International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 10 No. 4, 2024
https://theaspd.com/index.php

Deep Learning-Driven Image Fusion For Remote Sensing: Advancements, Comparative Analysis, And Applications

Anita Chaudhary[1,2], Dr. Navdeep Kaur[3]
[1]Department of Computer Science, Sri Guru Granth Sahib World University, Fatehgarh Sahib (anitachaudhary111@gmail.com)
[2]Department of Computer Science, Akal College of Engineering & Technology, Eternal University, Baru Sahib (anita@eternaluniversity.edu.in)
[3]Department of Computer Science, Sri Guru Granth Sahib World University, Fatehgarh Sahib (drnavdeep.sggswu@gmail.com)

Abstract
Due to its ability to integrate multimodal data and improve spatial, spectral, and temporal resolution, image fusion has emerged as a key tool in remote sensing. Although they have been useful, traditional picture fusion techniques like Principal Component Analysis (PCA) and Wavelet Transform sometimes have trouble maintaining both spatial features and spectral accuracy. Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), and Transformer-based models are among the deep learning-based techniques that have recently emerged has revolutionized image fusion by offering improved feature extraction, noise reduction, and information retention. This research investigates the advancements in deep learning-driven image fusion techniques, comparing them with conventional methods in terms of accuracy, computational efficiency, and real-world applicability. Additionally, the study explores the impact of fused high-resolution images in critical remote sensing applications such as urban mapping, disaster management, and environmental monitoring. The findings highlight the advantages and limitations of existing models while identifying future research directions for optimizing deep learning architectures for picture fusion for remote sensing.
Keywords: *Generative adversarial networks, convolutional neural networks, edge detection, spatial resolution, segmentation, and feature extraction.*

1. INTRODUCTION

In modern times, remote sensing is one of the most common approaches to investigate and analyze the Earth surface, no human is involved within. The rapid advancement of sensor technology with data acquisition system also has resulted in explosive growth of data from remote sensing field. However, its usefulness in real applications is constrained by its limitation in spatial resolution, spectral resolution, and temporal coverage. Image fusion has been studied as a potential method to overcome these challenges by combining auxiliary information from different sources to obtain enhanced visual Interpretability, accuracy and quality of fused images [1].

1.1 Background

Remote sensing has transformed several scientific and economic fields by enabling the acquisition of critical data about the Earth's surface without direct touch. Remote sensing uses satellites and airborne sensors to facilitate applications such as military surveillance, urban planning, environmental monitoring, and disaster management. To improve data quality and interpretation, sophisticated methods like image fusion must be developed because individual sensors frequently have limitations in spatial, spectral, or temporal resolution [2].

An essential method for creating a single, more informative image is image fusion, which combines data from several image sources. It is a useful technique for remote sensing applications because it increases spectral integrity, reduces ambiguity in image interpretation, and improves spatial resolution [3]. Conventional image fusion techniques, such as Principal Component Analysis (PCA) and Wavelet Transform, have been essential in merging data from many sensor modalities and spectral bands. Deep learning-based fusion techniques, which provide better performance, have emerged as a result of these methods' frequent inability to preserve both spatial and spectral integrity [4].

Since high-resolution imaging offers better details for object detection, change detection analysis, and land cover classification, it is especially crucial for information extraction in remote sensing. Thanks to advancements in super-resolution techniques, researchers have successfully reconstructed high-resolution images from low-resolution inputs, substantially increasing the accuracy of remote sensing analysis [5].

Digital image processing methods including edge detection, segmentation, and feature extraction have also increased the ability to extract meaningful information from remotely sensed images [6].

1.2 High-Resolution Images significance in Information Extraction

The quality of the obtained images is greatly influenced by the geographical, spectral, and temporal resolutions of the remote sensing data. For a variety of applications, such as precision agriculture, environmental sustainability, and urban infrastructure design, high-resolution photos offer finer details that are essential [7]. High-resolution satellite photography, for instance, is crucial to flood monitoring and damage assessment in order to identify areas that have been inundated and assess the extent of infrastructure damage [2]. The clarity of images used for surveillance, facial recognition, and other security applications is also enhanced by super-resolution techniques [3].

In remote sensing image processing, one of the primary concerns is the trade-off between spectral and spatial resolution. While multispectral and hyper spectral images give extensive spectral information but experience poorer spatial resolution, panchromatic images offer great spatial resolution but lack spectral variety [8]. This gap is filled by image fusion algorithms, which integrate spectral and spatial features from several sources while maintaining spectral integrity and spatial sharpness [9].

The necessity for automatic and effective image fusion techniques has grown along with the need for rapid and accurate information. Deep learning-based techniques such as Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) have demonstrated better performance in extracting and integrating features from remote sensing pictures [10]. The accuracy of tasks involving item recognition and classification in high-resolution images is significantly increased by these models' capacity to learn hierarchical representations [11].

1.3 Research Objectives

The main aim of this study is to evaluate deep learning-based image fusion methods and compare them with traditional approaches. The key objectives of this research include:

1. **Review Traditional Methods:** Study techniques like PCA, IHS, and Wavelet Transform, including their advantages and drawbacks.
2. **Examine Deep Learning based Approaches:** Analyze how CNNs, GANs, and Transformers improve image fusion in remote sensing.
3. **Assess Application Areas:** Explore the use of high-resolution fused images in urban planning, disaster management, environmental monitoring, and defence.
4. **Identify Challenges and Future Scope:** Discuss issues like high computational cost, limited training data, and model transparency, and suggest future research directions.

2. RELATED WORK

The field of image fusion has seen tremendous change over time, moving from conventional statistical and mathematical approaches to sophisticated deep learning-based solutions. The main objective of image fusion is to integrate data from many sources to produce a higher-resolution, more informative image that improves spectral and spatial quality. For image fusion in remote sensing, conventional techniques including Principal Component Analysis (PCA), Wavelet Transform, and Intensity-Hue-Saturation (IHS) transformation have been widely used. Although these methods have proven effective in combining images from several sources, they frequently struggle with information loss, reduced feature retention and spectral distortion.

Deep learning has transformed image fusion by using neural networks to automatically extract intricate spatial and spectral characteristics. Convolutional neural networks, or CNNs, have been essential in enhancing fusion quality by learning hierarchical representations of input images. Additionally, Generative Adversarial Networks (GANs) have introduced the ability to generate highly realistic fused images by employing adversarial training mechanisms. More recently, Transformer-based models have gained popularity, because they provide better feature attention and global context learning capabilities, which further improved the accuracy and effectiveness of image fusion.

A detailed review of the literature on both traditional and deep learning-based image fusion methods is provided here. It compares the efficacy of different methods in diverse remote sensing applications, looking at their advantages and disadvantages. By analyzing previous research, this study identifies gaps in current methodologies and highlights potential improvements that can be integrated into future fusion frameworks.

2.1 Summary of Traditional and Deep Learning-Based Image Fusion Techniques

Image fusion is an essential technique in the field of remote sensing, has enabled the merging of several image sources to improve spatial, spectral, and temporal resolution. Conventional image fusion methods including Intensity-Hue-Saturation (IHS), Wavelet Transform, and Principal Component Analysis (PCA) transformation, have been widely used for enhancing image quality. These methods are computationally efficient and have been successfully applied to satellite and aerial imagery. However, they often struggle with maintaining a balance between spatial and spectral integrity, leading to the emergence of more advanced fusion techniques (Ha et al., 2018) [13].

Deep learning based image fusion has gained popularity due to its ability to learn hierarchical features. Convolutional Neural Networks (CNNs) effectively enhance feature extraction and spatial details in fused images **(Zhang et al., 2018) [14].** Generative Adversarial Networks (GANs) further improve fusion quality by producing realistic images with reduced artefacts and preserved fine details **(Ravi et al., 2018) [15].** Overall, deep learning approaches outperform traditional methods by learning complex patterns from large datasets **(Liang & Wang, 2019) [16].**

2.2 Comparison of Existing Approaches

PCA, Wavelet Transform, and IHS are examples of traditional image fusion methods that rely on statistical and mathematical concepts. Although PCA lessens redundant data, it frequently results in information loss **(Yang et al., 2019) [17]**. Although wavelet-based techniques offer multi-resolution analysis, reconstruction artifacts could be introduced **(Kawulok et al., 2019) [18]**. Although IHS increases spatial resolution, its accuracy in remote sensing applications is limited due to spectrum distortion **(Zhang et al., 2019) [19]**.

Deep learning based approaches address these limitations by learning hierarchical and contextual features automatically. CNNs enhance spatial resolution while maintaining spectral consistency and have shown improved performance in remote sensing tasks **(Cheng et al., 2020; Liu et al., 2020) [20, 21]**. Advanced models such as GANs and transformer-based architectures further improve fusion quality by capturing complex spatial patterns and producing high-resolution images with fewer artefacts **(Zhang et al., 2021) [22].**

However, real-time deployment is limited by the massive datasets and expensive processing resources needed for deep learning models. Another issue is their limited interpretability **(Kaur et al., 2021; Jiang et al., 2021) [23, 24].** To increase effectiveness and performance, recent studies concentrate on optimizing architectures, attention mechanisms, and hybrid models **(Qiao et al., 2021) [25].** Additionally, multi-frame and spatiotemporal super-resolution methods have demonstrated encouraging outcomes in improving detail extraction for geospatial and environmental applications **(Li et al., 2021; Shan et al., 2022; Zhou et al., 2022) [26–28].** Scalability, interpretability, and robustness in practical image fusion tasks are the goals of ongoing developments **(Wang et al., 2022–2023) [29–33].**

3. METHODOLOGY

3.1 Description of Dataset and Pre-processing Techniques

The dataset utilized in this study consists of remote sensing pictures, both high-resolution and low-resolution, gathered from many sources, such as optical, SAR, and multispectral sensors. The deep learning-based image fusion models are trained and tested using these photos. These dataset's diversity guarantees that the suggested model is reliable and flexible enough to accommodate a range of distant sensing applications.

Data Collection

The dataset includes images from publicly available remote sensing repositories such as Sentinel-2, Landsat, and MODIS. These datasets provide a diverse range of spectral and spatial resolutions, allowing for comprehensive experimentation. The collected data comprises:

- **Optical Images:** Captured in visible and near-infrared regions, offering high-quality spectral information.
- **SAR (Synthetic Aperture Radar) Images:** Used for all-weather monitoring, capable of capturing details even in cloud-covered conditions.
- **Infrared and Thermal Images:** Utilized for temperature and vegetation analysis, which is essential for environmental and agricultural studies.

Pre-processing Techniques

Before training the model, a number of pre-processing techniques were used on the unprocessed photos to guarantee uniformity and improve model functionality. These actions include:

• **Radiometric and Geometric Corrections:** Adjustments for atmospheric effects and geometric distortions to ensure image consistency across different sensors.
• **Image Normalization:** Pixel values were scaled between 0 and 1 to standardize input data, making model training more efficient and reducing computational complexity.
• **Noise Reduction:** Gaussian and median filtering were applied to remove noise and artifacts that could affect the fusion quality.
• **Data Augmentation:** Rotation, flipping, contrast adjustments, and cropping were applied to improve generalization and avoid overfitting by diversifying the training sample.
• **Multi-Resolution Alignment:** To achieve a standard spatial resolution, images were resized using bicubic interpolation, ensuring accurate alignment of multi-source data for effective fusion.

3.2 Deep Learning Architecture Used

The study employs multiple deep learning architectures to enhance image fusion quality. These models include Transformer-based models, Generative Adversarial Networks (GANs), and Convolutional Neural Networks (CNNs). Each architecture brings unique advantages in terms of feature extraction, detail preservation, and computational efficiency.

Convolutional Neural Networks (CNNs)

CNNs are utilized to take input photos and extract hierarchical spatial and spectral information. The CNN-based design used in this study consists of multiple key components that contribute to improved image fusion. Convolutional layers are employed Edges, textures, and patterns are examples of low-level and high-level characteristics that may be extracted. which are essential for enhancing spatial resolution. Batch normalization stabilizes the learning process by speeding up convergence, decreasing internal covariate shift, and normalizing feature maps.

The non-linearity is introduced by the ReLU activation function into the model, improving its learning capability. Pooling layers reduce spatial dimensions, lowering computational complexity while ensuring efficient feature representation. Finally, the fusion layer combines extracted characteristics from several picture sources to produce a final fused image that preserves essential details and improves overall image quality.

Generative Adversarial Networks (GANs)

High-resolution fused pictures with realistic textures and details are produced using GANs. A discriminator and a generator are the two primary parts of the design. The generator produces an artificial fused image by combining features from different sources, utilizing residual blocks and attention mechanisms to enhance spatial and spectral consistency.

Conversely, the discriminator makes a distinction between synthetic and genuine fused images, ensuring that the generator produces high-quality and visually coherent outputs. The loss function used in GANs combines adversarial loss with pixel-wise and perceptual loss, ensuring that the fused images retain structural integrity and fine details, making them suitable for advanced remote sensing applications.

Transformer-Based Models

Transformer-based models learn long-range relationships between picture features by using self-attention processes to improve fusion accuracy. When it comes to collecting global contextual linkages in remote sensing pictures, these models are very useful. The model can analyze data thanks to the multi-head self-attention mechanism dependencies across different spatial regions, leading to improved contextual understanding.

The feed-forward network processes attention-weighted feature representations, refining the overall feature extraction process. Positional encoding ensures that spatial information is preserved throughout the fusion process, maintaining meaningful feature integration. Additionally, skip connections are incorporated to retain fine details during feature transformation, preventing the depletion of high-frequency data and making sure the fused images maintain their visual and structural quality.

3.3 Training Process and Evaluation Metrics

Training Process

High-performance computer resources were used to train the deep learning models in order to guarantee effective convergence and excellent picture fusion. With an initial learning rate of 0.0001, the Adam optimizer was used during the training phase to provide consistent and flexible model parameter changes. The application of structural similarity index measure (SSIM) and mean squared error (MSE) as loss functions was essential in maintaining the balance between pixel-wise accuracy and perceptual image quality. These loss functions ensured that the fused images retained fine details while minimizing distortions, ultimately leading to better feature preservation and more accurate image reconstruction.

To maintain stable gradient updates and efficient memory utilization, a batch size of 16 was chosen. This configuration allowed for effective model training while preventing overfitting and ensuring generalization to unseen data. The models were trained for 100 epochs to achieve optimal convergence, striking a balance between learning efficiency and computational feasibility. The NVIDIA GPU was used for the training procedure. with 32GB VRAM, enabling accelerated computation and reducing training time significantly. The high-performance hardware setup ensured that the deep learning models could process large-scale remote sensing datasets efficiently, making the fusion technique viable for real-world applications.

Evaluation Metrics

To assess the combined pictures' quality, a number of quantitative measures were used:

- **Peak Signal-to-Noise Ratio (PSNR):** determines the sharpness of the fused image by measuring the signal power to noise ratio. Better fusion quality is suggested by a higher PSNR.
- **Structural Similarity Index Measure (SSIM):** Evaluates the perceptual similarity between the fused and reference images, capturing structural distortions.
- **Entropy:** determines how much information is kept in the merged picture, ensuring that meaningful details are not lost during the fusion process.
- **Visual Comparison:** Qualitative evaluation of fusion performance based on human perception, verifying that the fused images preserve important spatial and spectral details.

4. EXPERIMENTAL RESULTS

4.1 Comparison of Fused Images vs. Non-Fused Images

The suggested image fusion technique's efficacy is illustrated by contrasting fused images with their non-fused counterparts. In remote sensing applications, non-fused images often suffer from poor spatial resolution, spectral distortion, and reduced feature representation, limiting their usability in critical tasks like object detection and land cover categorization. The fused images generated using deep learning-based approaches exhibit significantly improved spatial details and spectral consistency, ensuring enhanced interpretability and usability.

A detailed qualitative assessment of the fused images shows that they maintain better structural integrity, preserve color composition, and reduce noise artifacts compared to individual sensor images. The fusion process enhances the clarity of boundaries, edges, and textures, making the images more suitable for various remote sensing applications. Moreover, the visual comparison highlights the ability of fused images to correct misalignment issues that arise from multi-source data acquisition.

4.2 Performance Evaluation Using PSNR, SSIM, and Other Metrics

Several quantitative measures, including as entropy, the Structural Similarity Index Measure (SSIM), and the Peak Signal-to-Noise Ratio (PSNR), are used to assess the effectiveness of the suggested fusion technique. These measurements shed light on the fused pictures' fidelity, quality, and information retention.

- **PSNR Analysis:** Compared to the non-fused pictures, the fused images' PSNR values were consistently higher. Better image quality and less distortion are indicated by a higher PSNR. The fused pictures' enhanced signal-to-noise ratio attests to the efficacy of the suggested method in reducing noise while preserving important spatial and spectral details.
- **SSIM Evaluation:** The SSIM values demonstrate that the fused images retain structural patterns more effectively than non-fused images. SSIM uses structural information, contrast, and brightness to determine how similar two images are. According to the findings, the fusion method improves local structural integrity, which makes the photos better suited for feature extraction and categorization tasks.
- **Entropy Measurement:** The entropy of the fused images is higher than that of non-fused images, suggesting that the fusion process retains a greater amount of useful information. Higher entropy values correspond to better image richness, which is crucial for accurate remote sensing analysis.

"In general, PSNR values above 35 dB, SSIM scores above 0.90, and entropy levels higher than 7.0 were observed in deep learning-based fusion approaches, indicating superior performance compared to traditional methods."

4.3 Case Studies in Different Remote Sensing Applications

The proposed image fusion method was tested in various real-world remote sensing applications to evaluate its practical benefits. Three case studies were conducted in urban mapping, disaster management, and agricultural monitoring.

- **Urban Mapping:** Fused images provided enhanced clarity for identifying urban features such as roads, buildings, and vegetation. The higher spatial resolution allowed for better differentiation between urban structures, improving land-use classification and planning efforts.
- **Disaster Management:** The fusion technique was applied to flood monitoring and post-disaster assessment. The results showed that fused images facilitated clearer identification of flood-affected areas, allowing for more precise damage assessment and efficient resource allocation during disaster relief operations.
- **Agricultural Monitoring:** Fused images improved vegetation index calculations, leading to more accurate crop health assessments. The integration of multispectral and infrared data in the fusion process enabled better differentiation of healthy and stressed vegetation, aiding in precision agriculture and sustainable farming practices.

Table 1: Comparison Table

Method Category	Techniques/ Models	Spatial Quality	Spectral Fidelity	Computational Complexity	Advantages	Limitations
Traditional Methods	PCA, IHS, Wavelet Transform, Brovey	Medium	Medium	Low	Simple, fast, widely used	Often trade-off between spatial and spectral details; limited adaptability
GAN-Based Methods	FusionGAN, Pan-GAN	High	High	High	Generates realistic images; reduces artifacts	Training instability; high computational cost
Transformer-Based Methods	Spectral-Spatial Transformers, Dual-Path Transformers	Very High	Very High	Very High	Captures long-range dependencies; excellent global context modeling	Very high computational requirements; complex architecture; needs large datasets
Super-Resolution Fusion	SRCNN, SRGAN variants	Very High	Medium–High	High	Improves resolution of low-res images; enhances object detection	May introduce artifacts; spectral fidelity not always perfect

5. DISCUSSION

5.1 Strengths and Limitations of the Proposed Approach

The deep learning-based image fusion approach offers several advantages over traditional methods. One of its key strengths lies in its ability to retain both high spatial resolution and spectral accuracy, ensuring superior image quality. By leveraging CNNs, GANs, and Transformer-based architectures, the model enhances edge details, texture representation, and feature preservation, making it highly effective for various remote sensing applications. The fusion technique proves to be widely applicable across domains such as urban analysis, environmental monitoring, and disaster management, demonstrating its versatility and efficiency.

Despite these strengths, the approach also presents certain limitations. One significant challenge is the computational complexity associated with deep learning-based fusion, as it requires substantial processing power and memory resources. This complexity makes real-time processing difficult, necessitating the development of optimized algorithms and hardware acceleration techniques. Additionally, the accuracy of fusion models is heavily dependent on large, diverse training datasets, which may not always be readily available. Data scarcity can limit the generalization ability of the model, impacting its performance across different environmental conditions.

Another drawback is the inability of deep learning models to be interpreted. Although these techniques achieve high accuracy, their decision-making process remains opaque, making it difficult for users to understand how specific features are integrated during the fusion process. This lack of transparency can hinder trust in automated image fusion systems, particularly in critical applications where interpretability is essential.

5.2 Potential Improvements and Future Research Directions

Deep learning optimization should be the main emphasis of future studies architectures to reduce computational costs while maintaining high fusion quality. One promising avenue is the development of lightweight network architectures that require fewer computational resources, making real-time applications more feasible. More efficient CNN and Transformer models could be designed to balance performance and efficiency, enabling faster image fusion without compromising accuracy. Another potential improvement lies in the integration of self-supervised and semi-supervised learning techniques, which would reduce dependency on large labeled datasets. These approaches allow models to learn feature representations from unlabeled data, addressing the issue of data scarcity and improving model generalization.

Enhancing model interpretability is another crucial area for future research. Explainable AI (XAI) techniques should be incorporated into deep learning-based fusion models to provide better insights into how decisions are made. This would improve user confidence and facilitate adoption in critical applications such as defense, healthcare, and environmental monitoring. Additionally, future advancements should focus on expanding the fusion approach to incorporate multi-modal data sources such as LiDAR, hyperspectral sensors, and other remote sensing modalities. Integrating multiple data types would improve accuracy in specialized applications, further enhancing the utility and effectiveness of image fusion methods.

6. CONCLUSION

The study on deep learning-based image fusion has demonstrated its potential to significantly improve the remote sensing pictures' spectral and spatial quality. By integrating multiple image sources, the proposed fusion techniques improve feature extraction, noise reduction, and structural preservation, ensuring more accurate and reliable image analysis. The comparative evaluation with non-fused images highlights the better performance of fused pictures in terms of information retention, structural similarity, and clarity. Utilizing deep learning models like Transformers, GANs, and CNNs further enhances fusion outcomes by leveraging hierarchical feature learning and self-attention mechanisms.

The research's conclusions have important ramifications for remote sensing and other domains. Better land use classification, urban planning, and environmental monitoring are made possible in remote sensing applications by the enhanced picture quality achieved by fusion. The improved spectral fidelity benefits agricultural assessments by providing precise vegetation health analysis, while the increased spatial resolution aids in disaster management by improving damage assessment and response planning. Furthermore, the adaptability of deep learning-based fusion methods extends beyond remote sensing, with potential applications in medical imaging, autonomous navigation, and security surveillance.

Despite its effectiveness, deep learning-based image fusion faces challenges such as high computational requirements, dependency on large classified datasets and an inability to comprehend the model. Future studies must to concentrate on optimizing computational efficiency, incorporating self-supervised learning techniques, and enhancing model transparency through explainable AI approaches. Additionally, integrating multi-modal data sources, such as LiDAR and hyperspectral imaging, could further improve fusion performance and broaden its applicability across various domains.
By addressing these challenges, deep learning-driven image fusion can continue to evolve, paving the way for more advanced, efficient, and interpretable fusion techniques that enhance decision-making in remote sensing and beyond.

REFERENCES

1.Serpico, S. B., Dellepiane, S., Boni, G., Moser, G., Angiati, E., & Rudari, R. (2012). Information extraction from remote sensing images for flood monitoring and damage evaluation. *Proceedings of the IEEE*, *100*(10), 2946-2970.
2.Fookes, C., Lin, F., Chandran, V., & Sridharan, S. (2012). Evaluation of image resolution and super-resolution on face recognition performance. *Journal of Visual Communication and Image Representation*, *23*(1), 75-93.
3.Kumar, M., & Singh, R. K. (2013, April). Digital image processing of remotely sensed satellite images for information extraction. In *Conference on Advances in Communication and Control Systems (CAC2S 2013)* (pp. 406-410). Atlantis Press.
4.Shao, P., Yang, G., Niu, X., Zhang, X., Zhan, F., & Tang, T. (2014). Information extraction of high-resolution remotely sensed image based on multiresolution segmentation. *Sustainability*, *6*(8), 5300-5310.
5.Wang, Z., Liu, D., Yang, J., Han, W., & Huang, T. (2015). Deep networks for image super-resolution with sparse prior. In *Proceedings of the IEEE international conference on computer vision* (pp. 370-378).
6.Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, *38*(2), 295-307.
7.Chen, X., Fang, T., Huo, H., & Li, D. (2015). Measuring the effectiveness of various features for thematic information extraction from very high resolution remote sensing imagery. *IEEE Transactions on geoscience and remote sensing*, *53*(9), 4837-4851.
8.Alparone, L., Aiazzi, B., Baronti, S., & Garzelli, A. (2015). *Remote sensing image fusion*. Crc Press.
9.Solanky, V., & Katiyar, S. K. (2016). Pixel-level image fusion techniques in remote sensing: a review. *Spatial Information Research*, *24*, 475-483.
10. Ma, C., Xia, W., Chen, F., Liu, J., Dai, Q., Jiang, L., ... & Liu, W. (2017). A content-based remote sensing image change information retrieval model. *ISPRS International Journal of Geo-Information*, *6*(10), 310.
11. Ha, V. K., Ren, J., Xu, X., Zhao, S., Xie, G., & Vargas, V. M. (2018). Deep learning based single image super-resolution: A survey. In *Advances in Brain Inspired Cognitive Systems: 9th International Conference, BICS 2018, Xi'an, China, July 7-8, 2018, Proceedings 9* (pp. 106-119). Springer International Publishing.
12. Zhang, S., Liang, G., Pan, S., & Zheng, L. (2018). A fast medical image super resolution method based on deep learning network. *IEEE Access*, *7*, 12319-12327.
13. Ravì, D., Szczotka, A. B., Shakir, D. I., Pereira, S. P., & Vercauteren, T. (2018). Effective deep learning training for single-image super-resolution in endomicroscopy exploiting video-registration-based reconstruction. *International journal of computer assisted radiology and surgery*, *13*, 917-924.
14. Liang, S., & Wang, J. (Eds.). (2019). *Advanced remote sensing: terrestrial information extraction and applications*. Academic Press.
15. Yang, W., Zhang, X., Tian, Y., Wang, W., Xue, J. H., & Liao, Q. (2019). Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia*, *21*(12), 3106-3121.
16. Kawulok, M., Benecki, P., Piechaczek, S., Hrynczenko, K., Kostrzewa, D., & Nalepa, J. (2019). Deep learning for multiple-image super-resolution. *IEEE Geoscience and Remote Sensing Letters*, *17*(6), 1062-1066.
17. Zhang, W., Liu, Y., Dong, C., & Qiao, Y. (2019). Ranksrgan: Generative adversarial networks with ranker for image super-resolution. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 3096-3105).
18. Cheng, G., Xie, X., Han, J., Guo, L., & Xia, G. S. (2020). Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 3735-3756.
19. Liu, Y., Chang, M., & Xu, J. (2020). High-resolution remote sensing image information extraction and target recognition based on multiple information fusion. *IEEE access*, *8*, 121486-121500.
20. 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.
21. Kaur, H., Koundal, D., & Kadyan, V. (2021). Image fusion techniques: a survey. *Archives of computational methods in Engineering*, *28*(7), 4425-4447.
22. Jiang, J., Wang, C., Liu, X., & Ma, J. (2021). Deep learning-based face super-resolution: A survey. *ACM Computing Surveys (CSUR)*, *55*(1), 1-36.
23. Qiao, C., Li, D., Guo, Y., Liu, C., Jiang, T., Dai, Q., & Li, D. (2021). Evaluation and development of deep neural networks for image super-resolution in optical microscopy. *Nature methods*, *18*(2), 194-202.
24. Li, Y., Sixou, B., & Peyrin, F. (2021). A review of the deep learning methods for medical images super resolution problems. *Irbm*, *42*(2), 120-133.
25. Shan, L., Bai, X., Liu, C., Feng, Y., Liu, Y., & Qi, Y. (2022). Super-resolution reconstruction of digital rock CT images based on residual attention mechanism. *Advances in Geo-Energy Research*, *6*(2).

26. Wang, H., Wei, M., Cheng, R., Yu, Y., & Zhang, X. (2022). Residual deep attention mechanism and adaptive reconstruction network for single image super-resolution. *Applied Intelligence*, *52*(5), 5197-5211.
27. Zhou, L., Cai, H., Gu, J., Li, Z., Liu, Y., Chen, X., ... & Dong, C. (2022, October). Efficient image super-resolution using vast-receptive-field attention. In *European conference on computer vision* (pp. 256-272). Cham: Springer Nature Switzerland.
28. Lepcha, D. C., Goyal, B., Dogra, A., & Goyal, V. (2023). Image super-resolution: A comprehensive review, recent trends, challenges and applications. *Information Fusion*, *91*, 230-260.
29. Guerreiro, J., Tomás, P., Garcia, N., & Aidos, H. (2023). Super-resolution of magnetic resonance images using Generative Adversarial Networks. *Computerized Medical Imaging and Graphics*, 102280.
30. Gharibi, Z., & Faramarzi, S. (2023). Multi-frame spatio-temporal super-resolution. *Signal, Image and Video Processing*, *17*(8), 4415-4424.
31. Zhao, M., Yang, R., Hu, M., & Liu, B. (2024). Deep Learning-Based Technique for Remote Sensing Image Enhancement Using Multiscale Feature Fusion. Sensors, 24(2), 673. https://doi.org/10.3390/s24020673.
32. Chen, W., Chen, J., Wan, Y., Liu, X., Cai, M., Xu, J., Cui, H., & Duan, M. (2024). Land Cover Classification Based on Multimodal Remote Sensing Fusion. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-1-2024, 35-40. https://doi.org/10.5194/isprs-annals-X-1-2024-35-2024
33. Myasnikov, E. (2024). Fusion of Deep Learning-based and Spectral Features for Hyperspectral Image Analysis. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-2/W1-2024, 31-38. https://doi.org/10.5194/isprs-annals-X-2-W1-2024-31-2024
34. Wang, M., Sun, Y., Xiang, J., Sun, R., & Zhong, Y. (2024). Adaptive Learnable Spectral–Spatial Fusion Transformer for Hyperspectral Image Classification. Remote Sensing, 16(11), 1912. https://doi.org/10.3390/rs16111912.
35. Greza, M., Bhattacharya, I., Hoegner, L., & Jutzi, B. (2024). GAN-Based Dual Image Super Resolution for Satellite Imagery Decreasing Radiometric Uncertainty. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-3-2024, 155-2024.
36. Lia, Z., Li, J., Ren, L., & Chen, Z. (2024). Transformer-based Dual Path Cross Fusion for Pansharpening Remote Sensing Images. International Journal of Remote Sensing, 45(4), 1170-1200. https://doi.org/10.1080/01431161.2024.2306153.