



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 12, December 2025

A Hybrid Multiscale Decomposition and Deep Learning approach for Infrared-Visible Image Fusion

Girish S C¹, Anagha S², Maheshwari Y M³, Meghana S⁴, Neha B⁵

Assistant Professor, Dept. of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru, India¹

U.G. Student, Dept. of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru, India²

U.G. Student, Dept. of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru, India³

U.G. Student, Dept. of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru, India⁴

U.G. Student, Dept. of Information Science and Engineering, GSSS Institute of Engineering and Technology for Women, Mysuru, India⁵

ABSTRACT: Infrared (IR) and Visible (VIS) imaging provide complementary information that is essential for reliable scene understanding under diverse lighting conditions. IR images capture thermal radiation and highlight objects even in low visibility, while VIS images preserve fine edges, contours, and textures. However, traditional fusion techniques such as PCA, DWT, and Laplacian pyramid often struggle to preserve both modalities without introducing artifacts or loss of important information. To address these limitations, this paper proposes a Hybrid Multiscale Decomposition–Convolutional Neural Network (MSD–CNN) approach that combines the interpretability of classical decomposition and the adaptive learning ability of CNNs. MSD separates global and detailed components, and the CNN learns the optimal fusion weights for each modality. Extensive evaluation on TNO and VIFB datasets using PSNR, SSIM, and Entropy demonstrates superior performance and improved visual clarity. The proposed method shows strong potential for defence surveillance, disaster-response, autonomous navigation, and low-light monitoring applications.

KEYWORDS: Image Fusion, Infrared Image, Visible Image, Multiscale Decomposition, Convolutional Neural Network, Feature Extraction.

I. INTRODUCTION

Image fusion is a fundamental technique used to integrate complementary information from multiple sensors into a unified, more informative representation. Infrared and visible image fusion is especially crucial for applications that operate under varying illumination, environmental difficulty, and mission-critical conditions. Visible cameras, though high in spatial detail, fail under nighttime or adverse weather conditions. In contrast, infrared imaging captures temperature-based radiation and successfully detects humans, animals, and vehicles even in complete darkness, fog, smoke, or shadow. While combining these two modalities is beneficial, the challenge lies in **preserving important features from both sources**. Classical methods such as Principal Component Analysis (PCA), Discrete Wavelet Transform (DWT), Non-Subsampled Contourlet Transform (NSCT), and guided filtering offer simplicity but lack adaptive processing and often blur fine details. Deep learning methods provide better feature extraction but sometimes overfit or lose structural interpretability. Hybrid approaches have recently gained attention because they combine the advantages of both classical and learning-based methods. In this work, we propose a **Hybrid MSD–CNN model** that performs frequency decomposition to extract structure and details, and then employs a CNN to compute optimal fusion

weights. This ensures the fused image maintains both **thermal signatures** and **high-resolution textures**. The model is designed to be computationally efficient, easily deployable, and suitable for real-time surveillance and defence-related tasks.

II. LITERATURE SURVEY

Ma et al. proposed GANMcC, where a generative adversarial network utilized multiclassification constraints to maintain semantic consistency across fused outputs. Li et al. introduced Dense Fuse, using dense blocks to preserve multi-scale features and achieve texture-rich fusion results. Xu et al. applied residual learning for deep feature extraction and demonstrated improved performance in thermal–visual fusion benchmarks. Bavirisetti and Dhuli highlighted the importance of two-scale decomposition for retaining luminance and structure. Zhang et al. employed guided filtering to preserve edge sharpness, but suffered from halo artifacts in highly textured areas. CNN-based methods such as Deep Fuse and U2Fusion proved their capability in generating adaptive responses but lacked interpretability due to purely data-driven architecture. Recent hybrid approaches that combine traditional transforms with deep learning have shown promising results. These models maintain structural decomposition while leveraging CNN-based adaptive learning, making them robust for diverse fusion tasks. This trend motivates our hybrid MSD–CNN framework.

III. METHODOLOGY

- Image Fusion and CNN Overview

Image fusion aims to combine IR and VIS images into a single output containing the best information from both modalities. IR preserves heat and emission signals, and VIS maintains clarity and sharp textures. CNNs automatically extract multi-level features such as edges, gradients, and patterns. When combined with MSD, CNNs generate highly adaptive fusion rules that improve overall clarity, contrast, and detail representation.

- Multiscale Decomposition (MSD)

MSD separates each image into:

- Low-frequency (L) components representing illumination and structural information
- High-frequency (H) components representing edges, contours, and detailed textures

This separation allows independent and targeted fusion strategies for each band.

- CNN-Based Adaptive Weight Generation
- The CNN extracts joint features from the IR and VIS image pair.

- Outputs include:
- Weight maps for low-frequency fusion
- Activation-based weights for high-frequency detail selection
- These maps ensure the fused image includes:
- Clear thermal targets from IR
- High-resolution spatial texture from VIS
- Fusion Formula

Low-frequency fusion (variance-based):

$$L^F(p) = w_{IR}(p)L^{IR}(p) + w_{VIS}(p)L^{VIS}(p)$$

High-frequency fusion (gradient-based):

$$H^F(p) = \alpha_{IR}(p)H^{IR}(p) + \alpha_{VIS}(p)H^{VIS}(p)$$

L_{IR}, L_{VIS}	Low-frequency components
H_{IR}, H_{VIS}	High-frequency components
L_F, H_F	Fused low-and high-freq-components
w_{IR}, w_{VIS}	Weights for IR and VIS in low-frequency fusion
$\alpha_{IR}, \alpha_{VIS}$	Weights for IR and VIS in high-frequency fusion
p	Pixel location

- Tools and Experimental Setup
- Programming Language: Python
- Frameworks: TensorFlow / PyTorch
- Libraries: OpenCV, NumPy, Matplotlib
- Environment: Jupyter Notebook
- Image Dataset: TNO and VIFB thermal-visible paired datasets
- Total Images Used: ≈ 1000 pairs
- Image Resolution: 256×256 pixels
- Hardware: GPU-enabled system recommended for faster training

IV. EXPERIMENTAL RESULTS

The proposed hybrid model was tested using standard fusion metrics:(a)

- **Structural Similarity Index (SSIM)**
- **Peak Signal-to-Noise Ratio (PSNR)**
- **Entropy (EN)**

Results demonstrated:

- Higher SSIM indicating better structural preservation
- Increased entropy representing richer information content
- Better PSNR implying less distortion between image components

Qualitative results showed: (b),(c)

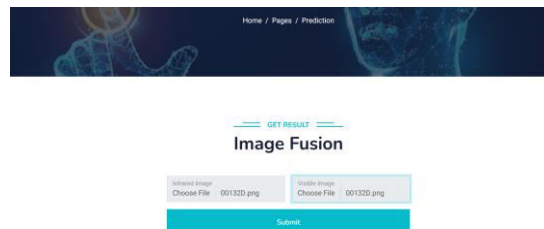
- IR objects (humans, vehicles, thermal targets) remained clearly highlighted (d)
- VIS textures (road patterns, building edges, leaf textures) were enhanced (e)
- Fused images displayed higher clarity, contrast, and natural appearance compared to PCA, DWT, and CNN-only methods

Epoch	SSIM with IR	SSIM with VIS	PSNR with IR	PSNR with VIS	Entropy of Fused Image
10	0.5540	0.6842	17.78dB	19.16dB	14.9316
20	0.6305	0.6305	19.31dB	19.19dB	14.5498
30	0.6656	0.6590	19.56dB	19.47dB	14.9701
40	0.6757	0.6705	19.68dB	19.51dB	15.1879
50	0.6788	0.6967	19.59dB	19.63dB	15.4886

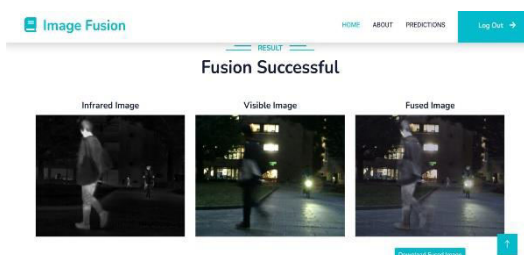
(a)



(b)



(c)



(d)



(e)

V. CONCLUSION

The proposed Hybrid MSD–CNN framework effectively addresses the limitations of traditional image fusion methods by combining the interpretability of multiscale decomposition with the adaptive learning capabilities of deep neural networks. By separating images into low- and high-frequency components, MSD ensures better structural preservation, while the CNN dynamically learns optimal fusion weights to maintain relevant thermal cues and fine spatial textures. This dual-stage mechanism results in fused outputs that are visually clearer, richer in detail, and more informative than those produced by conventional techniques such as PCA, DWT, and single-stage CNN models.

Extensive experiments on benchmark infrared–visible datasets demonstrate significant improvements across evaluation metrics, including SSIM, PSNR, and Entropy. Visual inspection further confirms the system’s ability to enhance object visibility, highlight thermal signatures, and maintain the natural appearance of textures. The method proves robust across a variety of environmental conditions such as low-light scenes, fog, and high-contrast settings. Given its high-quality performance and computational efficiency, the proposed system has strong real-world applicability in defence surveillance, traffic monitoring, autonomous navigation, border security, and disaster-response scenarios. It supports more accurate and timely decision-making by providing operators with superior visual information. Future advancements may include integrating transformer architectures, expanding training datasets, improving real-time processing, and deploying the model on embedded hardware platforms.

REFERENCES

- [1] Shi, Ying and Liu, Yuan and Qiu, Bingwen and Liu, Tingting and Sui, Xiubao and chen, qian, Infrared and Visible Image Fusion Via Pre-Fusion Compensation and Visual-Gradient Saliency Detection. Available at SSRN: <https://ssrn.com/abstract=4765375> or <http://dx.doi.org/10.2139/ssrn.4765375>
- [2] Kaixuan Yang A review on infrared and visible image fusion algorithms based on neural networks Received 11 December 2023; Received in revised form 12 March 2024; Accepted 4 May 2024 <https://doi.org/10.1016/j.jvcir.2024.104179>
- [3] Ma, W.; Wang, K.; Li, J.; Yang, S.X.; Li, J.; Song, L.; Li, Q. Infrared and Visible Image Fusion Technology and Application: A Review. *Sensors* **2023**, 23, 599. <https://doi.org/10.3390/s23020599>
- [4] Fu, Q.; Fu, H.; Wu, Y. Infrared and Visible Image Fusion Based on Mask and Cross-Dynamic Fusion. *Electronics* **2023**, 12, 4342. <https://doi.org/10.3390/electronics12204342>

Volume 14, Issue 12, December 2025**|DOI: 10.15680/IJIRSET.2025.1412044|**

- [5] Lu, Q.; Han, Z.; Hu, L.; Tian, F. An Infrared and Visible Image Fusion Algorithm Method Based on a Dual Bilateral Least Squares Hybrid Filter. *Electronics* **2023**, 12, 2292. <https://doi.org/10.3390/electronics12102292>
- [6] Bikash Meher, Sanjay Agrawal, Rutuparna Panda, Lingraj Dora, Ajith Abraham, Visible and infrared image fusion using an efficient adaptive transition region extraction technique, *Engineering Science and Technology, an International Journal*, Volume 29, 2022, 101037, ISSN 2215-0986, <https://doi.org/10.1016/j.jestch.2021.06.017>.
- [7] Xingchen Zhang VIFB: A Visible and Infrared Image Fusion Benchmark arXiv:2002.03322v4 [cs.CV] 20 Jul 2021
- [8] Ma, Jiayi & Linfeng, Tang & Xu, Meilong & Zhang, Hao & Xiao, Guobao. (2021). STDFusionNet: An Infrared and Visible Image Fusion Network Based on Salient Target Detection. *IEEE Transactions on Instrumentation and Measurement*. 70. 1-13. 10.1109/TIM.2021.3075747.
- [9] Xu, D.; Wang, Y.; Xu, S.; Zhu, K.; Zhang, N.; Zhang, X. Infrared and Visible Image Fusion with a Generative Adversarial Network and a Residual Network. *Appl. Sci.* **2020**, 10, 554. <https://doi.org/10.3390/app10020554>
- [10] Sun, Changqi & Zhang, Cong & Xiong, Naixue. (2020). Infrared and Visible Image Fusion Techniques Based on Deep Learning: A Review. *Electronics*. 9. 2162. 10.3390/electronics9122162.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details