y., & Jia, Y. (2021). RXDNFuse: A aggregated residual dense network for infrared and visible image fusion. *Information Fusion*, *69*, 128-141.
14. Ram Prabhakar, K., Sai Srikar, V., & Venkatesh Babu, R. (2017). Deepfuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs. In *Proceedings of the IEEE international conference on computer vision* (pp. 4714-4722).
15. Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).
16. Ma, J., Yu, W., Chen, C., Liang, P., Guo, X., & Jiang, J. (2020). Pan-GAN: An unsupervised pan-sharpening method for remote sensing image fusion. *Information Fusion*, *62*, 110-120.
17. Li, H., & Zhang, L. (2018, October). Multi-exposure fusion with CNN features. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (pp. 1723-1727). IEEE.
18. Li, J., Guo, X., Lu, G., Zhang, B., Xu, Y., Wu, F., & Zhang, D. (2020). DRPL: Deep regression pair learning for multi-focus image fusion. *IEEE Transactions on Image Processing*, *29*, 4816-4831.
19. Guo, X., Nie, R., Cao, J., Zhou, D., Mei, L., & He, K. (2019). FuseGAN: Learning to fuse multi-focus image via conditional generative adversarial network. *IEEE Transactions on Multimedia*, *21*(8), 1982-1996.
20. Zhang, H., Le, Z., Shao, Z., Xu, H., & Ma, J. (2021). MFF-GAN: An unsupervised generative adversarial network with adaptive and gradient joint constraints for multi-focus image fusion. *Information Fusion*, *66*, 40-53.
21. Zhang, H., Xu, H., Tian, X., Jiang, J., & Ma, J. (2021). Image fusion meets deep learning: A survey and perspective. *Information Fusion*, *76*, 323-336.
22. Jian, L., Yang, X., Liu, Z., Jeon, G., Gao, M., & Chisholm, D. (2020). SEDRFuse: A symmetric encoder–decoder with residual block network for infrared and visible image fusion. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-15.
23. Ahmed, S. T., Kumar, V., & Kim, J. (2023). AITel: eHealth Augmented Intelligence based Telemedicine Resource Recommendation Framework for IoT devices in Smart cities. *IEEE Internet of Things Journal*.
24. Li, H., Wu, X. J., & Kittler, J. (2018, August). Infrared and visible image fusion using a deep learning framework. In *2018 24th international conference on pattern recognition (ICPR)* (pp. 2705-2710). IEEE.
25. Lahoud, F., & Süsstrunk, S. (2019, July). Zero-learning fast medical image fusion. In *2019 22th international conference on information fusion (FUSION)* (pp. 1-8). IEEE.
26. Xu, H., Ma, J., Le, Z., Jiang, J., & Guo, X. (2020, April). Fusiondn: A unified densely connected network for image fusion. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 07, pp. 12484-12491).
27. Zhao, C., Wang, T., & Lei, B. (2021). Medical image fusion method based on dense block and deep convolutional generative adversarial network. *Neural Computing and Applications*, *33*, 6595-6610.
28. Ma, J., Jiang, X., Fan, A., Jiang, J., & Yan, J. (2021). Image matching from handcrafted to deep features: A survey. *International Journal of Computer Vision*, *129*, 23-79.

29. Yang, J., Ma, Y., Yao, W., & Lu, W. T. (2008). A spatial domain and frequency domain integrated approach to fusion multifocus images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *37*(PART B7).
30. Thouheed Ahmed, S., & Sandhya, M. (2019). Real-time biomedical recursive images detection algorithm for Indian telemedicine environment. In *Cognitive Informatics and Soft Computing: Proceeding of CISC 2017* (pp. 723-731). Springer Singapore.
31. Swamy, R., Ahmed, S. T., Thanuja, K., Ashwini, S., Siddiqha, S., & Fathima, A. (2021, January). Diagnosing the level of Glaucoma from Fundus Image Using Empirical Wavelet Transform. In *Proceedings of the First International Conference on Advanced Scientific Innovation in Science, Engineering and Technology, ICASISET 2020, 16-17 May 2020, Chennai, India*.
32. Ambika, B. J., Guptha, N. S., & Siddiqha, S. A. (2023). Anaemia Estimation for Patients Using Lasso And Ridge Regression Algorithms. *Milestone Transactions on Medical Technometrics*, *1*(2), 53-63.