

Two-scale decomposition and deep learning fusion for visible and infrared images

Ruhan Bevi Azad¹, Hari Unnikrishnan², Lokesh Gopinath¹

¹Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, India

²Department of Electronics and Communication Engineering, Saveetha Engineering College, Chennai, India

Article Info

Article history:

Received Apr 22, 2024

Revised Nov 26, 2024

Accepted Dec 2, 2024

Keywords:

Algorithm unravelling

Deep learning

Near-infrared image

Traditional method

Two-scale decomposition

Visible image

ABSTRACT

The paper focuses on the fusion of visible and infrared images to generate composite images that preserve both the thermal radiation information from the infrared spectrum and the detailed texture from the visible spectrum. The proposed approach combines traditional methods, such as two-scale decomposition, with deep learning techniques, specifically employing an autoencoder architecture. The source images are subjected to two-scale decomposition, which extracts high-frequency detail and low-frequency base information. Additionally, an algorithmic unravelling technique establishes a logical connection between deep neural networks and traditional signal processing algorithms. The model consists of two encoders for decomposition and a decoder after the unravelling operation. During testing, a fusion layer merges the decomposed feature maps, and the decoder generates the fused image. Evaluation metrics including entropy, average gradient, spatial frequency and standard deviation are employed to subjectively assess fusion quality. The proposed approach demonstrates promise for effectively combining visible and infrared imagery for various applications.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Ruhan Bevi Azad

Department of Electronics and Communication Engineering, SRM Institute of Science and Technology
Kattankulathur, Chengalpattu, 603203, Tamil Nadu, India

Email: ruhanb@srmist.edu.in

1. INTRODUCTION

In the field of image processing research, image fusion is an emerging topic. Adopting similar methodologies and strategies, enhances the effectiveness, interpretability, and reproducibility of image fusion. Additionally, evaluating models on benchmark datasets and making code and data openly available will contribute to the advancement of research in the field of remote sensing [1]. Incorporating insights from the base paper involves leveraging unsupervised learning techniques and advanced loss functions for image fusion. Enhancing model interpretability and evaluating its performance on benchmark datasets are crucial steps. Additionally, sharing code and data ensures transparency and reproducibility, facilitating further advancements in the field [2]. The refinement fusion approach achieves superior performance in terms of image quality, target region preservation and efficiency [3]. Evaluating the performance and validating it through experiments, enhances the effectiveness and robustness of image fusion. considering the similarities and refining approach and addressing specific challenges in infrared and visible image fusion [4]. There are potential areas for improvement and innovation in methodologies. Leveraging insights from each approach can contribute to the development of more effective and advanced image fusion techniques [5]–[8]. The concept of a unified fusion framework, adaptive information preservation, mitigation of deep learning limitations, and the use of benchmark datasets, enhances the effectiveness, versatility, and evaluation of

image fusion. Additionally, considering the similarities and differences between approaches can provide valuable perspectives for refining fusion models and addressing specific challenges in infrared and visible image fusion [9], [10]. Sparse regularization [11], [12], dictionary learning [13], [14], alternating optimization [15], and performance evaluation against state-of-the-art methods, enhance the effectiveness and efficiency of infrared and visible image fusion. Leveraging techniques from diverse domains, such as hyperspectral and multispectral image fusion, can lead to innovative solutions and improvements in fusion algorithms [16].

A multiscale fusion strategy, and encoder-fusion strategy-decoder framework, enhance the performance and robustness of infrared and visible image fusion [17]. Comparative analysis of fusion techniques, evaluation metrics, efficiency, and suitability for real applications, and the effectiveness and applicability of image fusion research, even if they operate in a different domain than medical imaging, can guide the evaluation process and ensure the reliability of the fusion approach [18]. A symmetric encoder-decoder architecture with residual blocks, attention mechanisms, and separation of training and fusion stages, enhances the performance and efficiency of image fusion [19].

The use of generative adversarial network (GAN) with multiclassification constraints, content loss mechanisms, and comprehensive evaluation methodologies, enhances the performance and effectiveness of image fusion [20]. Adaptive enhancement techniques, hybrid decomposition models, coupled dictionary-based fusion, and novel fusion schemes [21]. The use of residual network architectures, innovative loss functions, and two-stage training strategies, enhances the performance and efficiency of image fusion. The emphasis on task-specific fusion strategies and the adoption of residual network architectures, can guide the refinement of the fusion approach to better adapt to diverse fusion tasks [22]. Utilizing unsupervised end-to-end network architectures, designing tailored loss functions, and implementing convolutional layer decomposition networks, can enhance performance and effectiveness [23]. The comprehensiveness and effectiveness of these methods should be enhanced. This includes leveraging historical context, exploring various fusion techniques, incorporating rigorous evaluation processes, and considering practical applications and future prospects [24]. Figure 1 shows the proposed fusion model which consists of four stages:

- Input sources: Infrared images and visible images are fed into the model
- Decomposition: The images are decomposed into constituent features
- Feature fusion: Extracted features are fused to enhance information
- Image fusion: The fused features are fused together to generate the final image

The remaining sections of this work are structured as follows. Section 2 provides a comprehensive explanation of the methodology that has been devised to integrate near-infrared and visible images and the framework is discussed in section 3. A quantitative and qualitative evaluation of the algorithm is provided in section 4, and it is concluded in section 5.

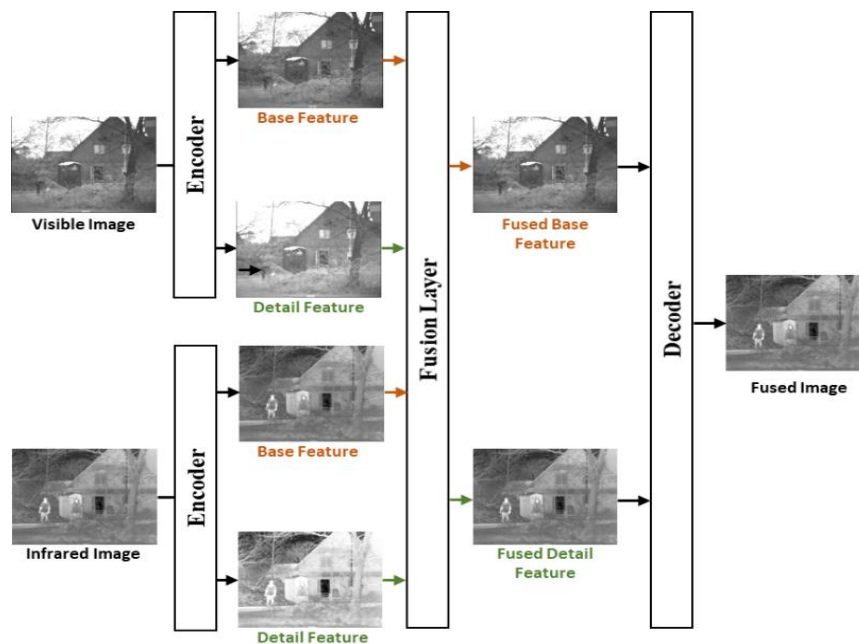


Figure 1. Proposed fusion model

2. PROPOSED METHOD

2.1. Two-scale decomposition method

Throughout the years, numerous fusion algorithms have been proposed, each aiming to enhance the quality and level of information in images. The primary objective of two-scale decomposition is to partition the original image into a series or ensemble of images, each of which highlights a specific attribute or characteristic. Instead of directly merging the source images, this decomposition method is employed. Decomposition prioritizes the fusion of the deconstructed images rather than making direct modifications to the source images. A more intricate and comprehensive fused image is generated by merging the unique information conserved in each deconstructed image.

This paper explores two distinct deep learning methodologies within the framework of multiscale decomposition for fusion. Although the first method is useful for extracting general features, it relies on a pretrained deep neural network that may not be optimized because of the intricate nature of multiscale decomposition tasks. This could result in suboptimal fusion outcomes due to the network's limited adaptability in various breakdown scenarios. On the other hand, the second method utilizes an autoencoder. Autoencoders are particularly well-suited for tasks that require accurate preservation of information, as they are specifically designed for extracting features and reconstructing images. An autoencoder's encoding process decomposes images into feature maps, enabling the efficient capture of information at different scales. An advantage of using an autoencoder-based approach, as opposed to a pretrained deep neural network, is its superior suitability for multiscale decomposition tasks due to its inherent adaptability and contextual relevance.

2.2. Optimization model

The proposed model utilizes optimization techniques to decompose an input image into a base image and a detail image. The acquisition of the basis image can be achieved by addressing the issue of low-frequency background information.

$$B_i^* = \arg \min \frac{\theta_{B_i}}{2} \|X_i - B_i\|_F^2 + \sum_{m=1}^n \|g_m^{B_i} * B_i\|_F^2 \quad (1)$$

where B_i^* is the disintegrated base image, $g_m^{B_i}$ ($m=1,2,\dots, n$) are high pass filters, X_i is the input image, $*$ is the convolution operation, θ_{B_i} is represents the tuning hyperparameter and

$$\sum_{m=1}^n \|g_m^{B_i} * B_i\|_F^2$$

is used to reduce the high frequency of B_i . Now, D_i represents the detailed image and means high frequency quality/texture and colour progression and is given by (2).

$$D_i^* = \arg \min \frac{\theta_{D_i}}{2} \|X_i - D_i\|_F^2 + \sum_{m=1}^n \|g_m^{D_i} * D_i\|_F^2 \quad (2)$$

here, $g_m^{D_i}$ ($m=1,2,\dots, n$) are low pass filters, and θ_{D_i} is again a tuning hyperparameter. To account for their flexibility, the weights for a central pixel must be normalized to sum to 1. Therefore, the base feature map is

$$B_i^{out} = B_i^{in} - \eta B_i \sum_{m=1}^n (g_m^{B_i})^T * (g_m^{B_i} * B_i^{in}) - \theta B_i (X - B_i^{in}) \quad (3)$$

where $(g_m^{B_i})^T$ represents the kernel of $g_m^{B_i}$ rotated by 180° and ηB_i is the step size.

2.3. Unravelling algorithm

The ‘‘unravelling’’ algorithm is a recently developed method that provides an appealing process for designing deep neural networks based on models. Algorithm unrolling is the process of transforming a repetitive algorithm into a deep neural network (DNN) by expanding its computational graph. This enables the training of predefined hyperparameters and unknown coefficients in a comprehensive way. The base convolution layer and detail convolution layer are replaced with filters $g_m^{B_i}$ and $g_m^{D_i}$ as (4).

$$B_i^{out} = B_i^{in} - \eta B_i [Conv_2^{B_i}(Conv_1^{B_i}(B_i^{in})) - \theta B_i (X - B_i^{in})] \quad (4)$$

where $Conv_m^{B_i}$ ($m=1,2$) which denotes the kernel size. Furthermore, we set the $Conv_1^{B_i}$ kernel equal to $Conv_2^{B_i}$ where this made a 180° turn. Similarly, the detailed feature map updating procedure is carried out as (5):

$$D_i^{out} = D_i^{in} - \eta D_i [Conv_2^{D_i}(Conv_1^{D_i}(D_i^{in})) - \theta D_i (X - D_i^{in})] \tag{5}$$

where, θB_i and θD_i are predefined hyperparameters and ηB_i and ηD_i are step sizes.

3. FRAMEWORK

The detail and base images are regarded as feature maps obtained from the source image. This approach involves organizing N detail convolution layer (DCL) and base convolution layer (BCL) as two encoders. The objective is to reproduce the repetitive procedure of conventional optimization models and extract fundamental and intricate feature maps. Subsequently, a supplementary decoder is generated using inputs that include the summation of two deconstructed feature maps. The result of this decoder is the reconstructed source image.

Figure 2 illustrates the network architecture during the training phase and Figure 3 depicts a single BCL, while a DCL has a similar structure but distinct parameters. The number of input and output channels for the first convolution units $Conv_1^{B_i}$ and $Conv_1^{D_i}$ is (1, H). The second convolutional units, $Conv_2^{B_i}$ and $Conv_2^{D_i}$ are set as (H, 1). The value of H is set to 64. DCL and BCL do not have any parameters that are shared between them. The Laplacian and blur filters are applied to the source image and the detail encoder D_i^0 and base encoder B_i^0 are initialized. The sigmoid function, a batch regularization layer, and a 3×3 convolution unit make up the decoder. The convolution unit has one input channel and one output channel. The restored image’s pixel values are normalized to a range of 0-1 by means of the sigmoid function.

Figure 4 illustrates the workflow of the testing framework. During the test phase, input pairs of infrared and visible images are given and the ultimate fusion results are subsequently obtained. Upon completion of the training process, two proficient encoders and a decoder are obtained. In this context, D_i^N , B_i^N , D_i^N , and B_i^N represent the infrared detail, base feature maps and visible detail, base feature maps, respectively.

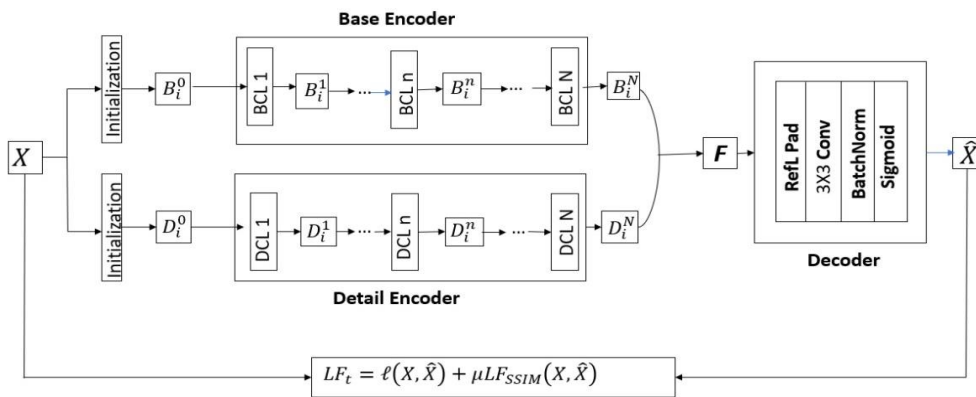


Figure 2. Training framework

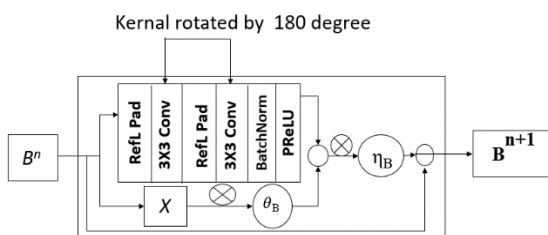


Figure 3. Single BCL layer

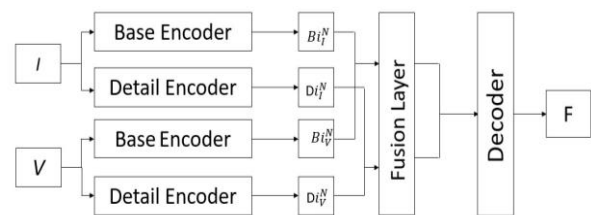


Figure 4. Training framework

4. EXPERIMENTS

The effectiveness of this strategy was evaluated through a comprehensive comparative analysis of the standard fusion algorithm and recently proposed fusion algorithm. Subjective assessments focused on visual quality and perceptual clarity, while objective assessments utilized established metrics, such as entropy and spatial fidelity. The results consistently demonstrate the superior performance of the proposed method, highlighting its potential for practical applications requiring high image quality.

4.1. Datasets

The FLIR and TNO datasets [25] were selected as the test subjects to thoroughly assess the efficacy of the proposed method. The TNO dataset consists of numerous prealigned pairs of near-infrared (NIR) and visible images. Similarly, the FLIR dataset consists of thermal infrared and visible images. The experimental setup involved selecting thirty pairs of images from the FLIR dataset. These images were then uniformly resized to 256×256 pixels to facilitate the analysis of multiscale transformations. A laptop featuring a Core i7 processor and 16 GB of RAM was utilized for conducting the evaluation process.

4.2. Quality metrics

Four objective assessment metrics were used to analyze the impact of the proposed method. They are entropy (EN) is the evaluation of image quality to quantify the level of information contained within the fused image. The average gradient (AG) metric assesses an image's visual clarity by examining its textural and contrast features. This evaluation determines how well the combined image preserves the intricate details and boundaries found in the original images. Images with higher AG scores are generally considered to have better perceptual quality. Spectral fidelity (SF) is a metric that quantifies the extent to which the spectral information of the input images is accurately maintained in the fused image, thus ensuring the fidelity of the fused image to the original spectral characteristics. Spatial distortion (SD) is a measure of the extent to which spatial distortion or misalignment occurs during the fusion process. It evaluates the degree to which those spatial details are maintained in the fused image.

4.3. Evaluation against rival algorithms using the FLIR and TNO datasets

The effectiveness of the proposed approach is substantiated through the application of the FLIR dataset. The fusion outcomes of the source image are presented in Figure 5, which depicts a scene featuring two people standing alongside their bicycle near the edge of an apartment complex road. The visible image clearly displays the specific features of the house and cycle. However, the individuals were not discernible in the visible image, whereas they were detectable in the NIR image because of their high sensitivity to thermal radiation, which captures data related to a person. Quantitative and qualitative metrics were used to assess the fusion performance of the proposed method.

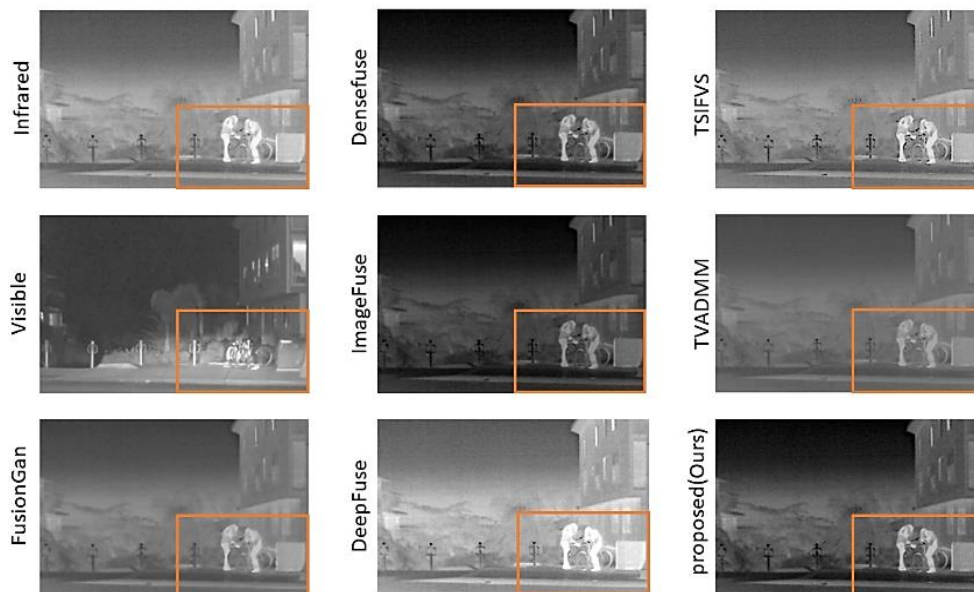


Figure 5. Fusion output

From a subjective standpoint, the image produced by FusionGAN appears to be of lower visual quality and lacks several elements found in the visible image. Images generated using DeepFuse generally appear blurry, with indistinct edges and considerable noise. This loss of edge information makes it difficult to discern the interior of the apartment. The DenseFuse algorithm fails to maintain consistent brightness levels, particularly noticeable on road surfaces. The output from total variation alternating direction method of multipliers (TVADMM) displays a degree of fogginess, while two-scale infrared and visible image fusion scheme (TSIFVS) results seem underexposed. In comparison to all other methods, the proposed approach demonstrates superior contrast in identifying prominent targets. The proposed methodology clearly produces a fused image that successfully displays the rich textures of two people near their cycles in the designated orange box.

As indicated in Table 1, the proposed method outperforms existing techniques in EN, AG, and SD, highlighting its superior performance in maintaining detail and clarity in the fused images. Specifically, the EN value of 7.45, AG value of 5.91, and SD value of 38.18 are the highest among all the compared methods, demonstrating the method's ability to preserve fine textures, edges, and spatial consistency, which are crucial for tasks like surveillance. The visual clarity of the proposed method, as shown in Figure 6, is markedly higher than that of alternative techniques, such as DeepFuse and FusionGAN, which show compromises in spatial consistency and edge detail.

Table 1. Average results obtained by applying multiple techniques to the FLIR dataset

Methods	EN	AG	SF	SD
DeepFuse	7.21	4.80	15.47	37.35
FusionGan	7.02	3.20	11.51	34.38
DenseFuse	7.21	4.82	15.50	37.32
TSIFVS	7.15	5.57	18.79	35.89
ImageFuse	6.99	4.15	14.52	32.58
TV-admm	6.80	3.52	14.04	28.07
Proposed	7.45	5.91	14.46	38.18

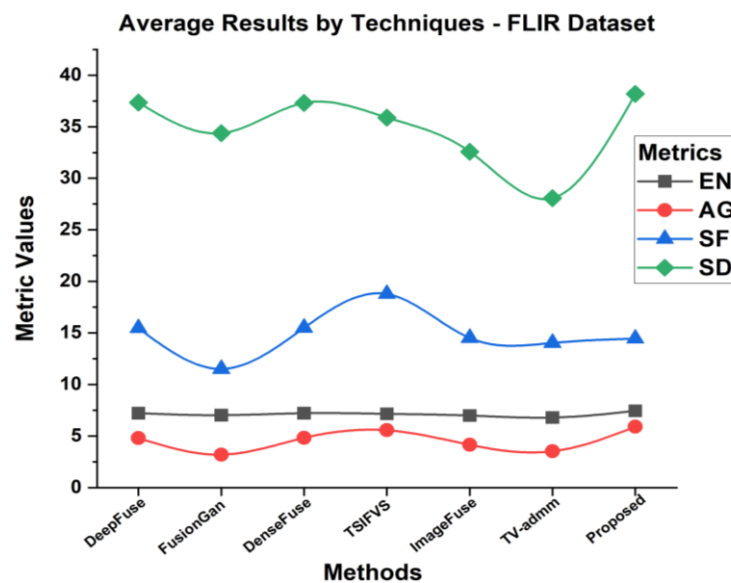


Figure 6. Average results by techniques - FLIR dataset

Using a TNO dataset, Table 2 displays the average values of six objective assessment criteria, with the red values emphasizing the highest values. Compared to DeepFuse and FusionGAN, our method consistently delivers better performance in terms of AG and EN. These findings align with previous studies, where multiscale decomposition techniques demonstrated an advantage in spatial clarity. However, as seen with SF, future improvements may focus on spectral fidelity, potentially through hybrid techniques that merge decomposition-based and spectral retention-focused methods. The proposed algorithm outperformed all alternatives in terms of fusion performance, with the exception of AG and SF, and achieved the maximum achievable score.

Table 2. Average results obtained by applying multiple techniques to the TNO dataset

Methods	EN	AG	SF	SD
DeepFuse	6.86	3.60	11.13	32.25
FusionGan	6.58	2.42	8.76	29.04
DenseFuse	6.84	3.60	11.09	31.82
TSIFVS	6.67	3.98	12.60	28.04
ImageFuse	6.38	2.72	9.80	22.94
TV-admm	6.40	2.52	9.03	23.01
Proposed	6.90	3.33	9.86	33.50

Figure 7 illustrates the comparison of average results for various techniques based on the TNO dataset, further highlighting the superior performance of the proposed method in most metrics. The results from both the FLIR and TNO datasets underscore the potential of the proposed method for applications requiring high spatial clarity, such as in surveillance and medical imaging. Future work could address the slight trade-off in spectral fidelity by exploring hybrid approaches that incorporate additional spectral loss functions, enabling a more balanced fusion across different modalities. Expanding the scope of this method to other data types, such as hyperspectral images, could reveal new applications and further enhance the robustness of the proposed fusion strategy.

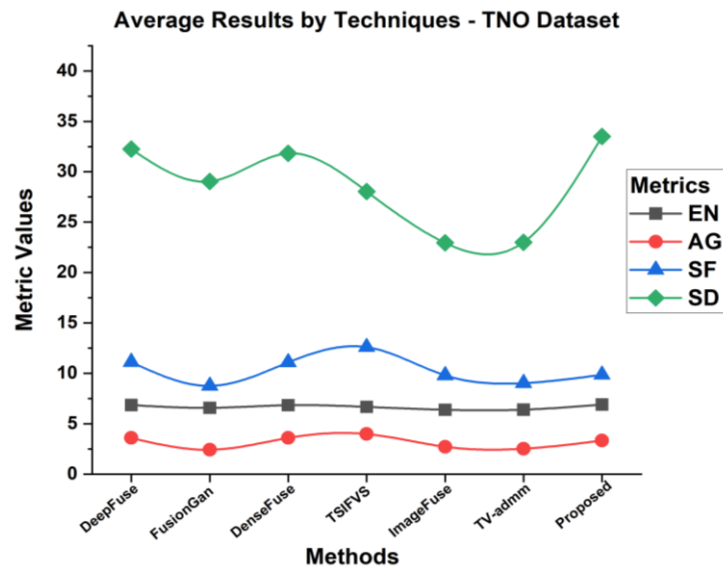


Figure 7. Average results by techniques - TNO dataset

5. CONCLUSION

The proposed method for visible and infrared image fusion, which combines two-scale decomposition with deep learning techniques, has demonstrated significant improvements in performance over existing approaches. Our approach effectively preserves both the spatial and thermal information from the input images, as reflected by the highest AG, EN, and SD values among all compared methods. These results indicate that our fusion model is particularly well-suited for applications where spatial detail and clarity are critical, such as surveillance, military, and medical imaging.

Beyond demonstrating its efficacy in image fusion, this research also opens the door for future investigations. While the method excels in maintaining texture and clarity, further work is required to enhance spectral fidelity. Exploring hybrid fusion techniques that incorporate spectral retention mechanisms alongside multiscale decomposition could address this limitation. Additionally, expanding the method to handle other multimodal image fusion tasks, such as hyperspectral or radar imagery, presents an exciting direction for future research.

In conclusion, our findings contribute to the ongoing advancements in image fusion techniques, offering a robust solution for combining visible and infrared data. The real-world implications of this method are vast, and its application across different domains can potentially lead to enhanced detection, monitoring, and imaging systems. Moving forward, improvements in spectral fidelity and extension to broader datasets will further solidify the utility of this approach within the image fusion research community.




ACKNOWLEDGEMENTS

The author expresses sincere gratitude to carry this extended research based on the financial support received from the SRM Institute of Science and Technology, Kattankulathur, India, under the Selective Excellence Research Initiative (SERI).




REFERENCES

- [1] H. V. Nguyen, M. O. Ulfarsson, J. R. Sveinsson, and M. D. Mura, "Unsupervised sentinel-2 image fusion using a deep unrolling method," *IEEE Geoscience and Remote Sensing Letters*, vol. 20, pp. 1–5, 2023, doi: 10.1109/LGRS.2023.3326845.
- [2] J. M. Ramirez, J. I. Martínez-Torre, and H. Arguello, "LADMM-Net: an unrolled deep network for spectral image fusion from compressive data," *Signal Processing*, vol. 189, Dec. 2021, doi: 10.1016/j.sigpro.2021.108239.
- [3] S. Zhang, X. Li, X. Zhang, and S. Zhang, "Infrared and visible image fusion based on saliency detection and two-scale transform decomposition," *Infrared Physics & Technology*, vol. 114, May 2021, doi: 10.1016/j.infrared.2020.103626.
- [4] L. Ma, Y. Hu, B. Zhang, J. Li, Z. Chen, and W. Sun, "A new multi-focus image fusion method based on multi-classification focus learning and multi-scale decomposition," *Applied Intelligence*, vol. 53, no. 2, pp. 1452–1468, Jan. 2023, doi: 10.1007/s10489-022-03658-2.
- [5] L. Gopinath and A. R. Bevi, "Anisotropic guided filtering and multi-level disintegration method for NIR and visible image fusion," in *International Conference on Cognitive Computing and Information Processing*, 2024, pp. 165–175.
- [6] K. Li, G. Liu, X. Gu, H. Tang, J. Xiong, and Y. Qian, "DANT-GAN: a dual attention-based of nested training network for infrared and visible image fusion," *Digital Signal Processing*, vol. 145, Feb. 2024, doi: 10.1016/j.dsp.2023.104316.
- [7] L. Gopinath and A. Ruhan Bevi, "A dimensionality reduction method for the fusion of NIR and visible image," in *International Conference on Image Processing and Capsule Networks*, 2023, pp. 629–645.
- [8] P. Zhu, W. Ouyang, Y. Guo, and X. Zhou, "A two-to-one deep learning general framework for image fusion," *Frontiers in Bioengineering and Biotechnology*, vol. 10, Jul. 2022, doi: 10.3389/fbioe.2022.923364.
- [9] H. Xu, J. Ma, J. Jiang, X. Guo, and H. Ling, "U2Fusion: a unified unsupervised image fusion network," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 1, pp. 502–518, 2022, doi: 10.1109/TPAMI.2020.3012548.
- [10] H. Unnikrishnan and R. B. Azad, "Nonlocal retinex based dehazing and low light enhancement of images," *Traitement du Signal*, vol. 39, no. 3, pp. 879–892, 2022, doi: 10.18280/ts.390313.
- [11] N. Anantirasirichai, R. Zheng, I. Selesnick, and A. Achim, "Image fusion via sparse regularization with non-convex penalties," *Pattern Recognition Letters*, vol. 131, pp. 355–360, 2020, doi: 10.1016/j.patrec.2020.01.020.
- [12] Y. Pan, T. Lan, C. Xu, C. Zhang, and Z. Feng, "Recent advances via convolutional sparse representation model for pixel-level image fusion," *Multimedia Tools and Applications*, vol. 83, no. 17, pp. 52899–52930, 2024, doi: 10.1007/s11042-023-17584-z.
- [13] Q. Hu, S. Hu, X. Ma, F. Zhang, and J. Fang, "MRI image fusion based on optimized dictionary learning and binary map refining in gradient domain," *Multimedia Tools and Applications*, vol. 82, no. 2, pp. 2539–2561, 2023, doi: 10.1007/s11042-022-12225-3.
- [14] Y. Jie, X. Li, H. Tan, F. Zhou, and G. Wang, "Multi-modal medical image fusion via multi-dictionary and truncated Huber filtering," *Biomedical Signal Processing and Control*, vol. 88, 2024, doi: 10.1016/j.bspc.2023.105671.
- [15] A. Camacho, E. Vargas, and H. Arguello, "Hyperspectral and multispectral image fusion addressing spectral variability by an augmented linear mixing model," *International Journal of Remote Sensing*, vol. 43, no. 5, pp. 1577–1608, 2022, doi: 10.1080/01431161.2022.2041762.
- [16] Q. Wei, J. Bioucas-Dias, N. Dobigeon, and J. Y. Tourneret, "Hyperspectral and multispectral image fusion based on a sparse representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3658–3668, 2015, doi: 10.1109/TGRS.2014.2381272.
- [17] H. Li, X. J. Wu, and T. Durrani, "NestFuse: an infrared and visible image fusion architecture based on nest connection and spatial/channel attention models," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 12, pp. 9645–9656, 2020, doi: 10.1109/TIM.2020.3005230.
- [18] H. Kaur, D. Koundal, and V. Kadyan, "Multi modal image fusion: Comparative analysis," in *Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing, ICCSP 2019*, 2019, pp. 758–761, doi: 10.1109/ICCSP.2019.8697967.
- [19] L. Jian, X. Yang, Z. Liu, G. Jeon, M. Gao, and D. Chisholm, "SEDRFuse: a symmetric encoder–decoder with residual block network for infrared and visible image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021, doi: 10.1109/TIM.2020.3022438.
- [20] J. Ma, H. Zhang, Z. Shao, P. Liang, and H. Xu, "GANMcC: a generative adversarial network with multiclassification constraints for infrared and visible image fusion," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, 2021, doi: 10.1109/TIM.2020.3038013.
- [21] W. Yin, K. He, D. Xu, Y. Luo, and J. Gong, "Adaptive enhanced infrared and visible image fusion using hybrid decomposition and coupled dictionary," *Neural Computing and Applications*, vol. 34, no. 23, pp. 20831–20849, 2022, doi: 10.1007/s00521-022-07559-w.
- [22] H. Li, X. J. Wu, and J. Kittler, "RFN-Nest: an end-to-end residual fusion network for infrared and visible images," *Information Fusion*, vol. 73, pp. 72–86, 2021, doi: 10.1016/j.inffus.2021.02.023.
- [23] J. Di *et al.*, "FDNet: an end-to-end fusion decomposition network for infrared and visible images," *PLoS ONE*, vol. 18, no. 9 September, 2023, doi: 10.1371/journal.pone.0290231.
- [24] Y. Luo and Z. Luo, "Infrared and visible image fusion: methods, datasets, applications, and prospects," *Applied Sciences (Switzerland)*, vol. 13, no. 19, 2023, doi: 10.3390/app131910891.
- [25] A. Toet, "The TNO multiband image data collection," *Data in Brief*, vol. 15, pp. 249–251, 2017, doi: 10.1016/j.dib.2017.09.038.




BIOGRAPHIES OF AUTHORS

Ruhan Bevi Azad    is an associate professor in the Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, KTR campus. She holds a master's degree in embedded system technologies from Anna University, Chennai and a Ph.D. in security in embedded systems from SRM Institute of Science and Technology. Her research interests include deep learning, image cognition, signal and image processing with machine learning, neural networks, IoT for automation, security in embedded systems, and reconfigurable computing. She has twenty-three years of teaching experience. She can be reached via email at ruhanb@srmist.edu.in.



Hari Unnikrishnan    is working selection grade assistant professor in Department of Electronics and Communication Engineering, at Saveetha Engineering College, Chennai. He has more than twenty-five years of teaching experience. He received his bachelor's degree in electronics engineering from Madras institute of Technology, Anna University, Chennai, in year 1993. He completed a master's degree in power electronics and drives from the College of Engineering, Anna University, Chennai. He completed Ph.D. at SRM Institute of Science and Technology, Tamil Nadu, India. He has published many research papers in national and international conferences and indexed journals. His research interests include signal processing, image processing, machine learning and wireless communications. He is a member of Institution of Engineers (IE), India. He can be contacted at email: hariu@saveetha.ac.in or uharih86@gmail.com.



Lokesh Gopinath    completed his B.E in electronics and communication engineering from TJ Institute of Technology, affiliated to Anna University, Chennai, India in 2013. He obtained his M.E degree in embedded system technologies from Anand Institute of Higher Technology, affiliated with Anna University, Chennai, India in 2016. Currently, he is pursuing a Ph.D. degree in electronics and communication engineering at SRM Institute of Science and Technology, Kattankulathur, India. His research focuses on image processing and deep learning. He can be reached at lg0654@srmist.edu.in.