Multi-Scale Fusion of Multi-Modal Astronomical Images for Enhanced Structural Analysis

Anamika Chowdhury[1], Monisha Kar[1], Snehangshu De[1], Sayan Roy Chowdhury[1], Sagnik Kar[1], Indranil Mukhopadhyay[1], Haimashree Ghosh[3], Hiya Chowdhury[3], Rupayan Das[2], Sutanu Ghosh[3], Subhabrata Banerjee[3]

Department of Computer Science and Engineering (IoT)[1], Institute of Engineering and Management, School of University of Engineering & Management, Kolkata, India

Department of Computer Science and Engineering[2], Institute of Engineering and Management, School of University of Engineering & Management, Kolkata, India

Department of Electronics and Communication Engineering[3], Institute of Engineering and Management, School of University of Engineering & Management, Kolkata, India

Abstract: Understanding celestial objects and astronomical data becomes easier when astronomical images from multiple modalities like optical images, infrared images, and radio images are combined through effective fusion. However, tradition deep learning methods faces limitations due to absence of large annotated datasets, which is scarce in astronomy and difference in modality characteristics of images. This study introduces a Hybrid Multi-Scale Weighted Laplacian Fusion (HMWLF) approach that functions without supervised training. By breaking images into a "pyramid" of layers, the system continuously captures fine-grained textures and broad structural shapes. The process uses a hybrid fusion strategy, which balances the weighted averaging for consistency with maximum selection to strengthen the most important features and prevent the loss of critical detail. By aligning everything up with the optical image, the system keeps the infrared and radio data clear and easy to understand. The experimental results of this framework confirm its efficiency and practical relevance in astronomical imaging with Structural Similarity Index (SSIM) of 0.742 and an average Peak Signal-to-Noise (PSNR) of 29.92 dB.

Keywords: Astronomical Imaging · Hybrid Multi-Scale Weighted Laplacian Fusion· Image Quality Assessment · Laplacian Pyramid Decomposition · Multi-modal Image Fusion · Multi-wavelength Astronomy · Structural Alignment ·

1 Introduction

Astronomical imaging is an important part of understanding different aspects of different celestial bodies. As different instruments are being improved for better observation, there is an increase in the collection of different kinds of observational data. This includes different kinds of observational data, such as optical, infrared, and radio. Each of these kinds of observational data contains different and unique kinds of information about different celestial bodies. However, it is a difficult task to combine all these different kinds of observational data into a single form.

The rest of the paper is organized as follows. Section 2 reviews literature work. Section 3 introduces system model. Section 4 presents the results and discussions. Section 5 concludes the paper.

2 Literature Survey

In recent times, different research works have been carried out to examine different methods for enhancing the quality of images and their analysis in astronomy. In the research work presented by the authors in [1], a deep learning model based on U-Net was proposed for denoising images in astronomy. The proposed method was able to enhance the image denoising process. In another research work presented by the authors in [6], an improved U-Net model was proposed for enhancing the performance of image denoising using U-Net.

A learning-based deconvolution technique was proposed in [2], and this technique was used for recovering details of high-resolution images from blurred astronomical images. This study revealed the potential of learning-based techniques for astronomical image restoration. The potential of computer vision techniques for astronomical image analysis was explored in [3], and this study was focused on feature detection and object identification. A traditional edge detection technique based on the Sobel operator was also explored [4].

Spatial regularization-based fusion was proposed by authors in [5], where they introduced an efficient framework for fusing multi-source astronomical images, maintaining spatial consistency. In [7], the authors proposed an improved TopHat (ITH) method for removing smear artifacts, thus enhancing the clarity of the image. The authors of [8] proposed a method for enhancing the stitching of astronomical images by utilizing a feature extraction method, i.e., XFeat, which is a part of deep learning.

Imaging technologies using radar were also discussed in [9], where the authors presented the ways for reconstructing the height dimensions of astronomical objects using earth-based radar technologies. In [10], the use of infrared optoelectronic sensing technologies was discussed, emphasizing their significance for astronomical observations.

The clustering-based methods were also studied in [11], where spectral clustering, K-means, and mean shift methods were used to efficiently cluster astronomical data. In [12], a modified convolutional neural network based on the MobileNet architecture was proposed for crescent moon detection. Therefore, the feasibility of using lightweight deep learning methods for astronomical data analysis was also shown.

The study in [13] aimed to introduce new advancements in the system for efficient astronomical image data analysis using a Docker framework for online processing.

More advanced methods for fusion and representation have also been discussed in some recent research. In [14], a cross-layer fusion network with adaptive alignment was proposed for better feature fusion, especially for multi-modal data. Similar to this idea of improving feature fusion for complex data, a spatially resolved embedding analysis was proposed in [15], especially for handling complex data like images, where a better understanding of spatial relationships in multi-dimensional data is required.

However, some problems remain to be solved in the field of multi-modal astronomical image fusion, even if some satisfactory results have been achieved by deep learning-based approaches, which are heavily data-dependent and may be prone to the dominance of certain modalities. The traditional approaches, although intuitive, are challenged by the difficulty of simultaneously maintaining the overall structures and details of images. Therefore, the problem of balanced fusion of complementary information among optical, infrared, and radio images without losing important information remains to be solved, and this paper proposes a novel Hybrid Multi-Scale Weighted Laplacian Fusion (HMWLF) framework for robust and high-quality fusion images.

3 System Model

The system introduces a Hybrid Multi-Scale Weighted Fusion (HMWLF) to designed to seamlessly blend data from optical, infrared, and radio imaging. To test the approach, we used the COSMOS multi-wavelength dataset. This method is designed into five-stage process for clarity and accurate fusion. The five-stage process consists of (1) Data Processing, (2) Multi-Scale Decomposition, (3) Hybrid Fusion Strategy, (4) Reconstruction and Post-Processing, (5) Evaluation

Metrics. The detailed flow diagram of our introduced work is explained in Fig. 1.

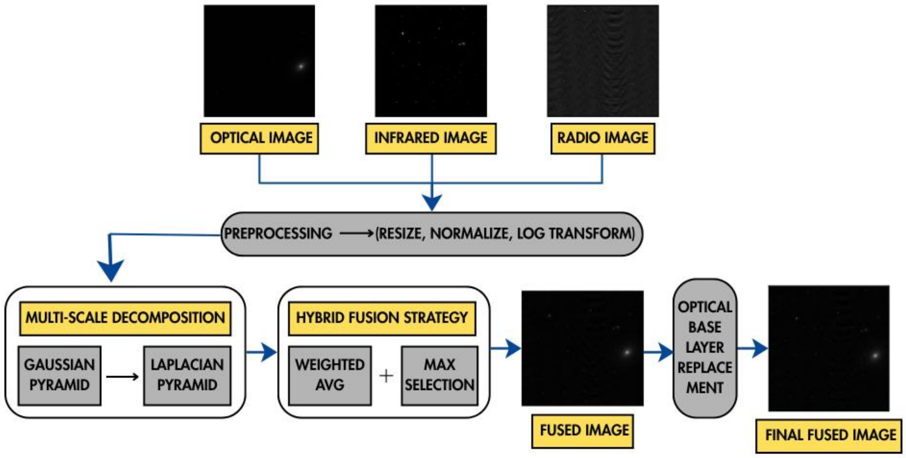


Fig.1: Architectural Framework of the System

3.1 Data Preprocessing

Each image has different resolutions, intensity distribution, and noise characteristics depending on the modality of the images, which makes it difficult to fuse the images directly. Preprocessing of the images is necessary for fusion, which involves several steps. First, there is a need to remove invalid pixel values, which keeps the data clean. Then, a logarithmic transformation is applied, which enhances the visibility of faint objects in the image. Then, all the images are normalized to a fixed size, and after that, they are normalized to a fixed range for efficient fusion.

3.2 Multi-Scale Decomposition

A multi-scale approach is constructed with the help of Gaussian and Laplacian pyramids to effectively obtain the large, coarse structures as well as the fine details from the given astronomical images. First, a Gaussian pyramid is constructed, where each input image is repeatedly down-sampled to obtain smoother versions of the image at different resolutions. This helps in extracting the large, coarse structures from the given images. Next, a Laplacian pyramid is constructed by repeatedly taking the difference between the higher resolution images of the Gaussian pyramid, which are up-sampled, and the corresponding images in the lower resolution. This gives a frequency decomposition of the image, where the image decomposes into a certain number of frequency bands. Here, the high-frequency bands represent the detailed features of the image, such as the presence of stars, and the low-frequency bands represent the overall illumination and large-scale variations in the background.

3.3 Hybrid Fusion Strategy

The fusion done over various levels of the Laplacian pyramid. This is done by using a hybrid method for effective fusion of information from different modal images. It uses a weighted average approach, where the weights are determined by the intensity values of the image so that stronger intensity values have a greater influence on fused image. In addition, maximum selection method is also used so that the important features are not compromised during the fusion process. Further, the lowest-frequency layer of the fused pyramid is replaced with the optical image layer to maintain overall structural consistency of the final fused image.

3.4 Reconstruction and Post-Processing

The final fused image is formed by reconstructing the Laplacian pyramid. This is done by iteratively up-sampling and merging the fused pyramid levels, which enables the image to be reconstructed to its original resolution with the features obtained in the multi-scale decomposition process. For further refinement a post-processing technique is used where the base layer is replaced with the optical layer while keeping the features from infrared and radio images.

3.5 Evaluation Metrics

To estimate the performance of the model structural similarity index (SSIM) and peak signal-to-noise ratio (PSNR), are calculated. As the ground truth for the fused image is not available, PSNR is computed separately for each of the input images from different modalities – optical, infrared, and radio images, then their average is taken as the final PSNR for a balanced evaluation. Additionally, qualitative analysis is carried out to visually examine how well the method is preserving fine details and enhancing faint features while minimizing the noises.

4 Results and Discussion

4.1 Dataset Description

The COSMOS dataset, a multi-wavelength astronomical dataset used in this study contains astronomical images from different spectral domains – optical, infrared, and radio spectra. Each modality corresponds to the same sky region but captures different insights due to wavelength sensitivity differences. The COSMOS dataset contains optical images captured using Advanced Camera for Surveys (ACS) provides detailed structural information. On the other hand, the infrared images in the COSMOS dataset captured using Infrared Array Camera (IRAC) reveals thermal characteristics which are not visible in the optical range, while the radio images from the Very Large Array (VLA) shows large-scale emission patterns. As the images from different modalities varies in resolution and dimension, all images are standardized through cropping and resized to 256x256 pixels after preprocessing.

4.2 Performance Evaluation

The performance of the models is evaluated using SSIM (Structural Similarity Index) and PSNR (Peak Signal-to-Noise Ratio), which measures the structural similarity and reconstruction quality, respectively. The results of WAF (Weighted Average Fusion) indicates poor preservation of structural details and has low potential to reduce noise. The performance of LPF (Laplacian Pyramid Fusion) significantly improves due to multi-scale approach but has limitations in adaptive feature fusion. The results obtained from MSAFNet (Multi-Scale Attention Fusion) which is a deep-learning based approach shows highest SSIM value, indicating strong structural preservation but it tends to prioritize the optical modality, thus producing a less balanced fusion. In comparison to these models, the HMWLF (Hybrid Multi-Scale Weighted Fusion) with an SSIM of 0.742 and PSNR of 29.92 dB demonstrates a balanced performance, effectively preserving features from all modalities. The quantitative comparison of different fusion models using SSIM and PSNR is presented in Table 1.

Table 1: Quantitative Comparison Of Image Fusion Methods

Model	**SSIM**	**PSNR(dB)**
WAF (Weighted Average Fusion)	0.119	19.97
LPF (Laplacian Pyramid Fusion)	0.625	31.76
MSAFNet (Multi-Scale Attention Fusion)	0.830	29.90
HMWLF (Hybrid Multi-Scale Weighted Laplacian Fusion)	0.742	29.92

4.3 Qualitative Visual Analysis

The qualitative results presented in Fig. 2 and Fig. 3 highlight the effectiveness of the proposed HMWLF framework in comparison to existing methods.

Each modality provides information that is complementary to the others. Optical images show crisp structural details, infrared images show faint and thermal features, and radio images show large-scale emission patterns that cannot be seen in other modalities.

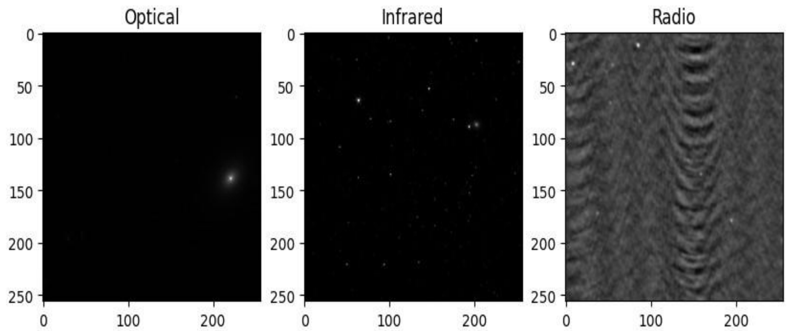
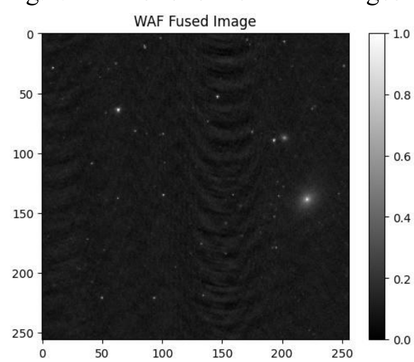
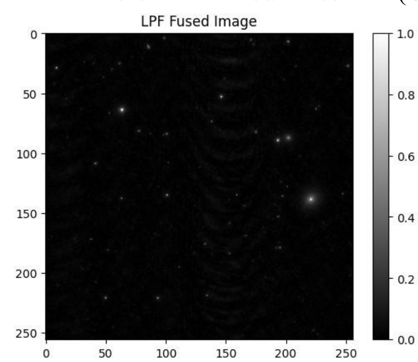


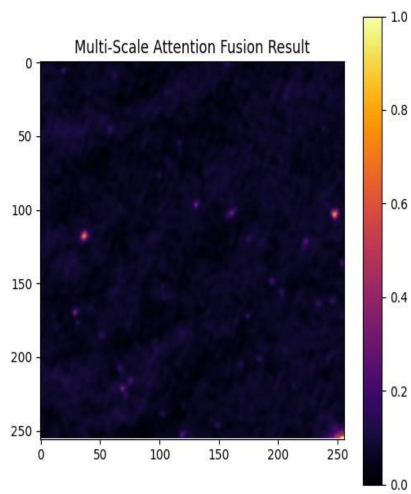
Fig.2: Astronomical Images From Different Modalities (Optical, Infrared, Radio)



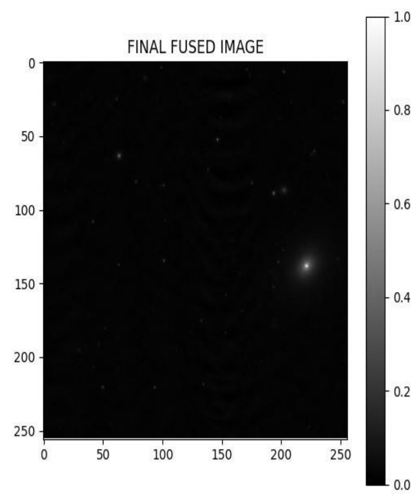
(a) WAF (Weighted Average Fusion)



(b) LPF (Laplacian Pyramid Fusion)



(c) MSAFNet (Multi-Scale Attention Fusion)



(d) HMWLF (Hybrid Multi-Scale Weighted Laplacian Fusion)

Fig.3: Comparison of different image fusion methods

As observed in Fig. 3, the WAF approach provides a blurred image due to its averaging process, which results in the loss of important structural details. The LPF approach provides better preservation of features by using multi-scale decomposition. However, it does not provide a well-balanced fused image with respect to different modalities. The MSAFNet approach provides better preservation of structures. However, it is more biased towards optical images and hence suppresses important information from infrared and radio images.

The proposed HMWLF approach provides a well-balanced fused image with respect to different modalities. Moreover, it provides better preservation of both fine structures and faint features simultaneously. The ability to differentiate between different characteristics of optical, infrared, and radio images in a single image provides evidence that our proposed approach is able to integrate different information without any bias towards any modality.

5 Conclusion

In this paper, a novel method for the fusion of multi-modal images in the field of astronomy is proposed, which is referred to as HMWLF. This method is efficient for handling the problems faced by the existing techniques, which need a large dataset for training a model. This method provides a strong fusion of feature for different modal images. Experimental results for the proposed method are performed, which prove that the proposed method provides a good trade-off between the structure, noise reduction, and enhancement of feature for different modal images. This method is efficient for handling the problems faced by the existing techniques, which are influenced by artifacts as well as the techniques of deep learning that are influenced by modality bias. This method provides visually appealing results for fusion images. The capability of the proposed framework for maintaining small and significant features in the field of astronomy makes this method highly efficient.

References

1. A. Vojtekova, M. Lieu, I. Valtchanov, B. Altieri, L. Old, Q. Chen, and F. Hroch,“Learning to denoise astronomical images with U-nets,” Monthly Notices of the Royal Astronomical Society, vol. 503, no. 3, pp. 3204–3215, Oct. 2021, doi: 10.1093/mnras/staa3567.
2. M. Long, Y. Soubo, S. Cong, N. Weiping, and L. Tong, “Learning deconvolutionsfor astronomical images,” Monthly Notices of the Royal Astronomical Society, vol. 504, no. 1, pp. 1077–1083, Mar. 2021, doi: 10.1093/mnras/stab956.
3. S. Karypidou, I. Georgousis, and G. A. Papakostas, “Computer vision for astronomical image analysis,” in Proc. 2021 IEEE Int. Conf. Progress in Informatics and Computing (PIC), Dec. 2021, pp. 94–101, doi: 10.1109/PIC53636.2021.9687023.
4. S. Khlamov, I. Tabakova, and T. Trunova, “Recognition of astronomical images using the Sobel filter,” in Proc. 2022 29th Int. Conf. Systems, Signals and Image Processing (IWSSIP), Jun. 2022, pp. 1–4, doi: 10.1109/IWSSIP55020.2022.9854425.
5. C. Guilloteau, T. Oberlin, O. Bern´e, and N. Dobigeon, “Informed spatial regularizations for fast fusion of astronomical images,” in Proc. 2022
IEEE Int. Conf. Image Processing (ICIP), Oct. 2022, pp. 2221–2225, doi: 10.1109/ICIP46576.2022.9897181.
6. J. Qi, M. Chen, Z. Wu, C. Su, X. Yao, and C. Mei, “An improved U-Net model forastronomical image denoising,” in Proc. 2022 China Automation Congress (CAC), Nov. 2022, pp. 1901–1905, doi: 10.1109/CAC57257.2022.10054887.
7. Z. Shen, H. Wang, S. Chen, T. Zhang, X. Zhang, and X. Yang, “ITH: ImprovedTopHat for smear suppression in astronomical images,” in Proc. 2024 Int. Conf. Image Processing, Computer Vision and Machine Learning (ICICML), Nov. 2024, pp. 276–279, doi: 10.1109/ICICML63543.2024.10958138.
8. T. Wang, X. Lin, and P. Zhang, “Enhancing astronomical image stitching withXFeat: A deep learning approach,” in Proc. 2024 Int. Conf. Image Processing, Computer Vision and Machine Learning (ICICML), Nov. 2024, pp. 708–712, doi: 10.1109/ICICML63543.2024.10958013.
9. K. Zhu, P. Liu, Z. Zhou, Z. Wang, Z. Dong, and Z. Ding, “Earth-based radar heightdimension imaging for astronomical object,” in *Proc. IEEE Int. Conf. Signal, Information and Data Processing (ICSIDP)*, Nov. 2024, pp. 1–5, doi: 10.1109/ICSIDP62679.2024.10868262.
10. W. Hu and D. Jariwala, “Infrared optoelectronic sensing technology,” IEEEJournal of the Electron Devices Society, vol. 13, pp. 1052–1059, 2025, doi: 10.1109/JEDS.2025.3617536.
11. V. S. Mohite, K. G. Patra, C. H. Sankar, and S. Jayan, “Astronomical imageclustering using spectral, K-means and mean shift,” in *Proc. Int. Conf. Computing for Sustainability and Intelligent Future (COMP-SIF)*, Mar. 2025, pp. 1–7, doi: 10.1109/COMP-SIF65618.2025.10969906.
12. M. H. Yuhandri, Yuhandri, and Sumijan, “Modified convolutional neural networks(CNN) with MobileNet-v2-186 architecture in crescent moon image detection,” in *Proc. Int. Conf. Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, Dec. 2025, pp. 32–37, doi: 10.1109/ICIMCIS68501.2025.11326934.
13. E. Hadzhyiev, S. Khlamov, A. Frolov, I. Tabakova, S. Lytvynenko, and M. Shypova, “Application of Docker Compose for constructing the infocommunication system for online processing of astronomical images,” in *Proc. 15th Int. Conf. Advanced Computer Information Technologies (ACIT)*, Sept. 2025, pp. 629–633, doi: 10.1109/ACIT65614.2025.11185586.
14. P. Peng, T. Xu, L. Song, M. Zhu, Y. Fang, and J. Li, “COXNet: Cross-layer fusionwith adaptive alignment and scale integration for RGBT tiny object detection,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 36, no. 1, pp. 596–608, Jan. 2026, doi: 10.1109/TCSVT.2025.3595147.
15. S. Warchol, G. Guo, J. Knitte, D. Freeman, U. Shalla, J. L. Muhlich, P. K. Sorger,and H. Pfister, “SEAL: Spatially-resolved embedding analysis with linked imaging data,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 32, no. 1, pp. 484–494, Jan. 2026, doi: 10.1109/TVCG.2025.3634794.