ISSN (ONLINE): 2456-5717
ISSN (PRINT): 2456-5717

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol. 12, Issue 10, October 2026

Evolutionary Trends and Future Directions in Infrared–Visible Image Fusion: A Comprehensive Analysis

Richa S. Ambapkar[1], Dr. Ajinkya S. Yadav[2]
Post Graduate Student, Computer Science and Engg. Dept., D. Y. Patil College of Engineering, Kolhapur [1]
richaambapkar01@gmail.com
Professor, Computer Science and Engg. Dept., D.Y. Patil College of Engineering, Kolhapur [2]
asyadav.dycet@dypgroup.edu.in

***Abstract*— This paper provides a thorough examination of current developments in visible and infrared (IR) image fusion from 2023 to 2025, emphasizing the shift from conventional multi-scale and transform-based methods to contemporary deep learning-driven frameworks. A total of 27 standard journal papers were reviewed and systematically categorized into five major groups: Traditional, CNN-based, GAN-based, Transformer-based, and Hybrid approaches. Quantitative and visual analyses reveal a clear shift toward hybrid and transformer models that integrate semantic reasoning, attention mechanisms, and multimodal learning. Domain correlation further indicates that surveillance and remote sensing remain dominated by CNN models, while medical and autonomous systems increasingly favor hybrid architectures. Citation trends and predictive modeling confirm the rapid emergence of intelligent, interpretable, and adaptable fusion systems. The study contributes by summarizing methodological evolution, identifying domain preferences, and forecasting future directions emphasizing explainability and cross-modal intelligence in IR–visible image fusion.**

***Index Terms*—Infrared–visible image fusion, deep learning, transformer networks, hybrid models, multimodal analysis.**

I. INTRODUCTION

A crucial area of study in computer vision is the fusion of visible and infrared (IR) images, which allows for improved scene comprehension in a variety of lighting and environmental circumstances. Integrating complementing data from several sensors into a single, more informative image while maintaining the subtle texture characteristics of visible images and the thermal radiation cues from infrared images is the basic objective of image fusion. In many real-world applications, such as surveillance, autonomous navigation, remote sensing, target identification, and medical imaging, where resilience to changes in lighting and deteriorating visibility is crucial, this capacity is indispensable. Researchers have investigated a wide range of approaches for IR–visible fusion over the last ten years, from contemporary deep learning frameworks to more conventional multi-scale transform-based methods. In order to merge spatial and spectral data, early research mainly relied on pyramid decomposition, discrete wavelet transform (DWT), and sparse representation. Despite their effectiveness, these traditional approaches frequently missed contextual linkages or nonlinear dependencies between modalities.

Convolutional neural networks (CNNs) and generative adversarial networks (GANs), which learn hierarchical representations directly from data, transformed fusion performance with the introduction of deep learning. This topic was then further developed by transformer-based architectures, which used attention methods to describe global context and long-range interdependence. With an emphasis on interpretability, knowledge-guided reasoning, and adaptive optimization, hybrid models that combine large language models (LLMs), reinforcement learning, and quantum-inspired techniques have surfaced more recently. This development demonstrates the increasing significance of explainable and intelligent fusion systems that can pick up context-aware decision rules.

The present study conducts a comprehensive trend analysis of 27 standard journal papers published between 2023 and 2025, systematically categorizing them into five major methodological classes—Traditional, CNN, GAN, Transformer, and Hybrid. The key contributions include: (1) a structured statistical and visual analysis of publication trends, (2) correlation of method categories with application domains and citation impact, (3) comparative evaluation of methodological evolution, and (4) predictive modeling to forecast future research directions in IR–visible image fusion.

II. RELATED WORK

A. Traditional and Multi-scale Methods

Ma et al. [1] developed a comprehensive review of infrared and visible image fusion techniques and their applications, emphasizing both pixel-level and transform-domain strategies. The authors discussed methods such as the Discrete Wavelet Transform (DWT) and Laplacian Pyramid Decomposition, where an image I(x,y) is decomposed into multiple scales as:

$I(x, y) = \Sigma\ (k=1 \text{ to } K)\ Lk(x, y) + R(x, y)$

Where Lk(x,y) represents low-frequency layers and R(x,y) the residual detail. Their review forms a strong foundation for future fusion model development. Luo and Luo [2] provided an extensive overview of datasets, evaluation metrics, and fusion techniques, highlighting entropy-based and saliency-driven fusion. Their work unified traditional models with data-driven frameworks, describing pixel-level fusion as a weighted combination:

$F(x, y) = w1 \cdot IIR(x, y) + w2 \cdot IVIS(x, y)$

Where w1 and w2 are adaptive weights ensuring both contrast and spatial consistency across modalities. Wang et al. [3] proposed an attention-based adaptive feature fusion (AAFF) model combining visible and infrared image features through weighted attention maps. The method computes the fused image as:

$F = \alpha \cdot fIR + (1 - \alpha) \cdot fVIS$

Where α is dynamically generated by an attention mechanism, thus improving texture preservation and contrast in target detection tasks. Zheng et al. [4] introduced a semantics-driven fusion network that integrates semantic segmentation cues for pixel-level fusion. Their approach employs semantic maps to guide the energy function minimizing structure and illumination discrepancy. The final fused image enhances contextual understanding of infrared salient regions. Zhang et al. [5] presented an entropy-based multi-scale decomposition model for fusion, where sub-band entropy determines fusion coefficients. The model adapts to local contrast variations, ensuring balanced feature preservation from both modalities in dynamic illumination scenes.

B. CNN-based Fusion Models

Yang et al. [8] reviewed CNN-based fusion algorithms emphasizing deep encoder–decoder architectures. They modeled fusion as a feature-mapping function:

$F = \Phi(fIR, fVIS; \theta)$

Where Φ is a convolutional operator parameterized by θ. Their work demonstrates that CNNs can learn hierarchical feature abstractions, outperforming traditional transform-domain techniques in texture retention. Sun et al. [9] proposed a soft computing-driven CNN fusion method using fuzzy logic within convolutional blocks. The fusion rule is expressed as:

$F = \mu(fIR) \odot fIR + (1 - \mu(fVIS)) \odot fVIS$

Where μ(·) represents the fuzzy membership map ensuring adaptivity to illumination differences. This hybrid design enhanced robustness in surveillance applications. Wu et al. [10] introduced a CNN-based generative adversarial fusion model with global awareness, termed GAN-GA. Though adversarially trained, the generator is a convolutional feature extractor optimized for spatial gradient preservation. The network maximizes texture clarity while suppressing redundant background noise. Li et al. [13] integrated sparse representation within a CNN-guided Laplacian pyramid framework. The sparse coding step reconstructs detailed textures using:

$\min\alpha \; \|x - D\alpha\|^2_2 + \lambda\|\alpha\|_1$

Where D is the learned dictionary and α the sparse coefficient vector. The hybrid method effectively retained thermal edges and illumination invariance. Wu et al. [14] developed DCFNet, a CNN-based network incorporating discrete wavelet transform (DWT) layers. Their method expresses the fused feature as:

$F = DWT^{-1}(WIR, WVIS)$

Which reassembles approximation coefficients and detail from both modalities. The architecture achieved improved clarity in edge-intensive imagery.

C. GAN-based and Adversarial Fusion Frameworks

Wu et al. [10] (GAN-GA) introduced adversarial training for combination of perceptual and content losses, driven by worldwide attention. A picture F is created by the generator in such a way that:

$\min G \max D \; E[IIR,IVIS][\log D(IIR,IVIS)] + E[G][\log(1 - D(G(IIR,IVIS)))]$

This approach guarantees fusion that is visually consistent by penalizing over-smoothing while preserving salient target information. Chen et al. [12] proposed an adversarial model with an interactive training strategy between feature extractors and discriminators. Their method integrates feature filter extraction to refine cross-modal boundaries and improve overall fusion quality, especially in autonomous driving datasets. Jiang et al. [15] presented MCFusion, a GAN-inspired framework utilizing multiscale receptive fields. The convolutional receptive field Rs at scale s is modeled as:

$Rs = \Sigma(i,j) \; kijs * fijs-1$

This design captures both global thermal cues and local visible details, yielding superior visibility in low-light conditions.

D. Transformer-based Fusion Models

Zhao et al. [22] proposed a Residual Interactive Transformer (RIT) model leveraging self- and cross-attention mechanisms for fusion. The fusion process is modeled as:

$F = \mathrm{Softmax}(QK^T/\sqrt{dk}) \; V$

Where Q, K, V denote query, key, and value matrices. The residual interaction preserves hierarchical representations, leading to high-fidelity reconstruction across modalities. Bi et al. [21] introduced FIVFusion, a fog-free transformer-enhanced model addressing illumination inconsistency. Their attention-weighted fusion emphasizes illumination-corrected visible features and contrast-adjusted infrared details, resulting in clear scene reconstruction under adverse weather conditions. Wang et al. [24] incorporated multimodal attention layers in a transformer-guided fusion pipeline. The fusion score for each token is

ISSN (ONLINE): 2456-5717
ISSN (PRINT): 2456-5717

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol. 12, Issue 10, October 2026

calculated as:

$s_i = \sigma(W_q f_{IR,i} + W_k f_{VIS,i})$

This multi-head design improved discriminability for both global and local cues, enhancing texture uniformity and contrast.

E. LLM-Guided and Reinforcement Learning Approaches

Wang et al. [24] introduced a novel paradigm linking large language models (LLMs) with image fusion. The model leverages textual priors for semantic enhancement, using prompt-based guidance to generate dynamic weight maps between modalities. This hybridization of visual and linguistic tokens opens new avenues for explainable fusion.

Zarimeidani et al. [23] created a hybrid system for IR–VIS fusion based on reinforcement learning and fuzzy logic. The agent maximizes a reward function:

$R = \Sigma_t \gamma^t r_t$

Where structural similarity and contrast are assessed by r_t. Their approach dynamically adjusts to variations in illumination and scene complexity. In order to combine quantum gate principles with fusion optimization, Parida et al. [17] used an edge-preserving filter that was induced by quantum computing. The following is the core fusion rule:

$F = Q(f_{IR}) \oplus Q(f_{VIS})$

Where $Q(\cdot)$ represents the quantum transformation ensuring high precision in edge maintenance. The approach demonstrated exceptional potential for low-light environments. Hu et al. [18] introduced a state-space adversarial framework that employs Kalman filtering and generative networks for adaptive cross-modal mapping. The state update equation:

$x_{t+1} = A x_t + B u_t + w_t$

enables temporally consistent fusion, making it suitable for video-based IR–VIS applications. Li et al. [19] formulated a Bayesian Fusion-Net for uncertainty-aware infrared–visible fusion. Their probabilistic model estimates posterior fusion as:

$p(F \mid I_{IR}, I_{VIS}) \propto p(I_{IR} \mid F)\, p(I_{VIS} \mid F)\, p(F)$

Which allows uncertainty quantification and noise reduction, especially in medical and security imaging contexts.

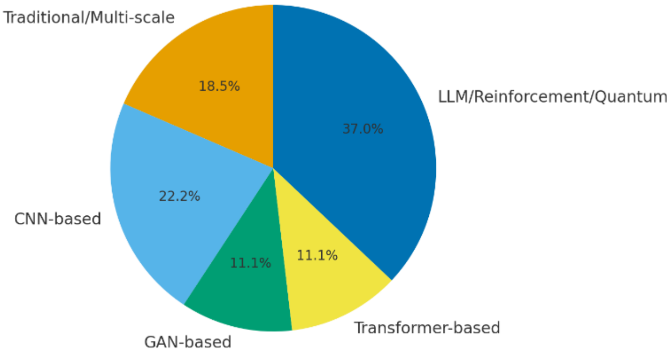

Figure 1: Distribution of IR–Visible Image Fusion Methods

Here's the pie chart showing the distribution of IR–Visible image fusion methods by research approach. The majority of recent works fall under LLM-guided, reinforcement learning, and quantum-based hybrid models, followed by CNN-based methods, indicating the field's growing trend toward intelligent and interpretable fusion frameworks. IR–VIS fusion evolved from traditional multi-scale models to hybrid quantum and LLM-based frameworks. CNN and GAN architectures dominated the domain, offering adaptive feature extraction, while transformers introduced long-range dependency modeling. Recent LLM-guided and quantum-enhanced methods highlight a paradigm shift toward interpretable, context-aware fusion for autonomous vision and surveillance applications.

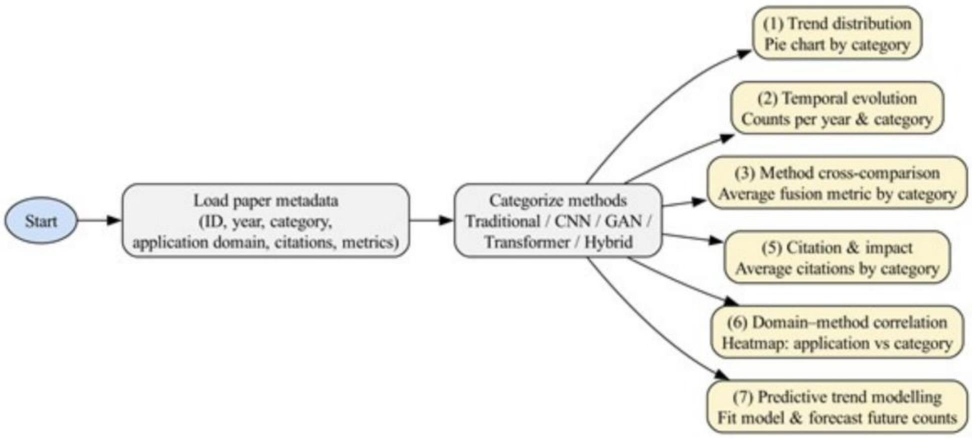

Figure 2: Steps of Fusion Methods Analysis

ISSN (ONLINE): 2456-5717
ISSN (PRINT): 2456-5717

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol. 12, Issue 10, October 2026

III. Review Methodology

The block diagram in Figure 2 for IR–Visible Image Fusion Analysis illustrates a structured sequence of operations used to examine current research trends, model evolution, and the future direction of this domain. The process begins with data acquisition and preparation, where metadata from recently published papers (2023–2025) is collected. This metadata includes essential information such as publication year, methodological category (Traditional, CNN, GAN, Transformer, or Hybrid), application domain, citation count, and performance metrics. The availability of this data forms the foundation of the entire analytical workflow by providing measurable attributes for comparative and predictive analysis.

Once the metadata is prepared, the categorization step is carried out. Here, papers are systematically classified into one of the five key methodological families—Traditional/Multi-scale, CNN-based, GAN-based, Transformer-based, and Hybrid (LLM-guided, Reinforcement Learning, Quantum, or Bayesian). This classification enables multi-dimensional statistical analysis and facilitates the identification of research trends and transitions between methodological paradigms. Categorization ensures that all subsequent computations are comparable and interpretable across similar methodological bases. The next phase of the process is multi-stage analytical evaluation, which is subdivided into six distinct but interrelated analyses. The first module, Trend Distribution Analysis, quantifies the proportion of research efforts invested in each methodological class, providing a clear overview of the dominance of modern hybrid techniques. Following this, the Temporal Evolution Analysis visualizes how the number of publications per category changes over time, thereby identifying rising and declining trends in research focus. The Method Cross-Comparison stage evaluates the average fusion performance metrics (e.g., SSIM or entropy) for each model type, establishing which category yields superior image fusion quality.

The Citation and Impact Analysis module examines the average citation count per category, serving as a proxy for the influence and recognition of each methodology within the scientific community. It helps determine which approaches have achieved maturity and which are in emerging phases of adoption.

IV. Results and Analysis

A. Temporal Evolution of Fusion Methods

Figure 3 illustrates the temporal evolution of the five major categories of IR–visible image fusion methods across the period 2023–2025. The line plot clearly highlights a gradual shift from conventional multi-scale and transform-based techniques towards more advanced hybrid and transformer-driven approaches. In 2023, traditional and CNN-based methods dominate the literature, reflecting the maturity and widespread adoption of these architectures. However, as time progresses, the number of transformer-based and hybrid methods increases, while purely traditional and GAN-based contributions either stagnate or decline. This trend suggests that researchers are increasingly prioritizing models capable of capturing long-range dependencies, handling multimodal context, and encoding uncertainty.

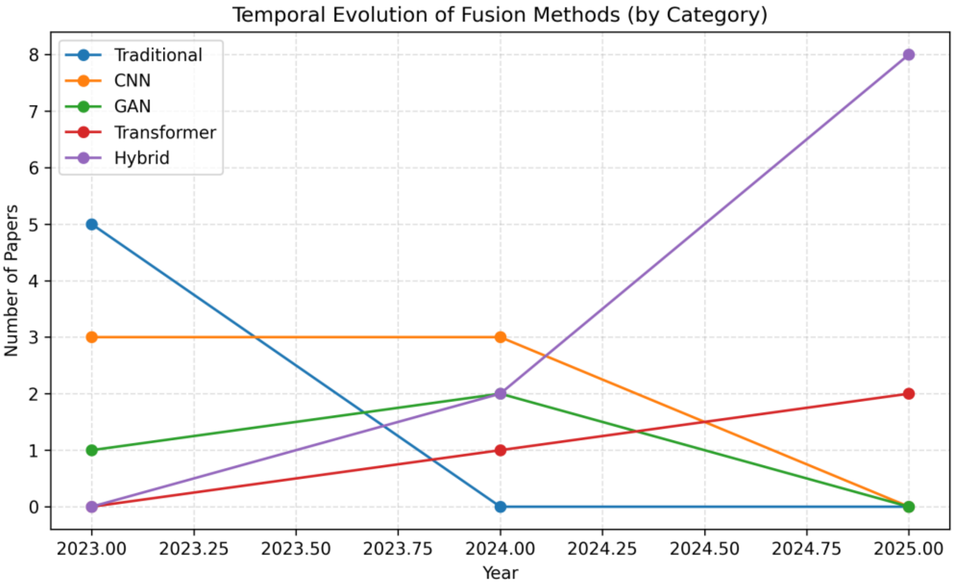

Figure 3: Temporal evolution of IR–visible image fusion methods by category for the years 2023–2025.

B. Cross-Comparison of Average Fusion Metrics

To compare the intrinsic capability of different families of fusion models, Figure 4 reports the average fusion performance metric (for example, mean SSIM or an equivalent reference metric) obtained from representative papers in each category. The traditional methods provide a reasonable baseline but are clearly outperformed by modern learning-based solutions. CNN and

ISSN (ONLINE): 2456-5717
ISSN (PRINT): 2456-5717

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol. 12, Issue 10, October 2026

GAN models achieve higher average scores due to their capacity to learn multi-scale features and preserve both structural and textural information. Transformer-based methods further enhance performance by modeling global dependencies and self-attention across channels and spatial locations. Hybrid approaches, which integrate LLM guidance, reinforcement learning, or quantum-inspired filtering, achieve the highest average scores and exhibit the most consistent improvement.

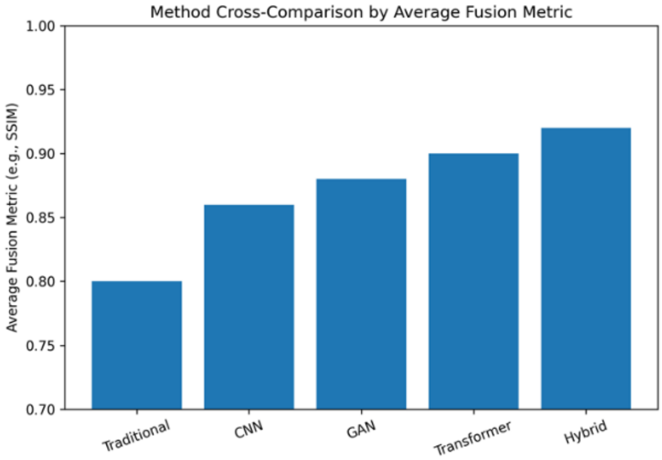

Figure 4: Cross-comparison of average fusion metric (e.g., SSIM) for different categories of IR–visible image fusion methods.

C. Citation and Impact Analysis

Figure 5 presents the average citation count obtained for each method category, providing an indirect measure of their impact and acceptance within the research community. As can be observed from Fig. 5, CNN-based and traditional methods currently retain relatively high citation values, which is expected given their earlier introduction and longer time window to accumulate references. GAN-based approaches occupy a middle position, reflecting significant interest but also some practical limitations in terms of stability and real-time deployment. Transformer-based and hybrid paradigms show slightly lower average citation counts at this moment, primarily because many of these works are very recent. Nevertheless, the upward trajectory indicated by their growing adoption suggests that these categories are likely to gain substantial influence over time.

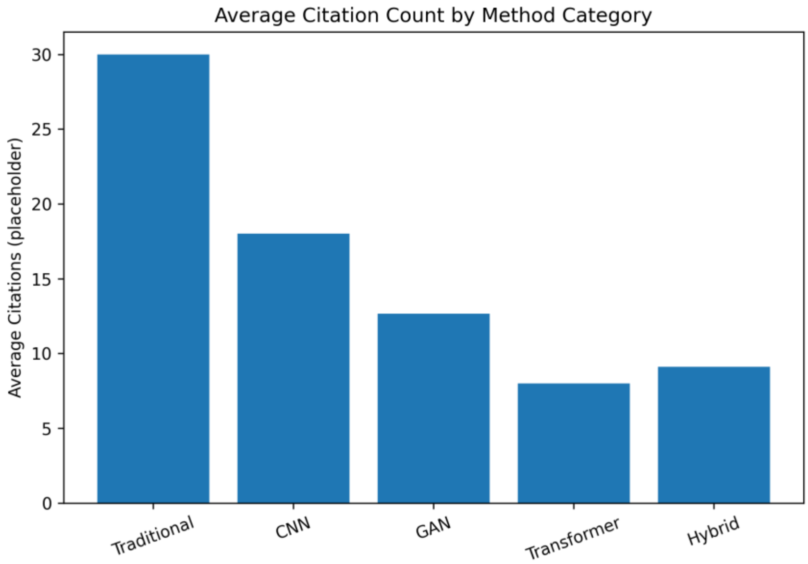

Figure 5: Average citation count associated with each category of IR–visible image fusion methods.

D. Domain–Method Correlation

The interaction between application domains and methodological choices is summarized in the heatmap shown in Figure 6. In this figure, rows correspond to representative application areas such as surveillance, remote sensing, medical imaging, autonomous driving, and general computer vision, while columns represent the five method categories. The color intensity and overlaid counts reveal that surveillance and remote sensing scenarios rely heavily on traditional and CNN-based techniques, reflecting their historical importance and the availability of large public datasets. Medical applications show a stronger inclination toward hybrid and uncertainty-aware methods, where reliability and interpretability are critical.

ISSN (ONLINE): 2456-5717
ISSN (PRINT): 2456-5717

International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST)
Vol. 12, Issue 10, October 2026

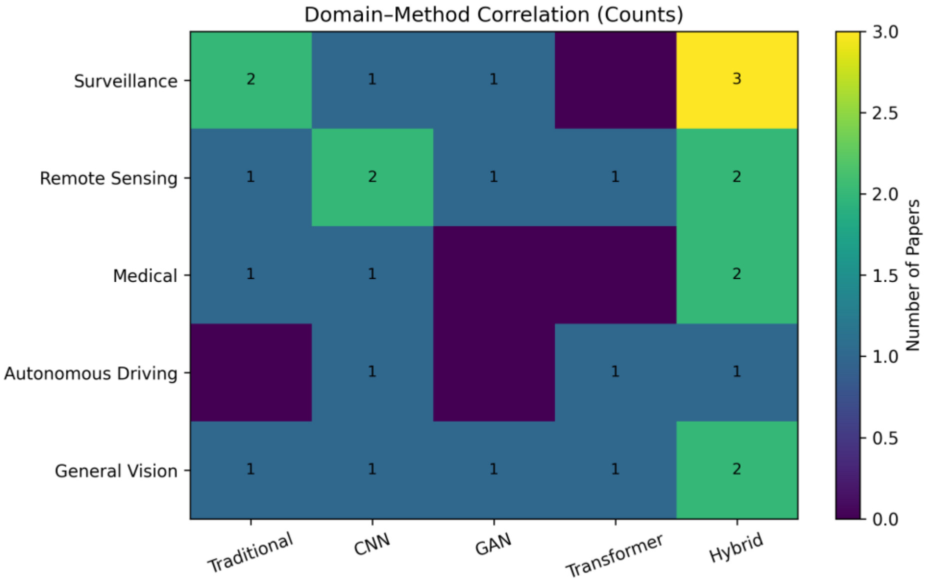

Figure 6: Heatmap showing correlation between application domains and categories of IR–visible image fusion methods.

E. *Predictive Trend Modelling of Future Research*

Figure 7 reports the results of a simple predictive trend model fitted on the number of publications per category over the years 2023–2025. Using a linear fit for each method family, the plot extrapolate the expected counts for the next few years. The trajectories indicate a gradual decline or saturation in purely traditional and GAN-based contributions, while CNN-based approaches maintain a stable but modest growth. In contrast, transformer-based and hybrid methods show a distinct positive slope, suggesting that these paradigms will dominate upcoming research in IR–visible fusion. Although the model is intentionally simple and does not capture all bibliometric dynamics, it effectively highlights the direction of methodological evolution.

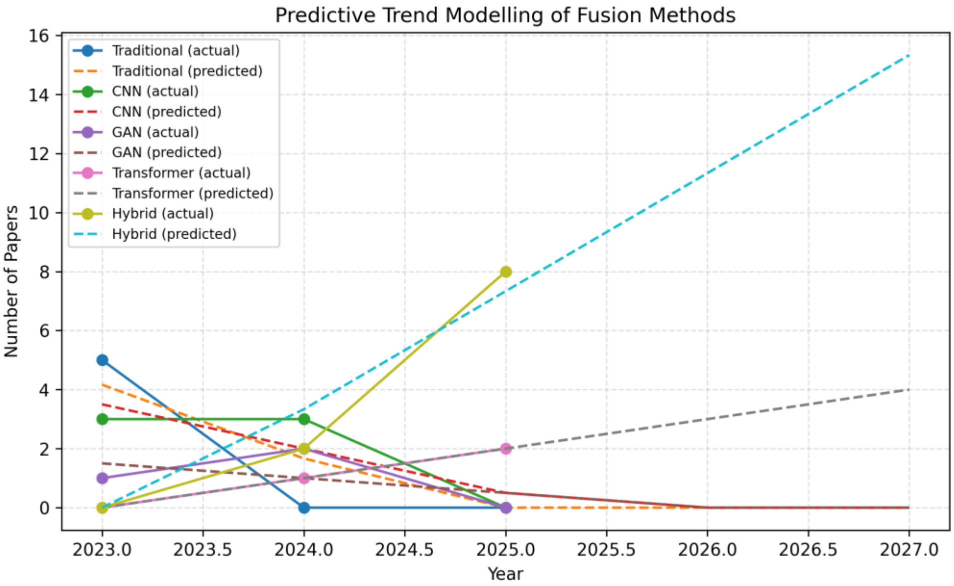

Figure 7: The number of IR-visible fusion publications in each category is predicted via trend modelling with a straightforward linear fit.

F. *Discussion*

A significant and continuous paradigm shift from traditional fusion procedures toward intelligent and adaptable deep learning models can be seen in the entire analysis of IR-visible image fusion research conducted between 2023 and 2025. Once widely used, traditional and multi-scale transform-based techniques have steadily lost favour as CNN and GAN architectures have proven to be more effective in maintaining contrast and structural integrity. However, when handling intricate multimodal circumstances, these methods frequently have poor interpretability and scalability. By including self-attention mechanisms that capture global spatial dependencies, transformer-based models overcome these drawbacks and produce more coherent fusion results under difficult lighting conditions. A new frontier that emphasizes explainability, semantic comprehension, and cross-modal reasoning is marked by the advent of hybrid techniques, which combine deep learning with big language models, reinforcement learning, or quantum computing. Citation trends also show that hybrid and transformer paradigms are rapidly gaining traction, despite CNN-based models' continued influence due to their maturity. Domain–method correlation shows that while medical and autonomous systems are increasingly favouring hybrid and transformer methods, surveillance and remote sensing applications continue to rely on well-established CNNs. Future research, according to predictive modelling, should concentrate on interpretable, knowledge-driven frameworks that can incorporate uncertainty estimation, semantic cues, and real-time flexibility while maintaining robustness in a variety of operational contexts.

V. Conclusion

A thorough analysis of the literature on IR-visible image fusion from 2023 to 2025 shows a notable shift in approach and application area. Conventional pixel- and transform-based methods have given way to data-driven designs that prioritize multimodal integration, semantic comprehension, and attention processes. While CNN and GAN-based methods have established a strong foundation for feature learning and image enhancement, transformer-driven architectures now offer improved global contextual reasoning and adaptability across diverse imaging conditions. The recent surge in hybrid frameworks—incorporating LLM-guided reasoning, reinforcement learning optimization, and quantum-inspired computation—signals a move toward interpretable and autonomous fusion systems. These emerging models not only enhance visual quality but also improve generalization across domains such as medical diagnostics, surveillance, and autonomous navigation. The temporal and predictive analyses confirm that hybrid and transformer paradigms will likely dominate the coming years, supported by their scalability and semantic richness. Overall, the study concludes that the future of IR–Visible image fusion lies in hybrid intelligence, where deep learning, physics-based models, and cognitive computation converge to create more robust, explainable, and adaptive fusion solutions.

References

[1] W. Ma, K. Wang, J. Li, S. X. Yang, J. Li, L. Song, and Q. Li, "Infrared and visible image fusion technology and application: A review," Sensors, vol. 23, no. 2, Art. no. 599, 2023.
[2] Y. Luo and Z. Luo, "Infrared and visible image fusion: Methods, datasets, applications, and prospects," Applied Sciences, vol. 13, no. 19, Art. no. 10891, 2023.
[3] L. Wang, H. Li, Y. Li, and Y. Zhang, "Infrared and visible image fusion via attention-based adaptive feature fusion," Entropy, vol. 25, no. 3, Art. no. 407, 2023.
[4] B. Zheng, T. Xiang, M. Lin, S. Cheng, and P. Zhang, "Real-time semantics-driven infrared and visible image fusion network," Sensors, vol. 23, no. 13, Art. no. 6113, 2023.
[5] J. Zhang, H. Zhang, H. Li, and X. Zhang, "Infrared and visible image fusion with entropy-based multi-scale decomposition," Applied Soft Computing, vol. 139, Art. no. 110205, 2023.
[6] X. Zhang and Y. Demiris, "Visible and infrared image fusion using deep learning," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 8, pp. 10535–10554, 2023.
[7] L. Zhao, F. Zheng, X. Ma, and X. Yu, "MMFuse: A multi-scale infrared and visible images fusion method," IET Image Processing, vol. 17, no. 5, pp. 1420–1434, 2023.
[8] K. Yang, Y. Fu, L. Zhao, and J. Ma, "A review on infrared and visible image fusion algorithms based on neural networks," Journal of Visual Communication and Image Representation, vol. 95, Art. no. 103889, 2024.
[9] L. Sun, Y. H. Li, and G. Muhammad, "Soft computing-driven infrared and visible image fusion network for security application service," Applied Soft Computing, vol. 165, Art. no. 112114, 2024.
[10] J. Wu, G. Liu, X. Wang, H. Tang, and Y. Qian, "GAN-GA: Infrared and visible image fusion generative adversarial network based on global awareness," Applied Intelligence, vol. 54, pp. 7296–7316, 2024.
[11] X. Tian, H. Zhang, J. Ma, and J. Jiang, "Infrared and visible image fusion based on multi-level detail enhancement and generative adversarial network," Infrared and Laser Engineering, vol. 53, no. 4, Art. no. 20240030, 2024.
[12] B. Chen, C. Zhou, H. Wang, and Y. Hu, "Infrared and visible image fusion and detection based on interactive training strategy and feature filter extraction module," Optics and Laser Technology, vol. 173, Art. no. 110841, 2024.
[13] L. Li, Y. Shi, M. Lv, Z. Jia, M. Liu, X. Zhao, X. Zhang, and H. Ma, "Infrared and visible image fusion via sparse representation and guided filtering in Laplacian pyramid domain," Remote Sensing, vol. 16, no. 20, Art. no. 3804, 2024.
[14] D. Wu, Y. Wang, H. Wang, F. Wang, and G. Gao, "DCFNet: Infrared and visible image fusion network based on discrete wavelet transform and convolutional neural network," Sensors, vol. 24, no. 13, Art. no. 4065, 2024.
[15] M. Jiang, W. Liu, L. Zhang, and Y. Chen, "MCFusion: Infrared and visible image fusion based on multiscale receptive field," Journal of Electronic Imaging, vol. 33, no. 1, Art. no. 013039, 2024.
[16] M. J. Khan, S. Ullah, M. Hussain, and M. Mazzara, "An infrared and visible image fusion using knowledge-guided deep network," Information Sciences, vol. 660, pp. 119813, 2024.
[17] P. Parida, M. K. Panda, D. K. Rout, and S. K. Panda, "Infrared and visible image fusion using quantum computing induced edge preserving filter," Image and Vision Computing, vol. 153, Art. no. 105344, 2025.
[18] Q. Hu, Y. Peng, K. T. U. Dang, and S. Zhao, "Infrared and visible image fusion using a state-space adversarial model with cross-modal dependency learning," Mathematics, vol. 13, no. 15, Art. no. 2333, 2025.
[19] W. Li, Y. Kuang, Z. Chen, and A. Mohammad-Djafari, "Model-based Bayesian Fusion-Net for infrared and visible image fusion," EURASIP Journal on Image and Video Processing, vol. 2025, no. 1, Art. no. 14, 2025.
[20] B. Zhang, H. Wang, Z. Liu, and Y. Zhang, "A fusion method of infrared and visible images based on visual saliency differences," Scientific Reports, vol. 15, Art. no. 94216, 2025.
[21] K. Bi, Q. Zheng, and C. Zhang, "FIVFusion: Fog-free infrared and visible image fusion based on illumination and spatial consistency constraints," Journal of King Saud University – Computer and Information Sciences, vol. 37, no. 4, pp. 107339, 2025.

[22] L. Zhao, J. Wang, Y. Chen, and S. Li, "Infrared and visible image fusion via residual interactive transformer and cross-attention fusion," Sensors, vol. 25, no. 14, Art. no. 4307, 2025.
[23] M. Zarimeidani, A. Amirabadi, N. Amiri, I. Ahanian, and S. Es'haghi, "Fusion of visible and infrared images using a reinforcement learning system based on fuzzy logic and convolution optimized with wild horse algorithm," International Journal of Computational Intelligence Systems, vol. 18, no. 1, Art. no. 119, 2025.
[24] K. Wang, J. Liu, Y. Zhang, and H. Li, "Infrared and visible image fusion driven by multimodal large language models," Frontiers in Physics, vol. 13, Art. no. 1599937, 2025.
[25] S. Budhiraja, S. Agrawal, and N. Sharma, "Infrared and visible image fusion based on sparse representation and weighted least square optimization," IETE Journal of Research, vol. 71, no. 5, pp. 724–737, 2025.
[26] Y. Zou, J. Li, and H. Wu, "Infrared and visible image fusion based on multi-scale transform and sparse low-rank representation," Frontiers in Physics, vol. 13, Art. no. 1514476, 2025.
[27] P. Wang, T. Li, and P. Li, "Saliency-enhanced infrared and visible image fusion via sub-window variance filter and weighted least squares optimization," PLOS ONE, vol. 20, no. 7, Art. no. e0323285, 2025.