Vol. 06, No. SI*, pp. 135 –160 (2025)
ISSN: 2708-0757

JOURNAL OF APPLIED SCIENCE AND TECHNOLOGY TRENDS

www.jastt.org

Adaptive Pan-sharpening via Contourlet Decomposition and Anisotropic Diffusion for Multispectral-Panchromatic 2D signals

Suraj Thakur[1], Ayush Dogra[1], Shalli Rani*[1], Bhawna Goyal[2], M Sravan Kumar Reddy[3]

[1]*Chitkara University Institute of Engineering and Technology, Rajpura, Punjab, 140401, India. suraj.3917@chitkara.edu.in; ayush.dogra@chitkara.edu.in; shalli.rani@chitkara.edu.in*
[2]*Department of Engineering, Marwadi University Research Centre, Marwadi University, Rajkot, Gujarat, India, 360003. bhawnagoyal28@gmail.com*
[3]*RGM College of Engineering and Technology (Autonomous), Nandyal Dist., AP, India. sravankumarreddy.m@rgmcet.edu.in*

**Correspondence: shalli.rani@chitkara.edu.in*

Abstract

Pan-sharpening is now indispensable in remote-sensing computing by means of its ability to render not only high-spatial-resolution images but also precise spectral fidelity to the images. These features are essential for competent environmental measurement and investigation. The paper presents a contourlet-based pan-sharpening algorithm, which aims to combine the high spatial resolution of the panchromatic (PAN) images with the multispectral (MS) ones to extract high spectral resolution of the fused image. The algorithm tends to enhance the details of the space and reduce the spectral distortion of a multi-stage fusion procedure to some extent. The structure initiates with contourlet decomposition images captured by both PAN and MS, thus resembling directional and multi-scale structure. At the high-frequency sub-bands, an anisotropic filter is used to reduce noise, leaving salient edges intact, and a maximum-absolute selection rule is applied iteratively to inject spatial details effectively. In order to have spectral fidelity and to maintain global contrast on low-frequency components, saliency maps are used in conjunction with adaptive weight maps. The sub-bands reconstructed progressively give a high-resolution, artifact-minimal fused image. The proposed framework is validated through benchmark remote sensing datasets. The proposed methodology is based on objective metrics. Performance evaluation based on a benchmark remote sensing dataset demonstrates that the proposed methodology outperforms existing approaches in terms of sharper edges, contrast, and spectral features and offers an efficient solution for environmental applications.

Keywords: *Pan-sharpening, multi-spectral, panchromatic image, contourlets, anisotropic, fused image.*
Received: October 20th, 2025 / Revised: December 14th, 2025 /Accepted: December 24th, 2025 / Online: December 31st, 2025

I. INTRODUCTION

Remote sensing has become one of the key technologies to address some of the most challenging problems faced by the world, such as climate change, the rapid pace of urbanization, and the destruction of the ecosystem. It is applied to manage the sustainable natural resources, to model the hazards predictively, and to preserve the biodiversity by non-invasive data collection of the surface properties of the Earth [1]. Aerial imaging and unmanned aerial vehicles (UAVs) combined with high-resolution satellite imagery provide real-time remote sensing geospatial information of large and non-uniform environments, effectively overcoming the limitations of geographical and climatic heterogeneity [2]. Remote sensing can be used to monitor large and heterogeneous land scopes almost in real-time due to the usage of satellite imagery, aerial photography, and unmanned aerial vehicles (UAVs), without considering geographical and climatic constraints. It has a combination of raw imagery with sophisticated computational processes, which render it an action environmental intelligence and therefore an indispensable means of environmental monitoring and analysis [3]. Remote sensing is an advanced technological system that contributes highly to the precision, uniformity, and extensiveness of environmental monitoring systems. PAN and MS cameras, their typical resolutions and spectral ranges, and their practical significance in Earth observation applications such as urban mapping, environmental monitoring, agriculture, and land-use analysis.

Representative examples from operational satellite systems (e.g., WorldView, Landsat, and Sentinel missions) have also been incorporated to ground the discussion in real-world contexts. Due to a high level of spatial coverage, it is capable of obtaining large geospatial data of large areas under one scan. The precise temporal coverage enabling the satellite constellations allows a periodic observation, which helps to detect a change of season and quickly evaluate a sudden phenomenon related to the environment, like a natural disaster. Moreover, being a non-invasive method of observation, remote

***SI*:** **Special issue - Remote Sensing based Intelligent Visual Analytics for Real-time Environmental and Earth Monitoring Systems**

doi:10.38094/jastt605678

sensing does not cause ecological disturbance, and multispectral and hyperspectral data can be measured in the invisible spectrum [4]. The quality of providing a time and space consistency of the environmental assessment is because of the standardization and reproducibility of the satellite imagery. Besides these, by connecting remote sensing data with Geographic Information Systems (GIS), Artificial Intelligence (AI), and cloud-based environments smoothly, one can perform more complex analytical operations, such as predictive modelling, automated classification, spatiotemporal changes detection, and analytics about the environment at a large scale [5]. Depending on the sensor, platform, and wavelength applied, different techniques of remote sensing will benefit certain environmental monitoring requirements.

Optical remote sensing identifies vegetation, land cover, and water resources using visible and near-infrared wavelengths. Thermal remote sensing detects emitted heat energy to analyse surface temperature patterns, urban heat islands, and volcanic activity. Microwave or radar sensing (SAR) operates day or night and through cloud cover, excelling in terrain mapping, flood monitoring, and soil moisture assessment. Hyperspectral imaging captures hundreds of narrow spectral bands to distinguish materials and diagnose environmental conditions. Lidar (Light Detection and Ranging) is a laser pulse technique that generates an elevation map and vegetation structure in detail, which is very accurate in terms of the structure. Pan-sharpening is a very essential 2 method in remote sensing, and it is used to overcome the trade-off between space and spectral resolution in satellite imaging systems.

PAN images contain great spatial details with little spectral information, and multispectral (MS) images are rich in spectral information but low in spatial resolution. Pan sharpening balances spatial and spectral resolution to create a single image with high spatial and spectral resolution, which is useful in environmental monitoring, urban planning, agriculture, surveillance, and medical research. However, pan-sharpening encounters various challenges, such as balancing spectral fidelity and spatial resolution, data heterogeneity because different sensors can have different characteristics, artifact prevention, and computational efficiency [6]. The conventional pan-sharpening techniques have tried to deal with these challenges by various techniques, such as Principal Component Analysis (PCA), Intensity-Hue-Saturation (IHS), and the Brovey Transform method. These techniques integrate MS data with spatial information provided by the PAN image [7].

The majority of the existing methods are prone to cause spatial distortions to the spectral factors, or vice versa, enrich spectral information at the expense of spatial information (typically at high resolution) [8]. These techniques are computationally efficient, but they can generate spectral distortions with a wide range of sensing conditions, which reduces their usefulness. The recent development in deep learning, especially with the Convolutional Neural Networks (CNNs), provides potential alternatives to the conventional pan-sharpening algorithms. CNNs are good at acquiring complex nonlinear correlations between PAN and MS images, enhancing the compromise between spatial and spectral quality [9]. Radiometric index maps (e.g., NDVI, SAVI) represent domain-specific priors, which, together with sensor-oriented augmentation strategies, dramatically enhance fusion accuracy without necessarily adding a lot of model complexity [CNN Custom]. These developments reduce the limitations of the traditional and modern model-based methods, especially where the sensors are heterogeneous, thus increasing robustness in the real world.

This paper presents efficient pan-sharpening methods that preserve spatial-spectral fidelity. The fact that the proposed method combines domain-sensitive input augmented with a simplified network structure effectively correcting the flaws of the traditional and CNN algorithms, enhancing uniform performance in a wide range of sensing scenarios [10]. It is designed to work effectively in diverse platforms; therefore, it is well adapted for practical monitoring of the environment where precision and effective fusion are required to make informed decisions [11].

Chipman in 1995 proposed a wavelet-based image fusion method preserving the multi-resolution features effectively [12]. Nevertheless, the method faces the weakness of computational efficiency and is not robust when it comes to artifacts of heterogeneous sensors. Simone, in 2002, emphasized enhancing the spatial and spectral information of the remote sensing fused image. Their techniques [13] are best in certain circumstances but cannot generalize to the modalities and cope with high-noise situations. Liu in 2017 developed a deep convolutional neural network to implement multi-focus image fusion [14], which described that the method produced a high range of detail maintenance. Although effective, the model is computationally intensive, and the optimization of the model to low-light or noisy environments is not established, thus restricting the usefulness of the model in practice. The article by Alparone in 2015 introduced the framework of remote sensing image fusion based on multi-scale and pan sharpening techniques developed by the authors of the article-in-focus and their integration into a single framework. Although it is effective to solve particular remote sensing problems, the framework cannot easily scale to large arrays of multiple sensors, is sensitive to multi-sensor artifacts, and necessitates data- driven solutions [15].

II. Related Work

Pan sharpening aims to combine complementary information in multi-modal sources to generate an image with high spatial resolution and spectral fidelity, used in a wide range of applications, including remote sensing, medical diagnostics, and surveillance. Although recent fusion techniques emphasise detail preservation, contrast enhancement, and computational efficiency quantitatively assessed through metrics including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Mutual Information Fusion (MIF. The algorithm suffers from modality-specific distortions, structural inconsistencies, and spectral degradation. To address these shortcomings, various studies have contributed diverse approaches. Li et al. (2013) provided a guided filtering-based visible infrared fusion method that employed edge preserving smoothing to eliminate artifacts and enhance cross-modality structure consistency [16]. In [17], Jie et al. designed a medical image fusion algorithm based on extended difference-of-Gaussians (DoG) and anisotropic edge-preserving filters, achieving superior contrast and diagnostic interpretability at high computational cost, restricting the scalability. Meng in 2022, proposed a vision transformer-based pan-sharpening

model, using the self-attention mechanisms to provide high spatial-spectral fidelity (high PSNR), but its computationally intensive implementation restricts its use in real-time applications in medical and multi-focus fusion. [18].

In 2015, Kumar et al. implemented a cross-bilateral filter-based visible-infrared fusion algorithm, utilizing spatial and intensity kernel weighting to perform robust edge preservation. [19]. Though the algorithm maintains edges, it has noisy scenes with low contrast, limiting its adaptability for medical imaging. Li et al. in 2018 [20] suggested a structure-aware fusion framework based on structural similarity measures, resulting in sharper and more distinct boundaries, yet lacking multiscale decomposition capability, thereby reducing robustness to noise. Zhang in 2022 employed local extreme maps to enhance multi-modal brain image fusion, using intensity- guided weighting to enhance contrast, though its generalisation beyond neuroimaging remains limited [21]. Jie et al. in 2024 [22] advanced medical image fusion with multi-dictionary sparse coding and truncated Huber loss, improving noise robustness and structural consistency with high SSIM, but the predefined dictionaries and high complexity hinder real-time use. Sufyan in 2022 proposed contrast and structure extraction for anatomical image fusion, optimizing structural 3 detail, but exhibit computational inefficiency under resource-constrained applications [23].

Additionally, Ma et al. in 2017 developed infrared-visible fusion using saliency maps and weighted least squares optimisation, enhancing salient regions with high perceptual quality. However, its applicability was limited to non-medical and non-multi-focus imaging scenarios [24]. In 2019, Qiu et al. introduced guided filter-based multi-focus fusion via focus region detection to preserve sharp areas, yet it is ineffective for infrared or medical modalities [25]. Bavirisetti in 2016 proposed a two-scale fusion for visible-infrared images using saliency detection in the wavelet domain, but risked information loss in non-salient areas [26]. Liu in 2015 presented a multi-scale transform and sparse representation framework, adaptable across modalities, though its effectiveness is hindered by high computational complexity and the absence of modal-specific optimization [27]. Despite these technological advancements, current remote sensing and image fusion methodologies still encounter several fundamental challenges. These include limited scalability when extended to heterogeneous multi-sensor environments, poor robustness against cross-sensor discrepancies and radiometric distortions, and reduced performance under adverse imaging conditions such as high noise levels or low illumination.

Furthermore, weak spatial–spectral interaction modelling, reliance on manual hyperparameter tuning, and absence of probabilistic uncertainty quantification significantly constrain their generalization and reliability. Such limitations emphasize the imperative for next-generation, data-driven, and optimization-oriented fusion architectures that can holistically reinforce structural fidelity, spectral preservation, and computational efficiency across diverse remote sensing modalities. The most current developments on the pan-sharpening deal with spectral and spatial anomalies in the synthesis of high-resolution panchromatic (PAN) and low-resolution multispectral (MS) images. Liu et al. in 2024 [38] address the problem of low inter-modal correlation and describe the injection errors together with multimodal texture correction (intensity, gradient, and deep A-PNN-contraband), adaptive degradation filtering, and edge detail fusion with high quality of Q-index, SAM, and ERGAS. By using a selective update block (SUB) with gated forgetting/filtering both in spatial-frequency space and frequency focal loss to restore high-frequency texture, Wang et al. (2025)[39] achieve state-of-the-art SSIM, UIQI, and LPIPS on multiple datasets. Zhang in 2025 [40] assess the suitability of algorithms to atmospheric, sensor, and scene variations, demonstrating limited spectral consistency and spatial fidelity under practical conditions. One can state that, collectively, these publications promote adaptive and distortion-resistant fusion paradigms needed in the reliable HRMS remote sensing.

III. PRELIMINARIES

A. *Contourlet Decomposition*

The contourlet transform is a multiscale and multidirectional image decomposition framework designed to efficiently represent two-dimensional signals with smooth contours, edges, lines, and curvilinear structures [28]. It integrates a low-pass (LP) and high-pass (HP) filter for multiresolution decomposition with a Directional Filter Bank (DFB) to achieve anisotropic and highly directional sub-band decomposition, enabling sparse representation of images with complex geometric features [29]. Unlike traditional wavelet transforms, which are limited to isotropic scaling and limited directional selectivity, the contourlet transform employs non-separable filter banks that capture textures and edges and directional information effectively with fewer coefficients. This makes it especially convenient in image enhancement and fusion applications under adverse illumination, and has been demonstrated to perform their wavelet- based counterparts in the ability to preserve and enhance meaningful geometric features such as edges and contours[18]. Mathematically, the contourlet decomposition of an image f is defined as:

$$a_S[p] = \langle f, L_{S,v} \rangle \quad (1)$$

$$d_{s,r}^{(m)}[p] = \left\langle f, \rho_{s,r,p}^{(m)} \right\rangle \quad (2)$$

Here, f denotes the input image, while $a_S[p]$ represents the represent the lowpass coefficient, $d_{s,r}^{m}[p]$ Refers to directional contourlet coefficients. The function $L_{S,v}$ is the lowpass basis at scale S and subband v, measuring coarse-scale intensity variations. In contrast, $p_{s,r,p}^{(m)}$ Represents the contourlet basis function at a given level of the decomposition m, scale s, direction r, and spatial position $p = [p_1, p_2]$. The operator $\langle ., . \rangle$ denotes the inner product of the image, f, and the basis function, measuring the degree of correlation.

Equation (1) provides the approximation coefficients of large-scale intensity variation at low frequency, and Equation (2) is an approximation of small-scale edges and contours of various orientations of the high frequency. Parameters s and r determine the magnitude of the resolution and the filter orientation, respectively, while m is the number of directional sub-bands, controlling the angular resolution of the decomposition. Figure 1, the Contourlet transform uses a Laplacian Pyramid (LP) framework to analyse the scale-space representation and the Directional Filter Bank (DFB) to extract

directional detail components, effectively producing bandpass directional sub bands with high directional sensitivity [29].

The overall model of the coefficients, $\{a_S[p], d^m_{s,r}[p]\}$, provides a representation of an image in compact and descriptive terms, preserving the high-energy structural intensity with finer directional features. Due to this property, direction consistency and structural sharpness are crucial properties in image fusion, which makes the Contourlet Transform quite useful in image fusion applications.

B. Anisotropic diffusion

Anisotropic diffusion is a nonlinear image enhancement technique that models spatially variant diffusion through partial differential equations (PDEs) to selectively smooth images while preserving prominent edges and structures [30]. In contrast to isotropic diffusion, which performs uniform smoothing across the image, anisotropic diffusion enables intra-region smoothing and suppresses inter-region blurring, thus maintaining edge sharpness even at coarse resolutions. This property makes it particularly useful for image denoising, enhancement, and pan sharpening tasks in remote sensing. The anisotropic diffusion process can be expressed as:

$$P_n(r,s) = aniso'\big(I_{out1n}(r,s)\big) \qquad (3)$$

Here, $Pn(r,s)$ is the anisotropically filtered output at a spatial point (r,s) by applying an anisotropic operator $aniso'(\cdot)$ on image $I_{out1n}(r,s)$. The operator tends to be spatially adaptive and nonlinear, regulating diffusion according to local gradient magnitude. Consequently, strong edges experience limited smoothing while homogeneous areas are denoised effectively [31].

Similarly, for a second intermediate image representation, the anisotropic filtering can be formulated as:

$$Qn(i,j) = aniso'\big(Iout2n(i,j)\big) \qquad (4)$$

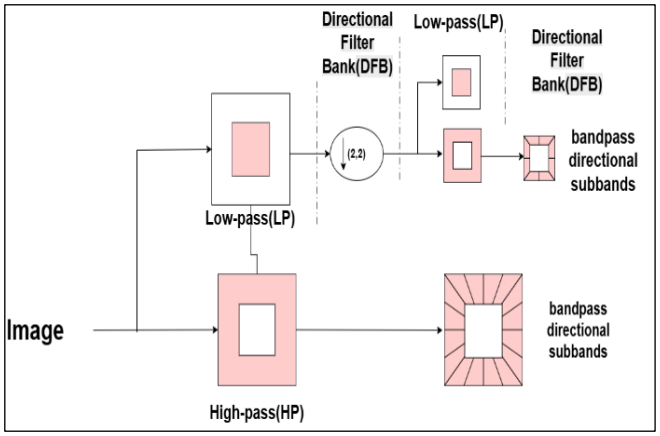

Fig. 1. Contourlet Decomposition

The equation $Qn(i,j) = aniso'(I^n_{out2}(i,j))$ has a similar formulation to the earlier one, except that the anisotropic. Operator $aniso'(\cdot)$ acts on a different intermediate representation $I^n_{out2}(i,j)$. In contrast to the fact that, unlike, $I^n_{out2}(i,j)$. The value of this component can be associated with a single stage, a single channel, or just a single feature map of the fused image. The value of $Qn(i,j)$ denotes an alternate or complementary input, possibly of a different spectral band, another spatial layer, or another processing path. The algorithm uses the same anisotropic transformation on this second input stream, returning us to edge-aware filtering uniformly across streams of data whilst adjusting itself to directional tone variations. The resulting coefficient $Qn(i,j)$ is therefore a parallel structural predictor to the $n(i,j)$, and incorporates extra frequency or contextual data. These outputs, when combined, can be complementary anisotropic responses, allowing the image fusion process to release finer spatial structures and spectral integrity as well as reduce noise and redundant artifacts. To capture residual details, the unfiltered high-frequency content can be obtained as:

$$R_n(k,l) = Iout1_n(k,l) - P_n(k,l) \qquad (5)$$

The equation $R_n(k,l) = I^n_{out1}(k,l) - P_n(k,l)$ defines the residual component $R_n(k,l)$ as the difference between the original input $I^n_{out1}(k,l)$ and the processed output term $P_n(k,l)$. If $P_n(k,l)$ indeed corresponds to the anisotropically filtered result $n(k,l)$ from the earlier formulation (with minor notational variations); this subtraction effectively measures the portion of the signal that is not captured by the anisotropic transformation. Practically, this kind of residual computation is at the heart of an image processing code and deep learning architecture: it isolates details of the image, noise patterns, or structural inconsistency, which is left behind after the operations of smoothing and transformation. In the pan sharpening or fusion task scenario, the residual term offers a natural way to restore lost fine-grained spatial information or emphasize discrepancies to correct, thereby increasing structural and spectral fidelity. A similar residual formulation for the secondary channel is given as:

$$S_n(v,k) = I_{out2n}(v,k) - Q_n(v,k) \qquad (6)$$

The second input channel is calculated using an equation known as the residual computation by $S_n(v,k) = \mathrm{I}^{\mathrm{n}}_{\mathrm{out2}}(v,k) - \mathrm{Q_n}(v,k)$ The calculation is quite similar to that in the third equation. In this case, after the anisotropic transformation in the second equation, the resulting value $Q_n(v,k)$ The optical process is removed in the form of the initial input $I_n\ out2(v,k)$. This gives the remaining value, $S_n(v,k)$, that has the information that is not contained in the anisotropically smoothed or transformed version of the second input. The wider framework of pan sharpening or fusion systems, that is, an accurate reconstruction of fine spatial details, is ensured by virtue of channel-specific subtleties at $S_n(v,k)$ and adds to the equal consideration of spectral and spatial features.

C. Max-Absolute Rule

The Max-Absolute rule is a widely adopted selection rule in image fusion, primarily applied on the high-frequency decomposition bands where edge and textural information dominate [32]. It operates on the principle that the coefficient

with the largest absolute value carries the most significant local detail, such as edges or texture variations. Formally, it is defined as:

$$F_k(s,d) = \begin{cases} P_k(s,d), & if\ |P_k(s,d)| \geq M_k(s,d)| \\ M_k(s,d), & if\ \ |M_k(s,d)| \geq P_k(s,d)| \end{cases} \quad (7)$$

In this formulation, $F_k(s,d)$ denotes the fused high-frequency coefficient of the spatial position (s, d) for the k^{th} directing sub-band of a multiscale transform. The parameters, $P_k(s,d)$ and $M_k(s,d)$, represent the respective high-frequency coefficients of the panchromatic (PAN) and multispectral (MS) images, respectively, in the same subband and at the same location. The rule simply selects the coefficient with the higher absolute magnitude, assuming it corresponds to a region with stronger edge or textural information. This mechanism effectively enhances local contrast and edge sharpness in the fused image. However, since it relies solely on intensity magnitudes, it may occasionally preserve noise or small artifacts. To overcome this limitation, advanced techniques such as anisotropic diffusion–based residual fusion adaptively refine coefficient selection using local gradients and spatial correlation, achieving superior perceptual and structural quality.

D. Weight Maps

Weight maps spatially varying structures used in image fusion to determine the contribution of each input image to the fused image at every pixel or region. Unlike the fixed weighting method, these maps are computed adaptively, based on local image properties including variance, gradient strength, contrast, or saliency, allowing the algorithm to dynamically identify and emphasize regions containing sharp edges or texturally rich features [33]. In multiscale decomposition-based fusion models, the weight maps are essential towards determining the optimal combination of low-frequency (structure) and high-frequency component (detail), ensuring that the resulting fused image inherits complementary spatial and spectral features of the source images. The weights are also normalized over all the sources of input so that the total of the weights adds up to unity over each spatial coordinate. This normalization imposes a balanced and unbiased contribution on each of the images, preventing the dominance of modalities. Consequently, weight maps function as a localized control mechanism, which prioritizes perceptually and structurally significant areas to enhance the overall fusion quality in terms of sharpness, contrast, and information content. The fused image can be mathematically represented as a weighted linear combination of the input images:

$$F\ (l,r) = \sum_{j=1}^{N} \lambda j\ (l,r)\ \cdot\ S_j\ (l,r) \quad (8)$$

Where F (l, r) is the fused image, intensity at spatial position (l, r). S_j (l, r) denotes the corresponding intensity of the j^{th} source image (panchromatic or multispectral), and λ_j(l, r) is the adaptive weight map of that source image at the same spatial location. These weight maps are spatially variant and are adaptively computed using local quantities like intensity variance, gradient energy, edge strength, or saliency. This adaptive nature enables the fusion algorithm to assign higher weight to regions with greater structural or textural significance, ensuring that the most informative parts of each input are retained in the final product. The total number of source images combined in the fusing process is termed N. To avoid preference for a specific image, and to ensure balance in terms of effects, the weight maps are typically normalized at every pixel position such that the sum of the weight maps is equal to one:

$$\sum_{j=1}^{N} \lambda_j\ (l,r)\ =\ 1 \quad (9)$$

Which guarantees that the total contribution of all input images at any spatial position remains equal to one. This constraint maintains radiometric balance, preventing disproportionate influence from any individual source and ensuring that the fused image exhibits consistent brightness, contrast, and visual coherence. In practical multiscale fusion pipelines, separate weight maps are typically employed for different frequency components. Low-frequency weight maps guide the integration of global structures and smooth intensity transitions, while high- frequency weight maps control the fusion of edges, textures, and fine spatial details. This frequency-specific weighting enables selective enhancement of both structural integrity and detail sharpness. Consequently, the fused image achieves an improved balance between spatial resolution enhancement and spectral fidelity preservation, leading to a perceptually natural and information-rich fusion output.

IV. EXPERIMENTS

This section elaborates on the details of the datasets that are used to validate the proposed fusion algorithm. The experiments are designed to confirm fairness in comparison with established methods, and Objective evaluation is performed using PSNR, SSIM, Entropy, SF, Corr, MIF, ERGAS, QABF, and LABF to comprehensively assess spatial, spectral fusion, as explained in Table I.

Step 1: Preprocessing

1. Read input PAN image A and MS image B.
2. Convert the PAN image to grayscale if required.
3. Resize both images to a common target size (e.g., 256 × 256).
4. Convert images to double precision and normalize intensities to [0,1].
5. Convert MS image B from RGB to YCbCr color space.
6. Extract luminance Y_B and chrominance components Cb_B.

Step 2: Multiscale Decomposition Parameters

7. Set the number of pyramid levels L = 3.
8. Initialize anisotropic diffusion parameters: number of iterations and gradient threshold.

Step 3: Contourlet-like Multiscale Decomposition

9. Decompose PAN image A into low and high-frequency components.

10. Decompose MS luminance image Y_B into low and high-frequency components.

Step 4: Low-Frequency Fusion

11. Apply Laplacian filtering on A and Y_B.
12. Apply Gaussian smoothing to Laplacian responses.
13. Compute saliency maps for PAN and MS luminance.
14. Normalize saliency maps and apply gamma correction.
15. Construct adaptive weight maps using normalized saliency.
16. Fuse low-frequency components using weighted averaging.

Step 5: High-Frequency Fusion

17. For each decomposition level l = 1 to L:
18. Apply anisotropic diffusion at the first level.
19. Fuse high-frequency subbands using the max-absolute selection rule.

Step 6: Luminance Reconstruction

20. Reconstruct the fused luminance image using inverse multiscale reconstruction.
21. Clip reconstructed luminance to [0,1].

Step 7: Color Reconstruction

22. Combine fused luminance with original chrominance components.
23. Convert the fused YCbCr image to the RGB color space.
24. Clip final RGB image to [0,1].

TABLE I. OBJECTIVE METRICS

Metric	Description	Equation	Interpretation / Role
PSNR	Peak Signal-to-Noise Ratio indicates the quality of reconstruction of the fused image compared to an actual image.	$PSNR = 10\,log_{10}(\frac{MAX_I^2}{MSE})$ where MAX_I The maximum pixel value and MSE is the mean squared error.	An increase in PSNR means there is an increase in the quality of fusion with a reduction in distortion and noise.
SSIM	Structural Similarity Index is a kind of perceptual similarity that is assessed with regard to luminance, contrast, and structure.	$SSIM(F,R) = \frac{(2\mu_F\mu_R + C_1)(2\sigma_{FR} + C_2)}{(\mu_F^2 + \mu_R^2 + C_1)(\sigma_F^2 + \sigma_R^2 + C_2)}$	SSIM values closer to 1 indicate stronger structural similarity and visual fidelity.
Entropy (EN)	Assesses the content of information in the combined picture by using the measurement of randomness.	$H(F) = -\Sigma_{i=0}^{L-1} p_i\, log2(p_i)$ where p_i Is the probability of gray level i.	Higher entropy reflects richer information content in the fused image.
SF	Spatial Frequency is used to quantify the amount of activity or fineness of detail in the combined image.	$SF = \sqrt{(RF^2 + CF^2)}$ Where RF and CF denote row and column frequencies.	Higher SF implies better texture and structural detail preservation.
Corr	The correlation coefficient is a measure used to determine the linear dependency between the fused picture and the source pictures.	$Corr = \frac{\Sigma(F - \mu F)(A - \mu A)}{\sqrt{\Sigma(F - \mu F)^2\Sigma(A - \mu A)^2}}$	Higher correlation indicates better similarity and consistency with source images.
MIF	Mutual Information Fusion is used to measure similar information between the fused image and the source images.	$MIF = MI(F,A) + MI(F,B)$	Higher MIF signifies effective retention of complementary information from source images.
ERGAS	Relative Global Dimensional Error in Synthesis evaluates spectral distortion in fused images.	$ERGAS = 100 \cdot \frac{h}{l}\sqrt{\frac{1}{N}\Sigma_{i=1}^{N}\left(\frac{RMSE_i}{\mu_i}\right)^2}$	Lower ERGAS values indicate better spectral fidelity and reduced distortion.
QABF	Edge-based Fusion Quality Metric assessing edge information transfer from source images to the fused image.	$\frac{Q_{AB}}{F} = \frac{\Sigma Q_0(A,F)\cdot w_A + Q_0(B,F)\cdot w_B}{\Sigma(w_A + w_B)}$	Higher QABF indicates superior edge preservation in the fused image.
LABF	Loss of Edge Information: was a measure of the loss of structuring information during fusion.	$LABF = 1 - ((QABF) + (NABF))$ (edge loss component)	Lower LABF values imply minimal edge information loss and better fusion performance.

A. Datasets

The proposed fusion algorithm was evaluated using publicly open pan sharpening datasets. Dataset-A PAN–MS image pairs obtained from the GitHub repository "Pan Sharpening Dataset" [34]. This dataset provides highly magnified panchromatic images alongside their corresponding multispectral counterparts, which are widely used for pan-sharpening experiments. Dataset-B Additional PAN–MS image pairs were obtained from the repository "Pan sharpening by Convolutional Neural Network" [35].

Dataset-A consists of PAN–MS image pairs obtained from our curated collection, containing high-resolution Panchromatic (PAN) images along with their corresponding Multispectral (MS) counterparts, widely used for pan sharpening experiments. Dataset-B contains more PAN- MS image pairs that represent a variety of natural and urban settings [36] and that are suitable to provide robustness in terms of both evaluation and

generalization with different settings. Dataset-C entails agricultural landscapes, thus, 6 allowing the evaluation of the performance of pan sharpening to be applied in vegetated and textured areas, in which spectral fidelity is essential. The urban infrastructures in Dataset-D and high-density urban areas are the areas of concern where spatial information can be effectively analyzed as per the high-frequency areas, which are represented by roads and buildings.

Dataset-E is a heterogeneous body of PAN-MS image pairs of both natural and urban conditions. It is an effective and extensive standard for assessing the flexibility, resilience, and generalization of the suggested approach in a wide range of environmental and imaging dialogs. The experiments were conducted using contourlet-based multiscale decomposition, anisotropic diffusion filtering, and low–high frequency fusion operations in MATLAB. The implementation utilized the Image Processing Toolbox for pre- processing, coefficient extraction, and evaluation of fusion performance metrics to ensure consistency and reproducibility.

All experiments were executed on a desktop workstation running Windows 11 Pro (Version 23H2, OS Build 22631.4169) with the following configuration: an Intel® Core™ i5 processor @ 2.72 GHz, 16 GB DDR4 RAM, 64-bit operating system on an x64-based architecture, and all computations were carried out without GPU acceleration.

V. Proposed Methodology

The proposed methodology aims at merging the high-spatial resolution data of the Panchromatic (PAN) image with the spectral properties of the Multispectral (MS) image. The algorithms used in the proposed method are contourlet multi-scale decomposition, low-frequency fusion (local variance), and high-frequency fusion (anisotropic diffusion filter) to ensure effective details enhancement and edge preservation. The fused luminance channel $\hat{Y}$ Is then integrated with the original chrominance components (Cb, Cr) from the MS image to form the final fused YCbCr image. The overall process is illustrated in Figure 2.

A. Preprocessing

In preprocessing, a PAN image P and an MS image M are both spatially aligned to the same size in the geometric dimension for further fusion operations. The resizing process is demonstrated as:

$$P_r = resize(P, T), M_r = resize(M, T) \quad (10)$$

Where T = [H, W] is the target size derived from the height H and width W. P_r and M_r are the resampled versions of PAN and MS images. The resizing is essential for performing fusion without spatial distortion. The resampled MS image, M_r, is then converted to the YCbCr colour space, which is denoted as:

$$M_{ycbcr} = RGB2YCbCr(M_r) \quad (11)$$

where M_{ycbcr} is the converted image with three components stated as:

$$M_{ycbcr} = \{Y_B, C_b, C_r\} \quad (12)$$

Where Y_B denotes the luminance channel, and C_bC_r are the chrominance channels.

This conversion derives intensity from colour, allowing the luminance component Y_B to be enhanced or fused with the high-resolution PAN image P_r for enhanced spatial detail while preserving the spectral fidelity carried by C_b and C_r.

B. Multiscale Decomposition

The PAN luminance, P_r, and the MS luminance Y_B are decomposed using a contourlet-based multi-scale residual pyramid to decompose the image across multiple frequency bands. On every level of decomposition, Gaussian blurring is used to extract the low-frequency component, and the standard deviation of the Gaussian filter is defined as:

$$\sigma = 2k \quad (13)$$

This expression gives a direct dependence between the pyramid level and the smoothing scale, stating that the lower levels (small *k*) retain high-frequency detail, whereas higher levels (large *k*) isolate broader structural information. By progressively increasing σ across levels, the multiscale decomposition attains a progressive separation of high and low-frequency data, which is crucial to integrate the PAN spatial data into the MS spectral data, preserving edge integrity as well as minimizing artifacts:

$$L_k = G\sigma(L_{k-1}), H_k = L_{k-1} - L_k \quad (14)$$

where L_0 is the input image and H_k represents the high-frequency residual at scale *k*. This is repeated iteratively over *K* pyramid levels.

C. Low-Frequency Component Fusion

The low-frequency bands, L_P (panchromatic image) and L_Y (multispectral luminance) are combined using a Gaussian-smoothed, saliency-based phase weighting strategy. This increases the local contrast, ensuring the smooth transition in the fused outcome through incremental computation of saliency maps, normalization and gamma correction, with derivation of fusion weights needed further.

1) Saliency-Map Computation: In order to obtain local contrast, both the panchromatic image, A (representing L_P), and the multispectral luminance, Y_B (representing L_Y), are operated using a Laplacian filter defined by the 3x3 kernel:

$$K = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

which is used to compute the high-frequency details as:

$$L_{ap1} = A * K, L_{ap2} = YB * K \quad (15)$$

where $*$ denotes convolution with boundaries replication. To suppress noise and smooth the saliency maps while preserving contrast, a Gaussian filter (15x15 kernel) with standard deviation σ = 2 is applied providing sufficient spatial coverage, noise suppression, and a balance between smoothness and detail, improving visual quality as:

$$P_1 = Lap_1 * G, \quad P_2 = Lap_2 * G \quad (16)$$

where G is the Gaussian kernel. The saliency maps are then

obtained by taking the absolute values as:

$$S_1 = |P_1|, \qquad S_2 = |P_2| \tag{17}$$

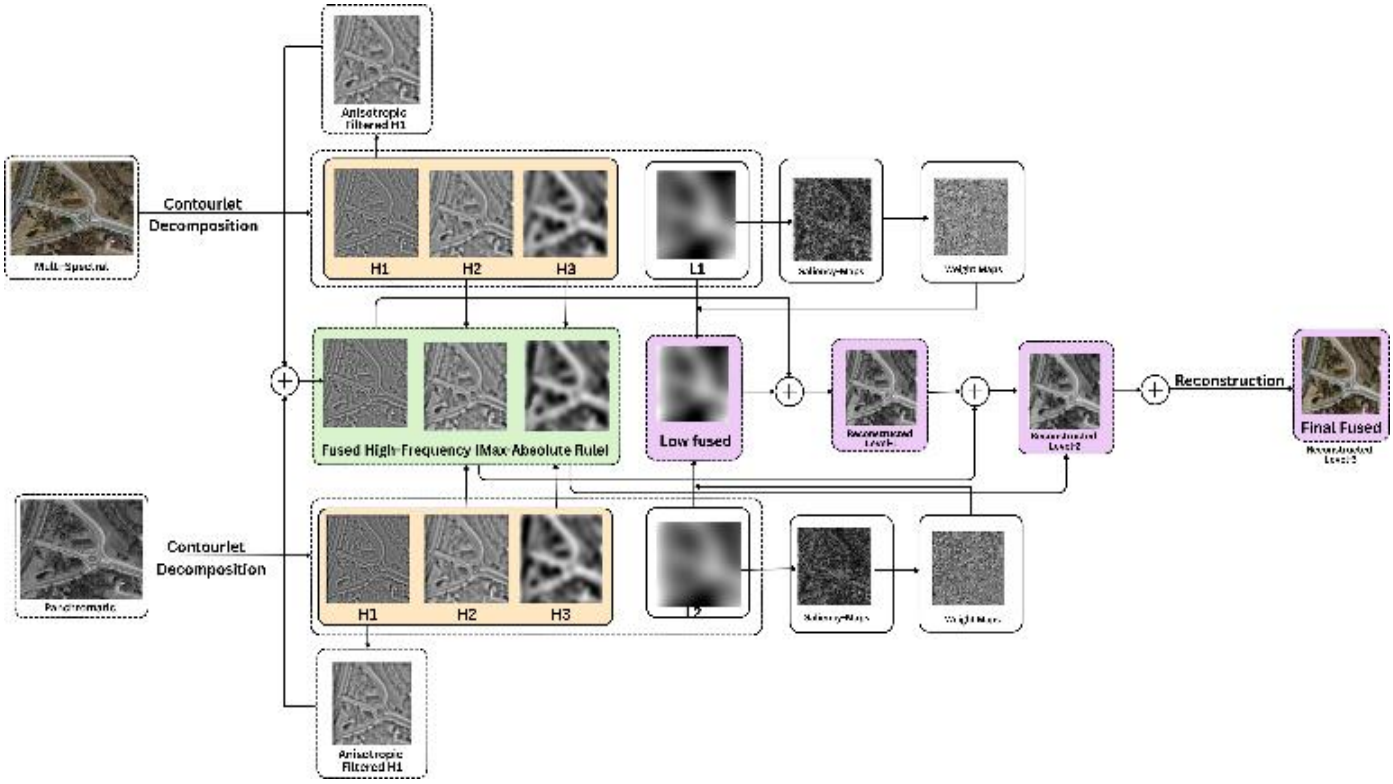

Fig. 2. Proposed Methodology Using Contourlet Decomposition and Anisotropic Diffusion.

2) Normalization and Contrast Enhancement: The saliency maps are normalized to the range [0, 1] using the following equation:

$$S1 = \frac{S1}{\max(S1) + \epsilon}, \quad S2 = \frac{S2}{max(S2) + \epsilon} \tag{18}$$

where S1 and S2 represent the saliency maps of the PAN and MS luminance images, respectively. ϵ = 10−6 is added to prevent division by zero in cases where a saliency map may contain all zero values. After normalization, to enhance local contrast and emphasize intermediate saliency values, a gamma correction with γ = 0.5 is applied as:

$$S_1 = S_1^{0.5}, S_2 = S_2^{0.5} \tag{19}$$

This ensures that the important regions in both images are more prominent, improving the effectiveness of the fusion process.

3) Weight Computation: The normalized, gamma-corrected saliency maps are used to compute the fusion weights as:

$$X_1 = \frac{T1}{T1 + T2 + \epsilon}, \quad X_2 = \frac{T_2}{T1 + T2 + \epsilon} \tag{20}$$

These weights ensure that regions with higher saliency contribute more prominently to the fused result.

4) Fusion: The low-frequency components L_P and L_Y are resized to match the weight map dimensions. The fused low-frequency component L_F is computed as:

$$L_F = W_1 \cdot L_P + W_2 \cdot L_Y \tag{21}$$

The weight maps W1 and W2 are generated from saliency or local contrast measures and have the same spatial dimensions as the original images at a given pyramid level. Since the low-frequency components may have been obtained at a course resolution during multiscale decomposition, resizing ensures that the element-wise multiplication of the weight maps with the low-frequency components are properly defined.

By resizing and performing the weighted fusion, the structural information from both PAN and MS images is accurately combined, while the Gaussian-smoothed weights ensure smooth transitions across regions, preventing misalignment or blocky artifacts in the fused image.

D. High-Frequency Fusion

For high-frequency detail enhancement, an anisotropic diffusion filter is first applied to the input images to sharpen edges while suppressing noise [37]. The process is controlled by the following partial differential equation:

$$\frac{\partial I}{\partial t} = \nabla \cdot (c(\| \nabla I \|)\nabla I) \quad (22)$$

where, I refer to image intensity, the diffusion time or iteration step is denoted by t and ∇ refers to the spatial gradient operator. The conduction coefficient (c(‖∇I‖)) varies the diffusion rate at each pixel with respect to local gradients, preserving sharp edge while smoothing edges homogeneous regions.

The quadratic conduction model is used to define (c(‖∇I‖)) as:

$$c(||\nabla||) = \exp\left(-\left(\frac{||\nabla||}{\mathrm{K}}\right)^2\right) \quad (23)$$

where, the size of the local image gradient is represented by the magnitude of the gradient (where ‖∇I‖ denotes the image) and K = 2 is the gradient threshold controlling diffusion extent such that gradient values higher than K cause less diffusion, preserving the edges' sharpness and suppressing noise in flatter regions. Following the application of the anisotropic diffusion the high-frequency sub bands in both the PAN and MS images are merged by max-absolute selection rule stated as:

$$H_F^{(a)}(y,s) = \begin{cases} H_P^{(a)}(y,s), & \left|H_P^{(a)}(y,s)\right| > H_Y^{(a)}(y,s)|, \\ H_Y^{(a)}(y,s), & otherwise, \end{cases} \quad (24)$$

where $H_P^{(a)}(y,s)$ and $H_Y^{(a)}(y,s)$ are the high-frequency coefficients of the PAN and MS images, respectively, at spatial position (y, s) and subband a, and $H_F^{(a)}(y,s)$ is the fused coefficient.

In the rule, at every pixel, one of the coefficients with the largest value is picked in the fused image, indicating sharper edges or textures while reducing noise.

E. Reconstruction

The fused luminance channel $\hat{Y}$, is reconstructed by fusing the fused low-frequency and high-frequency signals as:

$$\hat{Y} = L_F + \sum_{a=1}^{A} H_F^a \quad (25)$$

where $\hat{Y}$ represents the final fused luminance channel that will later be integrated with the original chrominance components to form the fused YCbCr image. The term L_F denotes the fused low-frequency component, which contains the global structural information and smooth intensity variations from both the PAN and MS images. The summation $\sum_{a=1}^{A} H_F^{(a)}$ aggregates the fused high-frequency sub bands across all decomposition levels a = 1, 2, . . ., A, where $H_F^{(a)}$ represents the high-frequency coefficients at subband a obtained using the max-absolute selection rule. The low-frequency component to the summation of the high-frequency subbands leads to reconstruction of the channel (Y) that almost perfectly preserves the total structure information along with fine details (edges and textures) so that a sharp and coherent fused image can be obtained.

F. Final Image Formation

The fused luminance channel $\hat{Y}$ is combined with the original chrominance components (C_b, C_r) of the multispectral (MS) image to construct the fused YCbCr representation as:

$$I_{ycbcr} = \{\hat{Y}, C_b, C_r\} \quad (26)$$

Then, an inverse colour space conversion is carried out to get the end result RGB fused image by:

$$I_{fused} = YCbCr2RGB(I_{ycbcr}) \quad (27)$$

This reconstruction method is designed in a way that the high spatial resolution details extracted from the PAN image are amalgamated with the MS image without affecting the spectral distortion. The low-frequency, variance-based fusion maintains global structural data and smooth intensity transition, whereas the high-frequency fusion through anisotropic diffusion and max-absolute selection enhances edges and fine textures. Hence, the resulting final fused image I_{fused} attains an optimal balance of spatial resolution and spectral fidelity, resulting in a sharp, visually consistent, and spectrally accurate pansharpened output.

VI. Result Analysis And Discussion

A. Visual result

In Dataset A as illustrated in Figure 3, the proposed algorithm provides the most visually superior fusion output amongst all comparative methods. The proposed methodology demonstrates higher spatially clarity and balanced spectral preservation than FGF-and-XDoG [17], CDIF [18] and CBF [19] which tend to introduce color distortions or excessively enhance local areas. The Structure-aware [20], MDHU [22], IMA [23], and VSM-and-WLS [24] generate grayish results with significant information loss, particularly in fine textural regions. Though, GFDFs [25] and Two-scale [26] methods retain spectral consistency along with some structural information. In contrast, the proposed technique gives a more well-balanced reconstruction, in which structural edges like building edges and vegetation areas are enhanced to have a sharper appearance, more natural contrast, and a better perceptual uniformity, representing better spatial spectral fidelity.

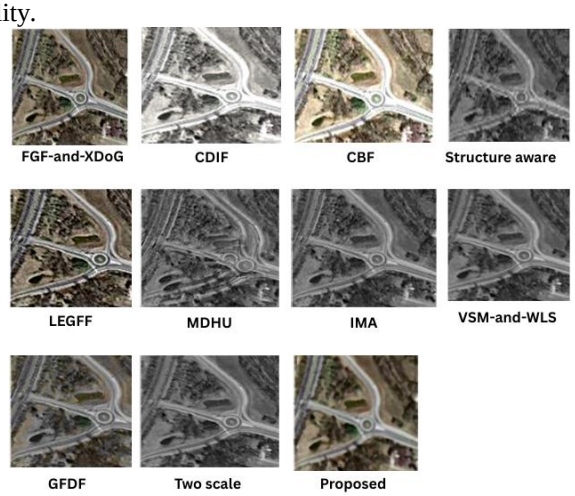

Fig. 3. Fused image of Dataset A with different techniques

In Dataset B presented in Figure 4, the proposed algorithm proves to be obviously superior. The hybrid methods like CDIF and CBF usually yield images that do not match the visual perception, with either over-saturation or under- saturation in selected regions; whereas the structure-oriented methods (e.g., Structure-aware, MDHU, IMA, and VSM-and- WLS) yield images with less vivid color schemes, with most images appear grayish and with low color maintenance. Even though LEGFF [21] has some spectral content, as does GFDF, they bring about blurring in the high-frequency areas. The suggested solution is useful in the sense of combining high- frequency detail with natural spectrum maintenance, which creates more distinct urbanization and a better definition of vegetation textures.

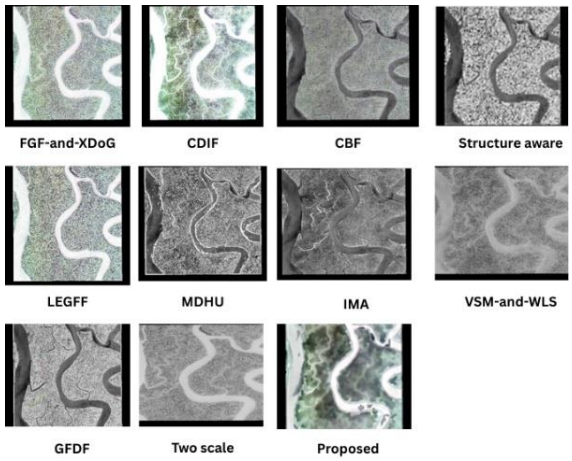

Fig. 4. Fused image of Dataset B with different techniques

Figure 5 The challenge of structural clarity and spectral naturalness. The existing techniques including CDIF and CBF have serious color degradation and artificial improvement. The three models, MDHU, IMA and VSM-and-WLS, are generally producing grayscale results that have a strong loss of spectral information. GFDF and Two-scale maintain structure though fine details are usually blurred and the proposed method is quite effective because it improves regions, edges and preserves the same spectral content.

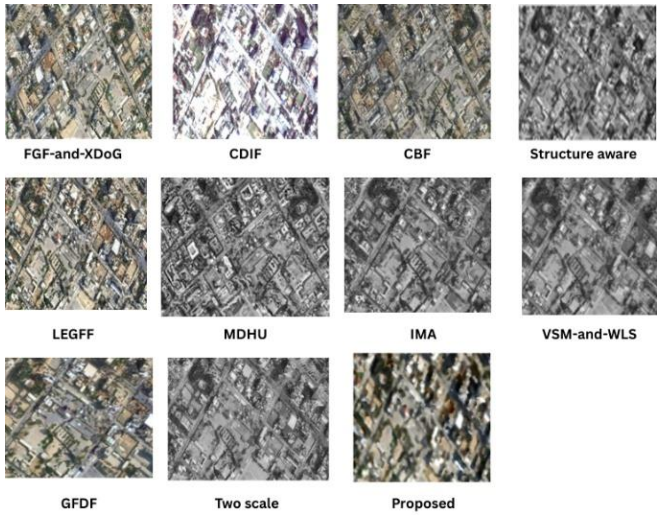

Fig. 5. Fused image of Dataset C with different techniques

In Dataset D of Figure6, with fine linear features like edges heterogenous structures, regions; the vast majority of hybrid approaches do not balance out. FGF-and-XDoG, CDIF, and CBF are distorting spectral content, whereas Structure-aware and MDHU are overdoing edges, making them look unnatural. IMA and VSM-and-WLS have spectrally lost outputs of lower quality. The edges river boundaries are not represented in GFDF and Two-scale. The proposed approach offers a natural appearance of water areas with distinct edges and clear vegetation textures, which is the most realistic of all approaches.

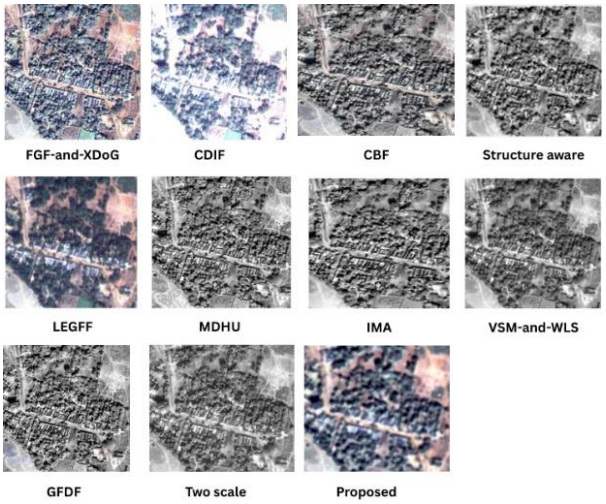

Fig. 6. Fused image of Dataset D with different techniques

Dataset E in Figure7 shows that the proposed algorithm is effective in Dataset E as is evidenced by Figure3. Competing fusion algorithms have several weaknesses - such algorithms as FGF-and-XDoG, CDIF and CBF will tend to generate unnaturally enhanced contrast, the algorithms like Structure-aware, MDHU, IMA and VSM-and-WLS are not able to reproduce colors correctly and the spectral distortion becomes apparent. LEGFF, as well as GFDF can preserve partial high-frequency detail, but with low global sharpness and Two-scale method results in blurred fine details because of inadequate edge reinforcement.

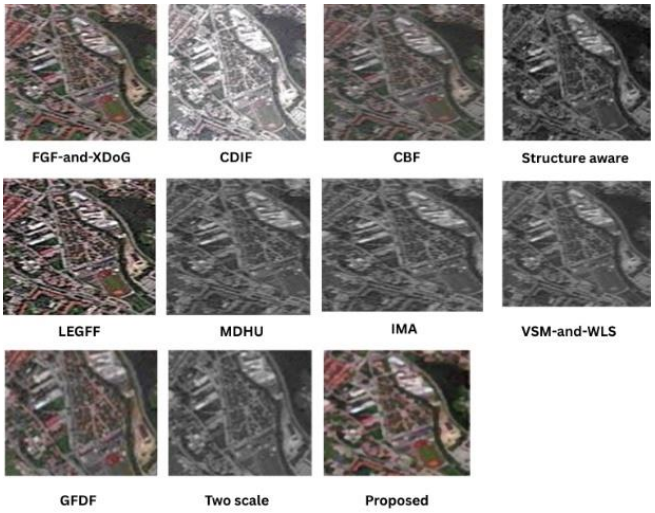

Fig. 7. Fused image of Dataset E with different techniques

Conversely, the proposed algorithm is also able to maintain spectral fidelity as well as spatial sharpness and produce a fused output, which is visually balanced and natural. The different features like road networks, vegetation textures, man-made structures are also highly maintained with few artifacts, which reflects the efficiency and flexibility of the method applied in challenging scenes.

B. Objective Results

As a way to thoroughly evaluate the fusion performance in Table II, various objective measurements were compared among various hybrid approaches including FGF-and-XDoG, CDIF, CBF, Structure-aware, LEGFF, MDHU, and IMA as well as VSM-and-WLS, GFDF, Two-scale, and the proposed one. The findings categorically show that the approach has better results in terms of the structural preservation, as the highest QABF score (0.6868) reflects. That proves its ability to preserve edges and fine details of the structure. The competitive performance in PSNR (22.91 dB) and SSIM (0.6995) is also shown by the method, indicating an equal trade-off of spatial fidelity and perceptual quality. CBF had the best PSNR (26.56 dB), SSIM (0.881) but it has a poorer QABF (0.5135), indicating that it has the propensity to over smooth edges at the expense of edge clarity. Equally, Structure-aware and VSM-and-WLS methods were relatively high (25.60 dB and 25.89 dB, respectively) but the QABF scores (0.5155 and 0.5212, respectively) were significantly lower than in the proposed method. Information-theoretically, the given methodology delivered a competitive entropy (7.2397) and an average MIF score (1.4878), which means that enough informational richness was not discarded.

TABLE II. COMPARISON OF DIFFERENT METHODS ON DATASET A

Metric	FGF-and-XDoG [23]	CDIF [24]	CBF [25]	Structure-aware [26]	LEGFF [27]	MDHU [28]	IMA [29]	VSM-and-WLS [30]	GFDF [31]	Two-scale [32]	Proposed
PSNR	22.471	10.5557	26.5626	25.6	19.2802	19.045	21.6114	25.8953	23.8754	26.5642	22.9102
SSIM	0.7981	0.7386	0.881	0.8677	0.6738	0.5987	0.6332	0.8676	0.7718	0.8742	0.6995
Entropy	7.453	7.4427	7.2244	7.1398	7.6103	7.2424	7.1916	7.14	7.243	7.164	7.2397
SF	0.1489	0.1492	0.1146	0.0948	0.2119	0.1445	0.1019	0.0978	0.1191	0.0968	0.0355
Corr	0.9179	0.9769	0.9608	0.9432	0.9025	0.7343	0.8684	0.9423	0.91	0.9548	0.8915
MIF	2.1528	0.6586	0.6481	2.2624	2.0055	2.0917	1.839	2.3459	2.2989	2.3035	1.4878
ERGAS	20.185	20.2828	3.2901	14.0833	28.3771	29.438	23.3795	3.4023	17.1724	12.6015	19.9439
QABF	0.4666	0.4744	0.5135	0.5155	0.5047	0.669	0.3162	0.5212	0.3929	0.5759	0.6868
LABF	0.5334	0.5168	0.4785	0.4754	0.4953	0.331	0.6838	0.4788	0.6071	0.4241	0.3132

Despite the fact that other approaches like GFDF and CBF performed better in terms of ERGAS (12.60 and 3.29, respectively), the proposed approach showed a high level of robustness in a broader range of metrics, especially the feature-based ones (QABF, LABF) whereby perceptual sharpness and structural consistency are the most important. On the whole, the experimental data supports the idea that the proposed method has an optimal balance in spectral fidelity, enhancement of spatial detail, and preservation of edges.

The approach shows excellent results on both quantitative and perceptual visual quality by effectively incorporating high-frequency spatial information with uniform spectral properties as compared to the current hybrid approaches. Objective measurements were analyzed in FGF-and- XDoG, CDIF, CBF, Structure-aware, LEGFF, MDHU, IMA, VSM-and-WLS, GFDF, Two-scale, and the proposed approach using the quantitative assessment of the shown performance in Table III.

The findings point to the fact that the suggested method achieves significant gains in mutual information retention, and the highest MIF index (1.6952) is reached. This indicates its good capability of maintaining complementary information of source images. Moreover, the approach displays an equal structural consistency, as the score of QABF is 0.5009, which is higher than a variety of more advanced options like LEGFF (0.2044) and GFDF (0.1232). To ensure preservation of details, the CBF algorithm pro- vided the best PSNR (18.89 dB) and SSIM (0.8729), which shows that all the edges are preserved and the structure is similar. Nevertheless, the relatively low values of the MIF and QABF indicate a balance between the spatial sharpness and the general structure of information. The same trend was noted in Structure-aware fusion approach, which presented significantly less good QABF results (0.4747) but much better PSNR (13.10 dB), indicating that the structural details were preserved, whereas the perceptual and spectral fidelity was less than ideal.

Comparatively, the proposed algorithm has provided a less competitive PSNR (13.07 dB) but with stronger information-theoretic and edge-based metrics. Entropy analysis also points to the fact that the proposed scheme (6.6539) has retained the higher informational content than CDIF (5.1806) and LEGFF (6.2467), yet this is closely related to IMA (7.0732), which returned the highest entropy. Even though the ERGAS values indicate that CBF (5.97) and VSM-and-WLS (12.52) were superior in spectral distortion, the suggested technique shows better performance on feature-sensitive ones (MIF, QABF), which are essential to retain perceptual sharpness.

TABLE III. COMPARISON OF DIFFERENT METHODS ON DATASET B

Metric	FGF-and-XDoG	CDIF	CBF	Structure-aware	LEGFF	MDHU	IMA	VSM-and-WLS	GFDF	Two scale	Proposed
PSNR	9.7821	7.8872	18.8974	13.1	10.3835	8.5178	12.9253	13.096	7.1954	12.4779	13.0793
SSIM	0.4085	0.7737	0.8729	0.6121	0.2896	0.2461	0.6028	0.6121	0.2105	0.5607	0.5834
Entropy	6.3929	5.1806	6.9538	6.4678	6.2467	6.9516	7.0732	6.4668	6.9885	6.4843	6.6539
SF	0.3748	0.2962	0.2826	0.2117	0.4577	0.3889	0.2628	0.2116	0.3557	0.2415	0.1066
Corr	0.7198	0.8815	0.9502	0.7738	0.7372	0.3793	0.8063	0.7738	0.176	0.7739	0.8466
MIF	1.1183	0.3693	0.6699	1.4055	0.9857	0.9656	1.3303	1.3887	1.1128	1.4615	1.6952
ERGAS	73.351	23.0681	5.9741	50.0874	68.3918	84.847	47.5272	12.5214	98.8008	53.7816	50.8013
QABF	0.1404	0.3771	0.456	0.4747	0.2044	0.2217	0.5992	0.2252	0.1232	0.1464	0.5009
LABF	0.8596	0.6136	0.5359	0.518	0.7956	0.7783	0.4008	0.7748	0.8768	0.8536	0.4991

On the whole, the comparative analysis proves that the suggested approach can guarantee an efficient trade-off between spectral fidelity, preservation of structural detail, and the preservation of mutual information, thus proving the reliability of this approach as a powerful fusion strategy in a variety of evaluation measures.

The comparative findings in Table IV offer a thorough evaluation of various fusion methods in terms of structural, spectral and information-theoretic measures. The GFDF method had the highest performance in all comparative methods of edges preservation, with the highest score of 0.9426 in QABF, indicating its high performance in terms of retaining all the structural information and preserving all structural edges. It also achieved the highest Mutual Information (MIF = 4.3397), which is effective in retaining complementary and informative material between images.

TABLE IV: COMPARISON OF DIFFERENT METHODS ON DATASET C

Metric	FGF-and-XDoG	CDIF	CBF	Structure-aware	LEGFF	MDHU	IMA	VSM-and-WLS	GFDF	Two scale	Proposed
PSNR	16.897	8.7071	18.238	17.4	16.7675	15.253	12.9253	17.4087	19.8215	18.2746	14.5665
SSIM	0.7348	0.7284	0.7205	0.6635	0.6453	0.5444	0.6028	0.6604	0.8638	0.715	0.5466
Entropy	7.5983	6.5278	7.4268	7.3086	7.7232	7.6553	7.0732	7.3158	7.5376	7.3858	7.7423
SF	0.1154	0.096	0.1089	0.0694	0.1354	0.1445	0.2628	0.0679	0.1095	0.0821	0.1522
Corr	0.7996	0.9173	0.7946	0.7665	0.7729	0.6332	0.8063	0.7657	0.8663	0.7985	0.6316
MIF	1.2208	0.2775	0.7153	0.8658	0.9274	1.4679	1.3303	0.8721	4.3397	1.5048	1.0584
ERGAS	28.708	21.2809	6.263	27.1042	29.5858	34.729	47.5272	6.7659	20.4998	24.4963	38.4077
QABF	0.8436	0.46	0.5231	0.56	0.7857	0.7548	0.5992	0.7322	0.9426	0.7949	0.2871
LABF	0.1564	0.5208	0.4575	0.4211	0.2143	0.2453	0.4008	0.2678	0.0574	0.2051	0.7129

Nevertheless, along with such merits, GFDF showed moderate PSNR (19.82 dB) and rather high ERGAS (20.49) implying shortcomings in spectral fidelity. However, CBF ensured the balance of high PSNR (18.23 dB) with high SSIM (0.7205), as well as competitive entropy (7.4268), showing good noise suppression and spectral consistency.

LEGFF and MDHU achieved rather high entropy (7.7423 and 7.65, respectively), yet their SSIM score (0.6453 and 0.5444) was significantly smaller, which implies less accurate perceptual similarity. Nevertheless, the proposed approach was not the best in PSNR (14.56 dB) and SSIM (0.5466) but showed a high level of entropy (7.7423) and spatial frequency (0.1522), which allows to infer a Nevertheless, the value of its QABF was rather low (0.2871) indicating that edge preservation was weakly better than in the case of GFDF and LEGFF. On the whole, Dataset C indicates the trade-offs of various fusion strategies. Whereas GDDF has an edge in edge-based and information-preservation measures, CBF and Structure-aware have a more balanced fidelity in PSNR and SSIM. The suggested approach focuses on the increase of entropy and detail, which possess mutually complementary advantages in situations where the value of information is valued more highly than the edge-based quality.

The evaluation presented in Table V highlights the relative performance of competing fusion methods across multiple metrics. The proposed approach attains a PSNR of 18.9827, which is comparable to GFDF (18.5479) and IMA (19.6891), and notably superior to methods such as CBF (14.6687) and MDHU (14.3530). However, Structure-aware fusion achieves the highest PSNR of 29.1947, setting a strong upper bound on

reconstruction fidelity. In terms of structural similarity, the proposed method reaches an SSIM of 0.7349, outperforming most conventional baselines such as CBF (0.3656) and MDHU (0.2972). Only LEGFF (0.9740) and Structure-aware (0.6146) demonstrate stronger structural preservation. Entropy analysis indicates that the proposed technique (7.8488) effectively preserves information content, second only to MDHU (7.8677), and higher than most alternatives, including Structure-aware (7.6018).

TABLE V. COMPARISON OF DIFFERENT METHODS ON DATASET D

Metric	FGF-and-XDoG	CDIF	CBF	Structure-aware	LEGFF	MDHU	IMA	VSM-and-WLS	GFDF	Two scale	Proposed
PSNR	14.216	7.3865	14.6687	19.69	29.1947	14.353	14.2398	19.6891	13.6682	18.5479	18.9827
SSIM	0.388	0.7153	0.3656	0.6146	0.974	0.2972	0.3452	0.6146	0.3007	0.6019	0.7349
Entropy	7.7323	5.8041	7.8222	7.702	7.6018	7.7617	7.8677	7.7123	7.6033	7.6649	7.8488
SF	0.1463	0.0862	0.1307	0.0855	0.081	0.155	0.1367	0.0855	0.1433	0.0911	0.0659
Corr	0.7263	0.876	0.6918	0.8847	0.9994	0.6665	0.6834	0.8847	0.612	0.8578	0.8788
MIF	0.8576	0.1922	0.0766	1.4933	4.7855	0.7027	0.8503	1.4848	0.6092	1.4532	1.6489
ERGAS	39.035	20.9846	9.086	20.7866	7.0666	38.422	38.149	5.1965	41.5733	23.7046	22.1051
QABF	0.1194	0.5451	0.6667	0.6778	0.9229	0.0696	0.0643	0.2809	0.0525	0.2249	0.5754
LABF	0.8806	0.4432	0.322	0.3113	0.0771	0.9305	0.9357	0.7191	0.9475	0.7751	0.4246

The comparison in Table VI shows the comparative performance of competing fusion strategies on various metrics. The proposed method achieves a PSNR of 18.9827 that is similar to GFDF (18.5479) and IMA (19.6891), but significantly higher than other methods, e.g., CBF (14.6687) and MDHU (14.3530). Nevertheless, Structure-aware fusion obtains the best PSNR of 29.1947 which is a powerful upper limit of reconstruction fidelity. The proposed method achieves a structural similarity of 0.7349, compared to most of the conventional baselines, e.g. CBF (0.3656) and MDHU (0.2972). It is only LEGFF (0.9740), and Structure-aware (0.6146) that exhibit more structural preservation. The entropy analysis has shown that the proposed method (7.8488) is effective in terms of information content (only MDHU (7.8677) does better) and better than most other methods, including Structure-aware (7.6018). The indicators of sharpness and the activity level show otherwise. Spatial frequency (SF) of the suggested technique is 0.0659, which is less than such techniques as MDHU (0.1550) and GFDF (0.1433), indicating that edge sharpness is not excessive. Equally, the correlation coefficient (0.8788) is also competitive, with its value just below LEGFF (4.7855), but much higher than standard baselines, including CBF (0.0766) and MDHU (0.7027). Mutual information analysis (MIF) indicates that the method proposed is 1.6489, which is very high, compared to the traditional baselines, including CBF (0.0766) and MDHU (0.7027). These results are further contextualized by the ERGAS of 22.1051, which is better spectral fidelity than other high-error methods such as the GFDF (41.5733) and MDHU (38.4220), but also IMA (5.1965) and LEGFF (7.0666) are even stronger. The given method shows the QABF score of 0.5754, which proves the obvious benefits when compared to the weak performers like MDHU (0.0643) and GFDF (0.0525). However, Structure-aware (0.9229) provides much higher boundary fidelity. The LABF of 0.4246 proves a medium-level stability to retain local edge structures, and the stability is lower than both GFDF (0.7751) and IMA (0.7191).However, the overall performance in Dataset D shows that the proposed method is a balance between similarity in structures and content as well as the moderate level of correlation and boundary fidelity, but it is outperformed by the dominant methods (LEGFF and Structure-aware).

Table VI summarize the comparative manner in which fusion algorithms behave under Dataset E. The given algorithm scores 24.9394 in PSNR, which is one of the more successful algorithms and surpasses the results of FGF-and-XDoG (24.3650), LEGFF (18.1240), and IMA (21.3798). However, the greatest PSNR values are achieved through GFDF (40.6560), MDHU (29.3050) and CBF (26.8643) that means that they are more effective in terms of structural preservation compared to IMA (0.5257) and LEGFF (0.6163). Structural similarity SSISM (0.7209) shows that the proposed approach preserves the structure moderately, better than IMA (0.5257) and LEGFF (0. The values of entropy indicate that the proposed technique (6.7194) has similar levels of information richness as CBF (6.7278) and Structure-aware (6.6890), but lower as compared to CDIF (7.3855) and LEGFF (7.5712).

As demonstrated in Table VI, the SF of the proposed method (0.0291) is much lower than other techniques like LEGFF (0.2502), CDIF (0.1841), and GFDF (0.1218) showing relatively weak edges and details. The same observation can be made on the correlation coefficient where the proposed method performs at 0.8798, which is satisfactory yet not as high as high performing baselines as GFDF (0.9971), CDIF (0.9778) and MDHU (0.9589). Mutual information analysis indicates moderate performance with the proposed method recording 1.3471, which is lower compared to CBF (0.6076) and LEGFF (1.7909). ERGAS Score of 15.9246 shows that spectral fidelity is acceptable, which is better than other methods with high errors (LEGFF (33.2816) and IMA (23.8915)) but worse than methods with high performance (GFDF (0.8880), VSM-and- WLS (3.1720), and CBF (3.1792)).

The proposed method features a QABF of 0.5820, which is better than IMA (0.15 The results of the LABF score (0.4180) indicates moderate local boundary fidelity, which is equal to

Structure-aware (0.4775) and CBF (0.4786), and worse than IMA (0.8459).In general, Dataset E results demonstrate that the proposed algorithm provides a balanced trade-off, preserving competitive PSNR, entropy, and edge fidelity scores, but has lower sharpness, correlation, and mutual information values than the hybrid baselines, such as GFDF and MDHU.

TABLE VI. COMPARISON OF DIFFERENT METHODS ON DATASET E

Metric	FGF-and-XDoG	CDIF	CBF	Structure-aware	LEGFF	MDHU	IMA	VSM-and-WLS	GFDF	Two scale	Proposed
PSNR	24.365	10.9817	26.8643	26.77	18.124	29.305	21.3798	26.7781	40.656	27.0353	24.9394
SSIM	0.8762	0.7683	0.8622	0.8485	0.6163	0.9107	0.5257	0.8485	0.9913	0.8535	0.7209
Entropy	7.1667	7.3855	6.7278	6.689	7.5712	6.8367	6.8537	6.7027	6.8938	6.708	6.7194
SF	0.144	0.1841	0.0967	0.0932	0.2502	0.1145	0.1148	0.0928	0.1218	0.0938	0.0291
Corr	0.932	0.9778	0.9378	0.9274	0.8677	0.9589	0.7507	0.9274	0.9971	0.9316	0.8798
MIF	2.1115	0.6385	0.6076	1.957	1.7909	3.5275	1.4354	2.0011	5.986	1.9913	1.3471
ERGAS	16.751	19.7583	3.1792	12.6928	33.2816	9.4842	23.8915	3.172	2.5671	12.3165	15.9246
QABF	0.6525	0.4859	0.5113	0.5125	0.5365	0.782	0.1541	0.5117	0.888	0.544	0.582
LABF	0.3475	0.5043	0.4786	0.4775	0.4635	0.218	0.8459	0.4883	0.112	0.456	0.418

VII. OPTIMIZATION OF PROPOSED ALGORITHM

For all optimization experiments of Dataset A , each image was decomposed into three pyramid levels which were similar to the contourlet structure. In the case of anisotropic diffusion, gradient threshold was (aniso_grad) set to 2. The guided filtering stage base layer radius (r1_default) was kept constant at 45 whereas regularization parameter (eps1) was established at 0.3. These parameter parameters were fixed so that they provided consistency over all optimization runs.

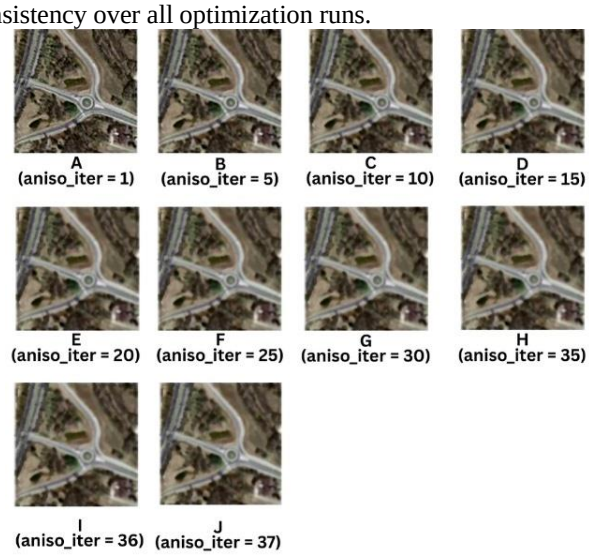

Fig. 8. Fusion quality variation across anisotropic diffusion iterations (Dataset A).

Anisotropic diffusion iteration count (aniso_iter) was analyzed to determine the effect of anisotropic diffusion on fusion quality in Dataset A in a systematic manner in terms of the QABF measure. The values of QABF grow steadily as 0.5231 at aniso_iter = 1 and high as 0.6868 at aniso_iter = 35, with an absolute change of 0.1637 (≈ 31.3%). This steady rise indicates that increasing the number of iterations improves the level of structural preservation because QABF lays stress on the quality of fine details and retention of edges as shown in Table VII and graphically in Figure 9.

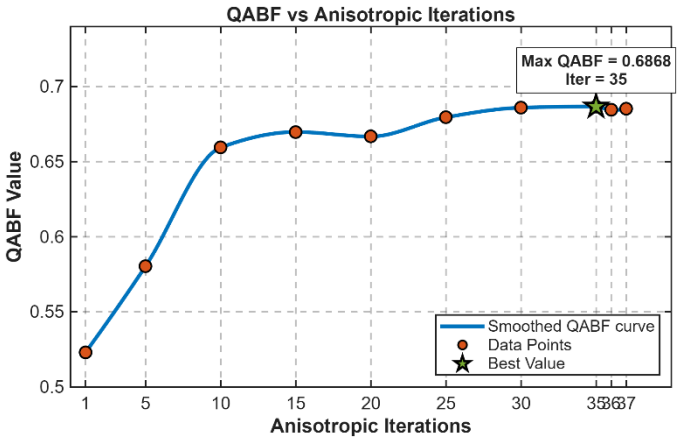

Fig. 9. Fusion quality variation across anisotropic diffusion iterations (Dataset A).

Nonetheless, there is a visual discrepancy between the metric performance and the quality of perception. The peak QABF occurs at an iterative value of aniso_iter = 35 and thereafter, the fused image has slight over accentuation of edges and halo artifacts which are not desired when considering the perceptual view point. However, the lower values of iteration like aniso_iter = 10 (QABF = 0.6595) or aniso_iter = 15 (QABF = 0.6698) produce much more natural and balanced images although they have slightly lower values of the QABF. This demonstrates that the structure-oriented QABF metric can effectively capture and quantify stronger edge contrast, even in scenarios where slight compromises in visual fidelity occur.

It is found that a useful trade-off to be explored is at aniso_iter values of 25 to 30 whereby QABF can be as large as possible (0.6796-0.6860) and the reconstructions are artifact free and appealing. Thus, we choose aniso30 as the most appropriate setting for this work that offers the combination of both objective preservation of edges and subjective picture quality. In all the optimized experiments of Dataset B as indicated in Figure

10, a multi-scale analysis using nine pyramid levels was used, based on a contourlet-based framework.

TABLE VII. OPTIMIZED OBJECTIVE RESULTS OF DATASET A PENDING

Metric	(aniso_iter=1)	(aniso_iter=5)	(aniso_iter=10)	(aniso_iter=15)	(aniso_iter=20)	(aniso_iter=25)	(aniso_iter=30)	(aniso_iter=35)	(aniso_iter=36)	(aniso_iter=37)
PSNR	23.498	23.4056	23.1625	23.0467	22.9798	22.9439	22.9208	22.9102	22.909	22.9081
SSIM	0.8131	0.7608	0.7278	0.714	0.7068	0.703	0.7007	0.6995	0.6993	0.6991
Entropy	7.4465	7.3524	7.3067	7.2818	7.2657	7.2544	7.2461	7.2397	7.2385	7.2373
SF	0.0723	0.0496	0.0425	0.0394	0.0378	0.0367	0.036	0.0355	0.0354	0.0353
Corr	0.9238	0.9116	0.9023	0.8977	0.895	0.8933	0.8922	0.8915	0.8914	0.8913
MIF	1.8254	1.6399	1.5613	1.5233	1.5071	1.4981	1.4901	1.4878	1.4873	1.4852
ERGAS	18.6292	18.8361	19.3719	19.6323	19.7843	19.8666	19.9195	19.9439	19.9467	19.9487
QABF	0.5231	0.5805	0.6595	0.6698	0.6668	0.6796	0.686	0.6868	0.6845	0.6852
LABF	0.4769	0.4195	0.3405	0.3302	0.3332	0.3204	0.314	0.3132	0.3155	0.3148

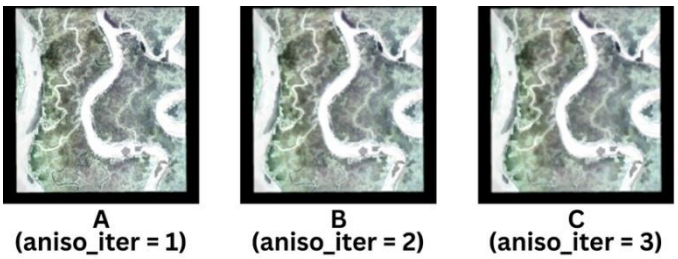

Fig. 10. Fusion quality variation across anisotropic diffusion iterations (Dataset B).

A gradient threshold of 10 was used to control the anisotropic diffusion process (computer (aniso_grad)). The guided filtering phase had a base layer radius, which was set at 45 (r1_default) and the regularization term, which was kept at 0.3 (eps1). These parameter choices were consistently applied across all optimization trials to ensure fair comparison. The effect of anisotropic diffusion iteration count (aniso_iter) on fusion quality in Dataset B shown in Table VIII was also analyzed using the QABF metric.

As shown in Figure 11 also, QABF values decrease slightly from 0.5009 at aniso_iter = 1 to 0.4921 at aniso_iter = 2 and 0.4827 at aniso_iter = 3. Although the highest QABF is obtained at aniso_iter = 1, the fused image appears relatively noisy and lacks smoothness in homogeneous regions. The parameter selections were uniformly used in all optimization experiments, so that the effect of anisotropic diffusion by iteration count on fusion quality inside Dataset B that is presented in Table VIII could be analyzed with the help of the QABF measure. As Figure 11 below also demonstrates, values of QABF change slightly with an increase of 0.5009 at the beginning of the iteration process (that is at aniso_iter = 1) to 0.4921 at the end of the iteration process aniso_iter = 2 and 0.4827 at the end of the iteration process aniso_iter = 3. The optimal QABF is achieved at an aniso before the value of aniso iter reaches 1 but the fused image is rather noisy and does not have smoothness in homogeneous areas.

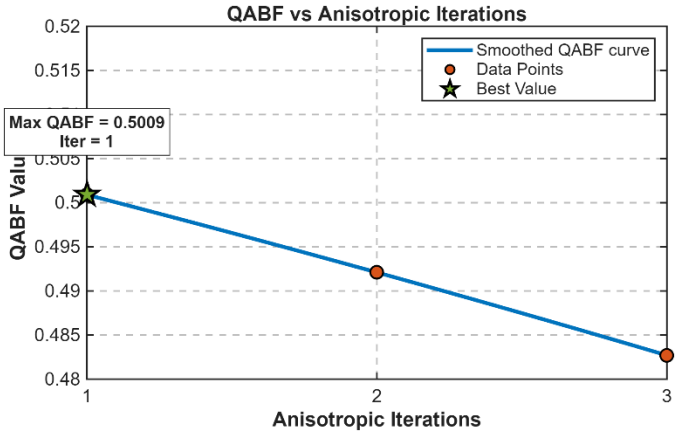

Fig. 11. Fusion quality variation across anisotropic diffusion iterations (Dataset B).

After increasing the number of iterations, the QABF score goes down slightly, but the visual quality is enhanced by the fact that noise is suppressed and the natural textures are better preserved.

TABLE VIII. OPTIMIZED OBJECTIVE RESULTS OF DATASET B

Metric	(aniso_iter = 1)	(aniso_iter = 2)	(aniso_iter = 3)
PSNR	13.0793	13.1571	13.1785
SSIM	0.5834	0.5808	0.5729
Entropy	6.6539	6.6786	6.683
SF	0.1066	0.0887	0.0806
Corr	0.8466	0.8479	0.8474
MIF	1.6952	1.7012	1.7019
ERGAS	50.8013	50.3478	50.2236
QABF	0.5009	0.4921	0.4827
LABF	0.4991	0.5079	0.5173

This once again demonstrates the difference between quantitative evaluation and perceptual one: although in QABF higher contrast of edges is desirable, noise or exaggerated textures are not graded. In this data set, there is a useful compromise at aniso_iter = 2 or 3 that generate more pleasing and balanced images even though they have lower scores on QABF.

The Dataset C results in Figure 12 one can be optimized experiments, the two-level pyramid decomposition used in the fusion was contourlet-based designed.

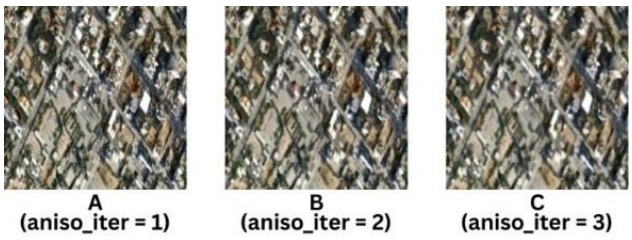

Fig. 12. Fusion quality variation across anisotropic diffusion iterations (Dataset C).

With a fairly low gradient threshold (computing: aniso_grad = 0.2), anisotropic diffusion was used, which put an emphasis on fine detail keeping. In the case of guided filtering, the radius of the base layer (r1_default) was fixed at 45 and the regularization parameter (eps1) was fixed to 0.3. With these constant settings, the results of the optimization could be directly compared between all the trials presented in Table VIII.

The effect of anisotropic diffusion iteration count (aniso_iter) on fusion quality in Dataset C was further examined using the QABF metric. As shown in Figure 13, QABF values decrease consistently from 0.2871 at aniso_iter = 1 to 0.2691 at aniso_iter = 2 and 0.2589 at aniso_iter = 3. Although the highest QABF is observed at aniso_iter = 1, the corresponding fused image suffers from noise and lacks smoothness in homogeneous regions.

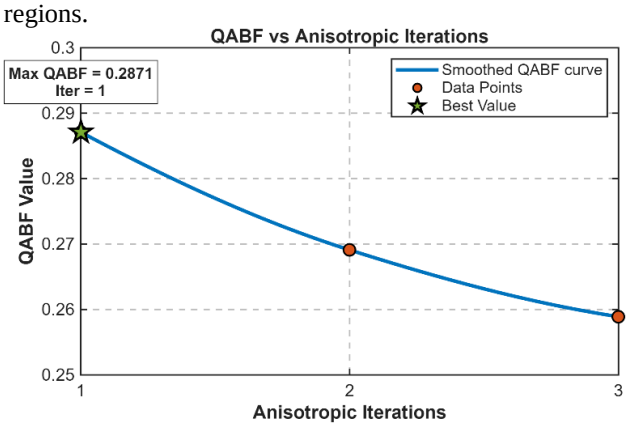

Fig. 13. Fusion quality variation across anisotropic diffusion iterations (Dataset C).

The impact of the number of counts of anisotropic diffusion iteration on the quality of fusion of Dataset C was also investigated aided by the concept of the QABF. Figure 13 of the appendix shows that QABFs drop steadily with an increase in aniso_iter, e.g., 0.2871 at aniso_iter = 1, 0.2691 at aniso_iter = 2 and 0.2589 at aniso_iter = 3. Even though the maximization of QABF occurs at a value of aniso_iter = 1, the resultant fused image has noise and is not smooth in homogeneous areas.

As the number of iterations increases, the QABF score reduces a little, but there is an increase in perceptual quality of the images as noise is reduced and the aesthetic look of the images becomes more ordinary. This again illustrates the difference between the quantitative assessment and the human eye view: on the one hand, QABF is concerned with edge retention, whereas on the other hand, it is not sufficient to represent the degradation due to noise. Hence in this dataset, the value of aniso_iter = 2 or 3 has a more perceptually acceptable result although it has smaller value of QABF. In the optimized experiments of Dataset D reported in Figure 14 Table IX, the contourlet-like structure was adhered to and a six- level pyramid decomposition was used.

TABLE IX. OPTIMIZED OBJECTIVE RESULTS OF DATASET C.

Metric	A (aniso_iter = 1)	B (aniso_iter = 2)	C (aniso_iter = 3)
PSNR	14.5665	14.6693	14.7223
SSIM	0.5466	0.5276	0.5085
Entropy	7.7423	7.7088	7.6811
SF	0.1522	0.1276	0.112
Corr	0.6316	0.6226	0.6144
MIF	1.0584	1.0347	1.0173
ERGAS	38.4077	37.9539	37.722
QABF	0.2871	0.2691	0.2589
LABF	0.7129	0.7309	0.7411

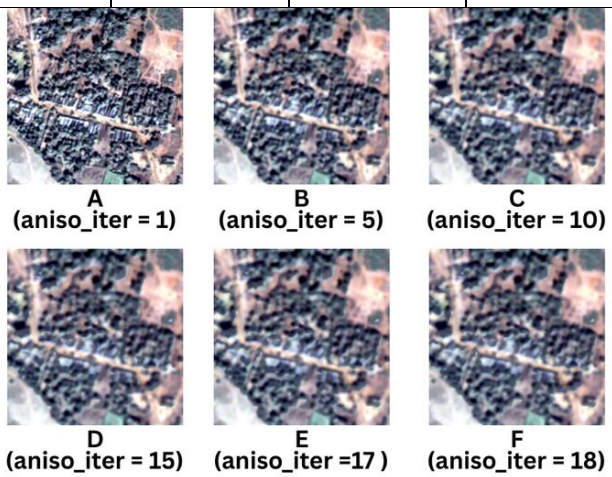

Fig. 14. Fusion quality variation across anisotropic diffusion iterations (Dataset D).

A gradient threshold of 2 (aniso_grad = 2) was used to represent the level of anisotropic diffusion, which offers the desired edge saving and noise reduction. At the guided filtering phase, the following parameters were defined based on the low-frequency soettings of the GFF approach; the base layer radius was kept to 45; in ther words, the parameter (r1_default) was default, and the regularization term was (eps1) regularization of 0.3. The consistency of these parameter values allowed the optimization results across all of the runs to be stable and reproducible.

Anisotropic diffusion iteration count (aniso_iter) affects fusion quality in Dataset D was also explored on the basis of the measurement of Quality of BFS (QABF) measure. Figure 14 indicates that the values of QABF rise significantly with the increase in 0.2747 at aniso_iter = 1 to a nabf value 0.4774 at aniso_iter = 5 and the value goes on increasing to 0.5466 at aniso iter = 10. It has its highest at the first instance of aniso_iter = 16 and QABF = 0.5800, then slightly decreasing to 0.5777 at aniso_iter = 17. The fact that this enhancement with additional iterations shows that anisotropic diffusion will work well in maintaining edges and structure to a specific extent.

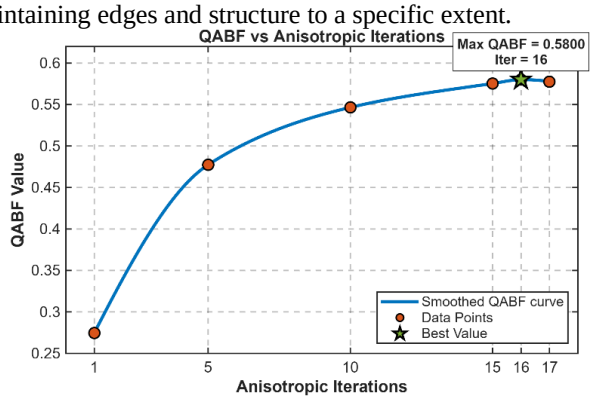

Fig. 15. Fusion quality variation across anisotropic diffusion iterations (Dataset D).

But on visual examination as depicted in Fig 15, one finds that the highest QABF is reached when the value of aniso_iter is set to 16, but one sees that there is a slight over-enhancing of edge in the fused image. Middle values, like aniso_iter = 10 or 15 of 1 /T/aniso (QABF = 0.5466 and 0.5754) have more visually balanced results with less artifacts, though. This once again illustrates the tradeoff between objective measures and perceived quality, where this time, with an aniso_iter = 15 gives a viable tradeoff between almost highest QABF and high visual quality. Dataset E In the optimized experiments in Figure 16, the input images were uniformly resized. It was decomposed into a two-level pyramid, which was based on a contourlet-like representation.

Fig. 16. Fusion quality variation across anisotropic diffusion iterations (Dataset E).

Anisotropic diffusion process controlled by a gradient threshold of 2 (aniso_grad = 2) was used to prevent the loss of structural detail but reduce excessive over-smoothing as in Table X. In the guided filtering step, the parameter radius of the base layer (which is set as r1 default comes out as 45 by default), and the optimized parameter (which is set as eps1 remains at 0.3 by default) were kept constant. Such parameter adjustments gave a uniform experimental starting point to assessing fusion performance as indicated in Table XI.

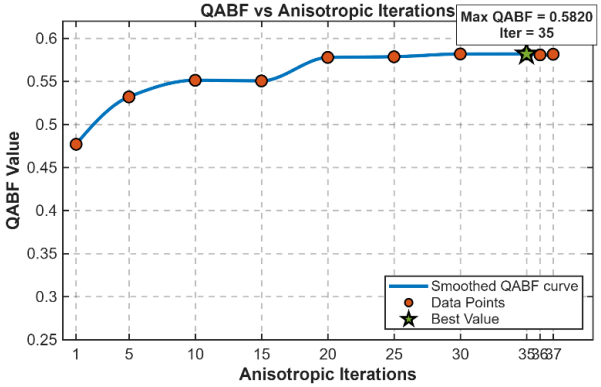

Fig. 17. Fusion quality variation across anisotropic diffusion iterations (Dataset E).

The anisotropic diffusion iteration count (aniso_iter) was empirically studied on the effect of fusion quality of Dataset E with the metric of QABF. Figure 17 indicates that QABF values are 0.4771 at a value of aniso_iter = 1 to 0.5322 at aniso_iter = 5 and 0.5514 at aniso_iter = 10.

The values keep on improving to 0.5779 at aniso_iter = 20 and serving its highest point to 0.5820 at aniso_iter = 35. The metric no longer changes at this step and at the following two steps, namely, at aniso_iter = 36 (0.5808) and aniso_iter = 37 (0.5817). This is a pointer that the more the iterations, the better the structural preservation, as represented by QABF until convergence is reached at about 30-35 iterations.

However, it is observed under the naked eye that the maximum QABF is achieved at aniso_iter = 35, however, in the fused image, the edges are over-enhanced slightly. However, the middle values like the aniso_iter = 20 or 25 (QABF = 0.5779 and 0.5788, respectively) produce images with the almost maximum QABF, but more natural look. Hence, it is possible to regard the practical optimum of the parameter aniso_iter = 25–30 the value offering the compromise between objective edge preservation and subjective visual quality.

VIII. Statistical Results

The results of the mean intensity values achieved on the fused images with the various fusion techniques in datasets A-E are reported in Table XII. When the mean is high, this would mean brighter fused images whereas values below the mean would make the images darker. Among the compared approaches, CDIF always has the largest mean values, which is an indication of its propensity to boost the overall image brightness and intensity representation. Conversely, the IMA, Structure-aware and VSM-and-WLS methods have a relatively lower mean value, revealing that it conserves the intensity. The Proposed method is balanced and the mean values are in-between extremes thus avoiding over brightness and having adequate levels of intensity. According to this balance, the offered approach successfully maintains image information

without excessive enhancement, and it is applicable to the creation of visually natural fusion products.

TABLE X. OPTIMIZED OBJECTIVE RESULTS OF DATASET D

Metric	A (aniso_iter = 1)	B (aniso_iter = 5)	C (aniso_iter = 10)	D (aniso_iter = 15)	E (aniso_iter = 16)	F (aniso_iter = 17)
PSNR	16.9135	18.4104	18.7958	18.9827	19.0381	19.0632
SSIM	0.656	0.7315	0.7348	0.7349	0.7353	0.7355
Entropy	7.9106	7.8872	7.8647	7.8488	7.844	7.842
SF	0.1356	0.0842	0.0712	0.0659	0.0645	0.064
Corr	0.8467	0.8736	0.8771	0.8788	0.8793	0.8796
MIF	1.6002	1.6798	1.6675	1.6489	1.6509	1.6497
ERGAS	28.035	23.6049	22.5843	22.1051	21.9649	21.9018
QABF	0.2747	0.4774	0.5466	0.5754	0.58	0.5777
LABF	0.7253	0.5226	0.4534	0.4246	0.42	0.4223

TABLE XI. OPTIMIZED OBJECTIVE RESULTS ACROSS DIFFERENT ANISOTROPIC ITERATIONS ON DATASET E

Metric	A (aniso_ iter =1)	B (aniso_ iter =5)	C (aniso_ iter =10)	D (aniso_ iter =15)	E (aniso_ iter =20)	F (aniso_ iter =25)	G (aniso_ iter =30)	H (aniso_ iter =35)	I (aniso_ iter =36)	J (aniso_ iter =37)
PSNR	25.3778	25.651	25.3219	25.1328	25.0358	24.9849	24.9561	24.9394	24.937	24.9349
SSIM	0.8151	0.7898	0.7549	0.7378	0.7294	0.7249	0.7224	0.7209	0.7207	0.7205
Entropy	6.9582	6.8523	6.7993	6.7695	6.7509	6.7373	6.7272	6.7194	6.718	6.7166
SF	0.0638	0.0429	0.0358	0.0328	0.0312	0.0302	0.0295	0.0291	0.029	0.0289
Corr	0.9087	0.9036	0.8925	0.8864	0.8831	0.8814	0.8804	0.8798	0.8797	0.8796
MIF	1.6088	1.4815	1.4091	1.375	1.3619	1.3544	1.3469	1.3471	1.3458	1.345
ERGAS	15.1409	14.672	15.2387	15.574	15.749	15.8416	15.8942	15.9246	15.9292	15.9331
QABF	0.4771	0.5322	0.5514	0.5507	0.5779	0.5788	0.5819	0.582	0.5808	0.5817
LABF	0.5229	0.4678	0.4486	0.4493	0.4221	0.4212	0.4181	0.418	0.4192	0.4183

Table XIII have the median intensity values of the fused images of dataset A-E. The median gives a good indication of central tendency, which is not as dependent on the exceptions as the mean, and therefore indicates the average intensity distribution of each fusion outcome.

Once again, CDIF has the highest median values of all the datasets, which implies that it has a steady tendency to shift the pixel intensity distribution towards brightness. Techniques like IMA, Structure-aware, and VSM-and-WLS however have relatively lower medians in common with the fact that they tend to maintain darker intensity ranges.

TABLE XII. MEAN

Technique	A	B	C	D	E
FGF-and-XDoG	101.8403	195.398	143.6299	150.4669	99.3771
CDIF	162.3344	203.4378	198.341	197.2284	157.9517
CBF	95.1402	136.4477	125.34	135.316	91.4348
Structure-aware	94.3523	142.2347	118.0005	130.9456	91.4022
LEGFF	98.502	193.2469	135.1212	127.2103	93.5756
MDHU	96.431	138.5512	121.9779	135.0653	92.1582
IMA	94.667	109.2937	116.7261	134.443	92.0464
VSM-and-WLS	94.1461	142.2892	117.7992	130.7454	91.316
GFDF	94.9752	151.2894	128.1592	135.7974	92.1704
Two scale	95.0264	163.5962	126.8719	135.944	91.6496
Proposed	100.4821	177.4762	118.9874	130.9038	86.0702

The Proposed method performs in the middle performing web where median values are neither too high nor too small. This implies that the resulting approach ensures that there is a balanced intensity distribution thus avoiding excessive brightening and at the same time creates adequate contrast. This nature is especially beneficial with image fusion, where visual naturalness and structural information should be maintained.

The values of mode of pixel intensities of datasets A through E are provided in Table IX of the appendix. The most common pixel value is the mode and therefore shows the dominant intensity level in the fused outputs. One can see a definite pattern that some of the methods, such as CDIF, LEGFF, MDHU, GFDF, and the Proposed method often obtain the maximum value (255) in at least one of the datasets, indicating saturation at the brightest intensity levels. This means that these techniques focus on high frequency details or exaggerate bright areas and this may result in the loss of subtle variations of intensity.

TABLE XIII. MEDIAN

Technique	A	B	C	D	E
FGF-and-XDoG	98	215	143	155	96
CDIF	161	228	212	208	153
CBF	92	141	123	141	89
Structure-aware	92	153	117	130	88
LEGFF	94	216	134	114	85
MDHU	93	132	119	140	89
IMA	91	111	115	140	89
VSM-and-WLS	91	153	117	129	88
GFDF	91	160	124	142	89
Two scale	92	174	128	136	89
Proposed	96	175	115	131	81

On the other hand, methods like CBF, Structure-aware, and VSM-and-WLS have a more moderate mode value as they are more natural in the intensity distribution. Interestingly, IMA shows a significantly different behavior, which IMA mode values can be as low as 1 in dataset B, which underlines the possibility of IMA bias to darker intensities in particular cases. In general, the Proposed technique is a compromise: even though it sometimes touches 255 in the bright areas, it does not have serious inconsistencies such as in IMA, so the balance of the level of intensity is balanced across the datasets.

The standard deviation (SD) values of all techniques in datasets A-E are provided in Table XV, as a statistical indicator of contrast variation in the fused images as well as intensity variation. An increased SD value represents more variation in pixel intensities, which is usually due to increased details and sharper contrasts, whereas a smaller SD value represents smoother and less contrasted results. CDIF and LEGFF are the two most successful in obtaining higher values of SD (e.g., 96.67 and 98.29 in dataset B), which confirms the fact that these two approaches do not lose or deteriorate the local features.

TABLE XIV. MODE

Technique	A	B	C	D	E
FGF-and-XDoG	87	255	255	255	82
CDIF	255	255	255	255	255
CBF	95	162	104	141	85
Structure-aware	90	163	106	147	82
LEGFF	78	255	255	255	50
MDHU	84	255	82	255	82
IMA	83	1	110	145	79
VSM-and-WLS	85	167	111	125	82
GFDF	83	255	80	255	77
Two scale	99	209	162	132	81
Proposed	89	255	255	255	68

Compared to it, other methods such as Structure-aware, VSM-and-WLS and Two scale have comparatively lower SD values implying smoother outputs with lower variability, but potentially compromising sharpness but more preferable to noise suppression. The SD values in the Proposed method are fairly high in all the datasets with SD of 58.14 in dataset C and 67.59 in dataset D which indicates a balanced performance that adds contrast to the data but does not add too much variations. This trade-off is especially important since an excessively high SD (as with LEGFF) can cause visual artifacts, whereas

excessively low SD can cause the loss of small details. On the whole, the Proposed method exhibits a regulated improvement technique, which guarantees both the quality of perception and perspective of various datasets.

TABLE XV. STANDARD DEVIATION

Technique	A	B	C	D	E
FGF-and-XDoG	45.061	94.5673	50.8681	61.4592	36.4615
CDIF	58.1626	96.6743	56.7094	57.3685	44.8635
CBF	37.7442	69.9226	43.0426	58.5484	27.0831
Structure-aware	35.3045	65.2266	39.797	52.9944	26.6951
LEGFF	53.5333	98.2925	56.0973	56.1303	51.6542
MDHU	38.4557	80.6694	50.6629	61.9297	29.2299
IMA	36.5095	61.9164	40.5135	62.4378	29.5838
VSM-and-WLS	35.1459	65.1748	39.7728	53.0151	26.7289
GFDF	38.8145	83.8269	49.882	62.4989	30.2229
Two scale	35.8177	75.6769	42.1295	51.6926	26.82
Proposed	41.9983	82.3467	58.1391	67.5879	34.2811

The values of the variance of the pixel intensities of various image fusion methods in test cases A-E are illustrated in table XVI below.

Variance is a measure of the distribution of intensities values with a larger value of variance indicating a high degree of contrast and the ability to represent details. LEGFF and MDHU demonstrate the largest variance values of the methods evaluated in a number of situations implying the great enhancement of local details, but in some instances, at the expense of increased noise. On the other hand, Structure- aware and VSM-and-WLS have lower variance, which means less variations in the outputs with lower intensity variation.

The Proposed method has shown competitive variance level in the cases C and D when it has been shown to have obtained values similar or better than the state-of-the-art techniques. This balance indicates the efficiency of the offered method towards increasing image detail without changing the intensity distribution in various situations.

IX. CHI-SQUARE TEST

In order to statistically confirm the consistency of the scores of the observed fusion qualities with ideal values, a Chi- Square. The Goodness-of-Fit test is carried out in the QAB/F metric of all the 11 methods (10 existing methods and 1 proposed method). Each method has an expected value Ei which is defined to be 1, the perfect score.

The Chi-Square test value is calculated. using:

$$\chi 2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \qquad (28)$$

where O_i is the observed QAB/F score of method *i*, and m = 11 is the number of methods. The degrees of freedom (DOF) are calculated as:

$$DOF = m - 1 = 11 - 1 = 10 \qquad (29)$$

The test is evaluated at a significance threshold of $\alpha = 0.05$, with the corresponding critical value (10,0.05) =18.307:

$$X^2_{crit} = 18.307 \qquad (30)$$

A Chi-Square value which is lower than this value will mean that the null hypothesis is accepted implying that there is no statistical significance in deviations to the expected value. Table XVII shows the Chi-Square analysis of Dataset-A in which the analysis of various fusion algorithms was tested with respect to the anticipated value. O (observed) is the results which the methods give and the expected value (E) was always 1 to make comparisons. The deviation of each technique to the expected result is measured by the difference between O and E, its squared form and the contribution to the Chi-Square.

The Proposed algorithm had the lowest Chi-Square contribution (0.09809) which means that it was nearest to the expected value among all the methods. Conversely, other approaches including IMA (0.46758) and GFDF (0.36857) had quite high deviations. The Chi-Square statistic of Dataset-A was 2.7304 which was the overall measure of the divergence between all techniques. This discussion shows that the Proposed algorithms are better in terms of stability and performance than the current methods. The computed Chi-Square value (χ^2 = 2.7304) is notably lower than the critical value of 18.307 for 10 degrees of freedom. Hence, the null hypothesis (H_0)—stating that there is no statistically significant difference between the observed and expected values.

TABLE XVI. VARIANCE

Technique	A	B	C	D	E
FGF-and-XDoG	1988.803	3779.956	2587.412	3777.251	1329.444
CDIF	3381.596	3834.419	3215.975	3291.155	2012.733
CBF	1424.623	2616.466	1852.674	3427.015	733.4955
Structure-aware	1246.405	1592.914	1583.81	2808.418	712.6264
LEGFF	2727.472	4302.255	3144.529	3150.626	2633.7
MDHU	1475.762	4204.075	2555.712	3701.157	854.3853
IMA	1331.235	2848.272	1641.301	3768.231	875.2034
VSM-and-WLS	1235.231	1562.33	1581.885	2810.608	714.4359
GFDF	1505.722	4297.957	2488.226	3750.666	913.426
Two scale	1282.91	2195.087	1774.903	2672.138	719.3132
Proposed	1763.349	2621.866	3376.203	4473.839	1175.194

Table XVIII indicates the Chi-Square of Dataset-B. The expected value was again to be 1 in all the methods. The Proposed method had a Chi-Square contribution of 0.2491, which is less than most other methods meaning that it remains relatively close to the expected value. Conversely, wrong or ineffective methods like FGF-and-XDoG (0.73891) and GFDF (0.76878) have bigger deviations, which means that the methods do not align. The general Chi-Square statistic of Dataset-2 is 5.4450, which is indicative of more penetration of performance over methods. Such findings indicate that some assessment measures are not stable and similar in different scenarios; whereas the Proposed approach shows a greater level of consistency and stability, regardless of the assessment it is undergoing. The computed Chi-Square value ($\chi^2 = 5.445$) is notably lower than the critical value of 18.307 for 10 degrees of freedom. Hence, the null hypothesis (H_0)—stating that there is no statistically significant difference between the observed and expected values.

Table XIV Chi-Square analysis of Dataset-C, respectively. In this case, the Proposed method has a bigger contribution (0.50823) than some other methods, however, the overall Chi-Square value is comparatively low at 1.6291. This shows that in the majority of the methods, the differences deviating the anticipated value are lower than in the previous data sets. Strategies like GFDF (0.003295) and LEGFF (0.045924) are also very close to the expected result though the Proposed method is also performing rather well within reasonable margins. In general, Dataset-C indicates more consistency in the results of the methods, and smaller variation in techniques.

The computed Chi-Square value ($\chi^2 = 1.6291$) is notably lower than the critical value of 18.307 for 10 degrees of freedom. Hence, the null hypothesis (H_0)—stating that there is no statistically significant difference between the observed and expected value.

TABLE XVII. CHI-SQUARE CALCULATION FOR DATASET-A

Technique	Observed (O)	Expected (E)	O - E	(O - E) ²	Chi-Square Contribution
FGF-and-XDoG	0.4666	1	-0.5334	0.28452	0.28452
CDIF	0.4744	1	-0.5256	0.27626	0.27626
CBF	0.5135	1	-0.4865	0.23668	0.23668
Structure-aware	0.5155	1	-0.4845	0.23474	0.23474
LEGFF	0.5047	1	-0.4953	0.24532	0.24532
MDHU	0.669	1	-0.331	0.10956	0.10956
IMA	0.3162	1	-0.6838	0.46758	0.46758
VSM-and-WLS	0.5212	1	-0.4788	0.22925	0.22925
GFDF	0.3929	1	-0.6071	0.36857	0.36857
Two scale	0.5759	1	-0.4241	0.17986	0.17986
Proposed	0.6868	1	-0.3132	0.09809	0.09809
Total Chi-Square Statistic					2.7304

TABLE XVIII. CHI-SQUARE CALCULATION FOR DATASET-B

Technique	Observed (O)	Expected (E)	O – E	(O – E) ²	Chi-Square Contribution
FGF-and-XDoG	0.1404	1	-0.8596	0.73891	0.73891
CDIF	0.3771	1	-0.6229	0.388	0.388
CBF	0.456	1	-0.544	0.29594	0.29594
Structure-aware	0.4747	1	-0.5253	0.27594	0.27594
LEGFF	0.2044	1	-0.7956	0.63298	0.63298
MDHU	0.2217	1	-0.7783	0.60575	0.60575
IMA	0.5992	1	-0.4008	0.16064	0.16064
VSM-and-WLS	0.2252	1	-0.7748	0.60032	0.60032
GFDF	0.1232	1	-0.8768	0.76878	0.76878
Two scale	0.1464	1	-0.8536	0.72863	0.72863
Proposed	0.5009	1	-0.4991	0.2491	0.2491
Total Chi-Square Statistic					5.445

Table XX shows the Chi-Square analysis of Dataset-D. In the present case, a number of approaches, including MDHU (0.86564) and GFDF (0.89776), have a rather large variance around the expected value. The Proposed method, on the other hand, adds only 0.18029 to the Chi-Square statistic, which is more consistent. The computed Chi-Square value (χ^2 =5.1403) is notably lower than the critical value of 18.307 for 10 degrees of freedom. Hence, the null hypothesis (H_0)—stating that there is no statistically significant difference between the observed and expected values.

TABLE XIX. CHI-SQUARE CALCULATION FOR DATASET-C

Technique	Observed (O)	Expected (E)	O – E	(O – E) ²	Chi-Square Contribution
FGF-and-XDoG	0.8436	1	-0.1564	0.02446	0.02446
CDIF	0.46	1	-0.54	0.2916	0.2916
CBF	0.5231	1	-0.4769	0.22743	0.22743
Structure-aware	0.56	1	-0.44	0.1936	0.1936
LEGFF	0.7857	1	-0.2143	0.04592	0.04592
MDHU	0.7548	1	-0.2452	0.06012	0.06012
IMA	0.5992	1	-0.4008	0.16064	0.16064
VSM-and-WLS	0.7322	1	-0.2678	0.07172	0.07172
GFDF	0.9426	1	-0.0574	0.0033	0.0033
Two scale	0.7949	1	-0.2051	0.04207	0.04207
Proposed	0.2871	1	-0.7129	0.50823	0.50823
Total Chi-Square Statistic					1.6291

TABLE XX. CHI-SQUARE CALCULATION FOR DATASET-D

Technique	Observed (O)	Expected (E)	O – E	(O – E) ²	Chi-Square Contribution
FGF-and-XDoG	0.1194	1	-0.8806	0.77546	0.77546
CDIF	0.5451	1	-0.4549	0.20693	0.20693
CBF	0.6667	1	-0.3333	0.11109	0.11109
Structure-aware	0.6778	1	-0.3222	0.10381	0.10381

LEGFF	0.9229	1	-0.0771	0.00594	0.00594
MDHU	0.0696	1	-0.9304	0.86564	0.86564
IMA	0.0643	1	-0.9357	0.87553	0.87553
VSM-and-WLS	0.2809	1	-0.7191	0.5171	0.5171
GFDF	0.0525	1	-0.9475	0.89776	0.89776
Two scale	0.2249	1	-0.7751	0.60078	0.60078
Proposed	0.5754	1	-0.4246	0.18029	0.18029
Total Chi-Square Statistic					5.1403

Table XXI offers the results of Dataset-E. The Proposed approach has a Chi-Square contribution of 0.17472 which is one of the smallest ones when compared with other methods. IMA (0.71555) and CDIF (0.2643), however, have more deviations thus indicating their inconsistency. The Chi-Square statistic of Dataset-5 equals 2.4731 in total which indicates the moderate level of variability of techniques used. These findings once again support the relative soundness of the Proposed approach which always appears to be nearer to the desired performance benchmark. The computed Chi-Square value ($\chi^2 = 2.4731$) is notably lower than the critical value of 18.307 for 10 degrees of freedom. Hence, the null hypothesis (H_0)—stating that there is no statistically significant difference between the observed and expected values.

X. Time Complexity Analysis

To obtain a global picture of the computational effectiveness of the suggested technique and the 10 base techniques, we used stacked bar chart to visualize the elapsed time in five conditions of experiments (A, B, C, D and E) as presented in Table XXII. This graph uses a logarithmic y- axis because the elapsed time range of the data used, is quite big, with the lowest and highest times being 0.024 seconds (LEGFF in condition A) and 1.3966 seconds (IMA in condition B) respectively.

The log scale is suitable at compressing large outliers (e.g., 1.1182s and 1.3966s of IMA in condition A and B, respectively; 0.8588s of GFDF in condition E) and at the same time has clear resolution to small values (barely 0.024the 0.1 seconds). In Dataset C the representation of each of the 11 techniques is a single bar with stacked segments corresponding to the elapsed times of conditions A-E represented by different colours (blue, red, teal, orange and purple, respectively). Custom ticks (0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10 seconds) have been used on the y-axis to make readings easy, even in a log scale. This design uses the minimum number of visual cues, whereas a grouped bar chart uses 55 bars to display the data, whereas appearance and instead allows both the relative total cost of computation and the contribution of each condition to each technique to be displayed in Figure 18.

TABLE XXI. Chi-Square Calculation for Dataset-E

Technique	Observed (O)	Expected (E)	O – E	(O – E) ²	Chi-Square Contribution
FGF-and-XDoG	0.6525	1	–0.3475	0.12076	0.12076
CDIF	0.4859	1	–0.5141	0.2643	0.2643
CBF	0.5113	1	–0.4887	0.23883	0.23883
Structure-aware	0.5125	1	–0.4875	0.23766	0.23766
LEGFF	0.5365	1	–0.4635	0.21483	0.21483
MDHU	0.782	1	–0.2180	0.04752	0.04752
IMA	0.1541	1	–0.8459	0.71555	0.71555
VSM-and-WLS	0.5117	1	–0.4883	0.23844	0.23844
GFDF	0.888	1	–0.1120	0.01254	0.01254
Two scale	0.544	1	–0.4560	0.20794	0.20794
Proposed	0.582	1	–0.4180	0.17472	0.17472
Total Chi-Square Statistic					2.4731

TABLE XXII. ELAPSED TIME (SECONDS) FOR DIFFERENT TECHNIQUES ACROSS DATASETS A–E

Technique	A	B	C	D	E
FGF-and-XDoG	0.0443	0.0499	0.0442	0.0379	0.0646
CDIF	0.0884	0.0561	0.0424	0.0432	0.0542
CBF	0.0334	0.0305	0.0294	0.035	0.0359
Structure-aware	0.046	0.0282	0.0341	0.026	0.032
LEGFF	0.024	0.0315	0.0349	0.0367	0.1986
MDHU	0.0514	0.053	0.0347	0.0514	0.0419
IMA	1.1182	1.3966	0.4685	0.5224	0.0499
VSM-and-WLS	0.0402	0.0276	0.0351	0.0342	0.0459
GFDF	0.0548	0.0391	0.0347	0.0371	0.8588
Two scale	0.0441	0.0497	0.0448	0.0856	0.0436
Proposed	0.0393	0.0289	0.0394	0.0321	0.0306

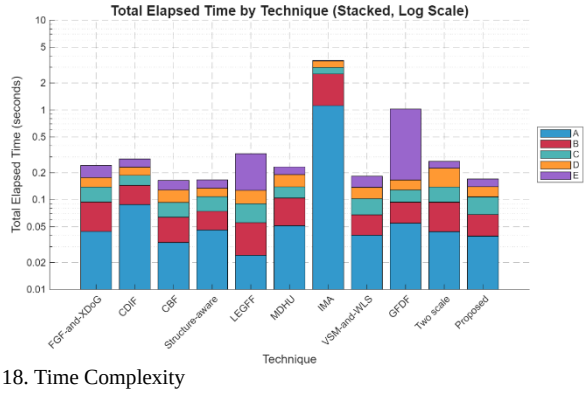

Fig. 18. Time Complexity

The tallest bar, which belongs to IMA, has the total elapsed time of about 3.56 seconds, most of which is contributed with large values of conditions A (1.1182s) and B (1.3966s), which implies that the computational overhead is very high and may become a limiting factor in time-sensitive applications. Equally, the length of the bar of GFDF in case E (0.8588s) is considerably longer, which points to input-related inefficiencies. By contrast, a technique CBF and Structure-aware have consistently short bars with a total time of between 0.16 seconds indicating their efficacy under all climatic conditions. The most feature of the proposed technique is that it has the smallest bar (61 seconds on average), and it shows the best and consistent performance due to the contribution of each condition being between 0.0289 seconds (condition B) and 0.0394 seconds (condition C).

It is worth noting that the chart shows condition-specific trends. E.g., in condition B, VSM-and-WLS was the fastest (0.0276s) and IMA the slowest (1.3966s), whereas fluctuation is apparent in condition E, in which LEGFF (0.1986s) and GFDF (0.8588s) record spiking errors compared to their mean speeds. These findings highlight the competitive advantage of the proposed technique since the size of the segments across all conditions is uniform (unless in techniques with sharp variability) with LEGFF and Two scale. The following visualization will be a complement to the highly detailed quantitative findings in Table 1 by providing a succinct and intuitive overview of computational efficiency. The fact that the stacked bar chart sums times through conditions that are shared by any given technique into a single bar makes this chart especially useful in terms of comparing overall performance and still adding the insight of contribution by individual condition. To analyze them further in the future, it may be applied to other metrics, for example, such as the use of memory or energy to explain further the trade-offs between techniques of processing images.

XI. FUTURE CHALLENGES

Although effective, the proposed framework has a number of weaknesses. The experimental assessment process was restricted to benchmark datasets, and this aspect can be a limitation to the applicability to heterogeneous real-world data. Computational complexity, despite analysis, was defined at a controlled hardware condition and might change depending on system setups. Moreover, the evaluation was mostly based on chi-square performance and elapsed time; these measures, even though productive, do not provide full coverage of the quality of perception and application needed.

The direction of the future work will include making the evaluation applicable to large- scale and high-resolution datasets to enhance the robustness analysis. In addition, alternative fusion and decomposition techniques, including shift-invariant and directional approaches such as D-SHIFT and related multiscale representations, are explicitly proposed as potential extensions to further enhance spatial detail preservation and reduce artifacts.

To reduce the computation overhead, algorithmic optimisation will be sought after to ensure that the quality is not compromised. The use of more objective measures, such as perceptual quality indices, and subjective assessments will further increase the validity of the performance assessment. In addition, their focus will be on changing the framework to be deployed in real-time and look into the domain-specific applications, such as remote sensing, medical imaging, and autonomous navigation.

XII. ABLATION STUDY

To analyze the importance of each and every constituent of the proposed pan-sharpening framework, an ablation study was thoroughly done by systematically altering the important parameters of the algorithmic strategies and fusion, and holding the rest of the pipeline constant. Objective quantitative findings that were reported in Tables VI-VIII and matching visual observations, as discussed in the experimental discussions, form the cornerstone of the analysis. Anisotropic diffusion iteration count was examined initially, as it is the parameter that dictates the trade-off between maintaining edge and reducing noise in high-frequency sub-bands. In Dataset A, the aniso_iter was changed to 1-37 with aniso_grad kept constant at aniso-grad = 2, pyr-levels kept constant at 3, and guided filtering parameters radius r= 45 and regularization 0.3.

A greater number of iterations led to a continuous increase in the structural preservation, where the QABF measure ranged from 0.5231 at aniso_iter= 1 to the maximum of 0.6868 at

aniso_iter= 35, as it was reported in Table VI. This is a relative improvement of about 31.3, and that concludes the fact that iterative anisotropic diffusion improves edge continuity and structural visibility. But the gains dwindled further, and the eye could see slight over-accentuation of ends. Thus, the most appropriate settings of Dataset A were chosen to be aniso_iter = 35, which offers a trade-off between edging objectives and the perceptions. In Dataset B, two tests were investigated; namely, the effect of anisotropic diffusion with a smaller iteration range (aniso_iter = 13) but the same guided filtering parameters (r = 45, ε = 0.3) with a larger gradient threshold to make sure higher noise levels were also taken into account.

Table VII presents the optimum value of QABF (0.5009) at anisoiter =1. Nevertheless, the homogeneous regions had unresolved noise in the corresponding fused images. When the iteration number was increased to 2 or 3, it had a small negative effect on QABF (0.4921 and 0.4827, respectively) but a large positive effect on perceptual smoothness and texture continuity. This fact causes the weakness of the edge-based metrics only and suggests that aniso is better to be chosen as 2-3 as a perceptual optimum configuration of Dataset B.

The same tendency is traced in Dataset C, in which the number of iterations of anisotropic diffusion was once again changed to 1-3, with the gradient threshold being low (aniso_grad = 0.2) to focus on preserving the fine details. Table IX indicates that the maximum QABF value was 0.2871 at aniso iter = 1, but this gave a significant noise artifact in the resultant image. More and more increases in the iteration number minimised QABF to 0.2691 and 0.2589 with aniso iteration of 2 and 3, respectively, and enhanced visual smoothness and noise amplification. These findings validate that the diffusion number will bring higher quality perceptual results at the expense of a minor drop in quantitative edge-based measures.

The interaction of the depth of multiscale decomposition was also analysed in all the datasets. In the majority of experiments, a contourlet pyramid of pyr_levels = 3 was used because it was effective in separating between global and fine-spatial luminance information. Empirical tracking showed that the number of levels of decomposition led to a lack of sufficient space detailing abduction, and that deeper pyramids led to a greater cost of computation and slight ring artifacts without much metric enhancements. Thus, pyr levels 3 have been used as a powerful and efficient setup in computation. It was demonstrated that saliency-guided low-frequency fusion played an important role in comparison with uniform averaging.

In this approach, Laplacian-based saliency extraction is used, then, with Gaussian smoothing of 15 x 15 and standard deviation, ions s =2 with gamma correction (gamma =0.5). These parameters remained constant in all datasets so that they could behave the same way. Adaptive weighting technique performed far better in terms of entropy and mutual information values indicated in Tables VI, V, and V does not create rough transitions between intensities and also does not exhibit luminance biases, which were observed when plain averaging was applied.

Lastly, the max-absolute selection measure was evaluated on high-frequency coefficient integration. In all datasets, the application of this rule was substituted by the use of coefficient averaging that led to a decrease in spatial frequency and edges. The max-absolute rule was always more effective in raising the sharpness of the boundary and structurization, especially with anisotropic diffusion. The positive interaction is manifested in better values of QABF and LABF, as described in Tables VI through VIII, which supports the idea that the rule is effective in retaining major spatial patterns and diffusion in eliminating noise.

Altogether, the ablation experiment that was facilitated by Table VI8 shows that every component and parameter option in the suggested framework has a significant impact on the final quality of the fusion. The choice of parameter values, pyr levels = 3, aniso = 2, aniso-iter = 30-35 (depending on the dataset) approximate and balanced between spatial quality, spectral fidelity and perceptual naturalness, the Gaussian kernel size of 15 times 15 with s = 2, gamma = 0.5, and guided filtering parameters of, r = 45, epsilon = 0.3.

By eliminating or under-mishandling any of these factors, significant deteriorations in the fusion performance are observable, justifying the jointness of the offered approach.

XIII. CONCLUSION

This paper presented a Contourlet-based pan-sharpening model designed for remote sensing, especially in the area of environmental analysis and monitoring. The proposed approach integrates the high spatial, high-resolution content of panchromatic images with the spectral richness of multispectral images, thereby suppressing spectral distortion and fusion artifacts.

The proposed method utilises simple averaging to preserve low-frequency components, and Laplacian-based saliency weighting to inject high-frequency details in order to maintain a balance between spatial fidelity and spectral integrity. Comprehensive experiments conducted on multiple benchmark datasets demonstrated that the proposed method consistently outperforms the compared algorithms. Visual analysis ensured that the fused images preserve sharper edges, enhanced contrast, and better depiction of the environmental features like vegetation, water bodies, and urban buildings.

Furthermore, quantitative evaluation with metrics such as PSNR, SSIM, ERGAS, and QABF states the capability of the proposed method to achieve high-quality fusion with minimal spectral distortion. Moreover, the Computational analysis also states that the proposed method is efficient, making it suitable for large-scale real-world applications.

In summary, the proposed framework provides an effective and interpretable solution for high-fidelity pan sharpening, advancing the field of remote-sensed image fusion. It enhances the precision of environmental surveillance and supports informed decision-making in domains such as land-use management, resource planning, and disaster assessment. Future work will focus on integrating adaptive parameter tuning and using deep learning-based feature extraction modules to improve the performance, robustness, and make it applicable to hyperspectral and real-time satellite imaging systems.

REFERENCES

[1] F. Paciolla, G. Popeo, A. Farella, and S. Pascuzzi, “Agronomic information extraction from UAV-based thermal photogrammetry using MATLAB,” Remote Sensing, vol. 17, no. 15, p. 2746, 2025.

[2] Y. Xue, J. Wang, Y. Wang, C. Wu, and Y. Hu, “Preliminary study of Grid computing for remotely sensed information,” International Journal of Remote Sensing, vol. 26, no. 16, pp. 3613–3630, 2005.

[3] Y. Ma, H. Wu, L. Wang, et al., "Remote sensing big data computing: Challenges and opportunities," Future Generation Computer Systems, vol. 51, pp. 47–60, 2015.

[4] G. Aloisio, M. Cafaro, I. Epicoco, and G. Quarta, "A problem solving environment for remote sensing data processing," in Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC 2004), IEEE, vol. 2, pp. 56–61, 2004.

[5] Z. Wu, J. Sun, Y. Zhang, Z. Wei, and J. Chanussot, "Recent developments in parallel and distributed computing for remotely sensed big data processing," Proceedings of the IEEE, vol. 109, no. 8, pp. 1282–1305, 2021.

[6] S. Krishnamoorthy and K. Soman, "Implementation and comparative study of image fusion algorithms," International Journal of Computer Applications, vol. 9, no. 2, pp. 25–35, 2010.

[7] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, "A new pansharpening algorithm based on total variation," IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 1, pp. 318–322, 2013.

[8] M. Ciotola, G. Guarino, G. Vivone, et al., "Hyperspectral pansharpening: Critical review, tools, and future perspectives," IEEE Geoscience and Remote Sensing Magazine, 2024.

[9] Q. Xu, Y. Zhang, and B. Li, "Recent advances in pansharpening and key problems in applications," International Journal of Image and Data Fusion, vol. 5, no. 3, pp. 175–195, 2014.

[10] L. He, Y. Rao, J. Li, et al., "Pansharpening via detail injection-based convolutional neural networks," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 4, pp. 1188–1204, 2019.

[11] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art," Information Fusion, vol. 19, pp. 4–19, 2014.

[12] L. J. Chipman, T. M. Orr, and L. N. Graham, "Wavelets and image fusion," in Proceedings of the International Conference on Image Processing, IEEE, vol. 3, pp. 248–251, 1995.

[13] G. Simone, A. Farina, F. C. Morabito, S. B. Serpico, and L. Bruzzone, "Image fusion techniques for remote sensing applications," Information Fusion, vol. 3, no. 1, pp. 3–15, 2002.

[14] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," Information Fusion, vol. 36, pp. 191–207, 2017.

[15] L. Alparone, B. Aiazzi, S. Baronti, and A. Garzelli, Remote Sensing Image Fusion, CRC Press, 2015.

[16] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," IEEE Transactions on Image Processing, vol. 22, no. 7, pp. 2864–2875, 2013. DOI: 10.1109/TIP.2013.2244222.

[17] Y. Jie, X. Li, F. Zhou, H. Tan, et al., "Medical image fusion based on extended difference-of-Gaussians and edge-preserving," Expert Systems with Applications, vol. 227, p. 120301, 2023.

[18] X. Meng, N. Wang, F. Shao, and S. Li, "Vision transformer for pansharpening," IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–11, 2022. DOI: 10.1109/TGRS.2022.3166972.

[19] B. K. S. Kumar, "Image fusion based on pixel significance using cross bilateral filter," Signal, Image and Video Processing, vol. 9, no. 5, pp. 1193–1204, 2015. DOI: 10.1007/s11760-013-0556-9.

[20] W. Li, Y. Xie, H. Zhou, Y. Han, and K. Zhan, "Structure-aware image fusion," Optik, vol. 172, pp. 1–11, 2018. DOI: 10.1016/j.ijleo.2018.06.096.

[21] Y. Zhang, W. Xiang, S. Zhang, et al., "Local extreme map guided multi-modal brain image fusion," Frontiers in Neuroscience, vol. 16, p. 1055451, 2022. DOI: 10.3389/fnins.2022.1055451.

[22] Y. Jie, X. Li, H. Tan, F. Zhou, and G. Wang, "Multimodal medical image fusion via multi-dictionary and truncated Huber filtering," Biomedical Signal Processing and Control, vol. 88, p. 105671, 2024. DOI: 10.1016/j.bspc.2023.105671.

[23] A. Sufyan, M. Imran, S. A. Shah, H. Shahwani, and A. A. Wadood, "A novel multimodality anatomical image fusion method based on contrast and structure extraction," International Journal of Imaging Systems and Technology, vol. 32, no. 1, pp. 324–342, 2022. DOI: 10.1002/ima.22621.

[24] J. Ma, Z. Zhou, B. Wang, and H. Zong, "Infrared and visible image fusion based on visual saliency map and weighted least square optimization," Infrared Physics & Technology, vol. 82, pp. 8–17, 2017. DOI: 10.1016/j.infrared.2017.02.005.

[25] X. Qiu, M. Li, L. Zhang, and X. Yuan, "Guided filter-based multi-focus image fusion through focus region detection," Signal Processing: Image Communication, vol. 72, pp. 35–46, 2019. DOI: 10.1016/j.image.2018.11.003.

[26] D. P. Bavirisetti and R. Dhuli, "Two-scale image fusion of visible and infrared images using saliency detection," Infrared Physics & Technology, vol. 76, pp. 52–64, 2016. DOI: 10.1016/j.infrared.2016.01.009.

[27] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," Information Fusion, vol. 24, no. 1, pp. 147–164, 2015. DOI: 10.1016/j.inffus.2014.09.002.

[28] D.-Y. Po and M. N. Do, "Directional multiscale modeling of images using the contourlet transform," IEEE Transactions on Image Processing, vol. 15, no. 6, pp. 1610–1620, 2006.

[29] P. Feng, Y. Pan, B. Wei, W. Jin, and D. Mi, "Enhancing retinal image by the contourlet transform," Pattern Recognition Letters, vol. 28, no. 4, pp. 516–522, 2007.

[30] M. G. Reddy, P. V. N. Reddy, and P. R. Reddy, "Medical image fusion using integrated guided nonlinear anisotropic filtering with image statistics," International Journal of Intelligent Engineering and Systems, vol. 13, no. 1, pp. 25–34, 2020.

[31] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of MRI data," IEEE Transactions on Medical Imaging, vol. 11, no. 2, pp. 221–232, 1992.

[32] V. V. Romanyuk, "Maximum-versus-mean absolute error in selecting criteria of time series forecasting quality," vol. 1, no. 96, pp. 3–9, 2021.

[33] M. W. Law and A. C. Chung, "Weighted local variance-based edge detection and its application to vascular segmentation in magnetic resonance angiography," IEEE Transactions on Medical Imaging, vol. 26, no. 9, pp. 1224–1241, 2007.

[34] S. Sharma et al., "A deep learning approach for the detection of vehicles in satellite images," IEEE Access, vol. 10, 2022. [Online]. Available: https://www.researchgate.net/publication/358880789_A_deep_learning_approach_for_the_detection_of_vehicles_in_satellite_images. DOI: 10.1109/ACCESS.2022.3219174.

[35] T. Wang Wei Hong, Pansharpening by Convolutional Neural Network, GitHub repository, 2025. [Online]. Available: https://github.com/ThomasWangWeiHong/Pansharpening-by-Convolutional-Neural-Network.

[36] S. Amburose, Pansharpening, GitHub repository, 2021. [Online].Available:https://github.com/amburosesekar/PanShaperning

[37] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pp. 629–639, 2002

[38]Liu, D., Wang, E., Wang, L., Benediktsson, J.A., Wang, J. and Deng, L., 2024. Pansharpening Based on Multimodal Texture Correction and Adaptive Edge Detail Fusion. Remote Sensing, 16(16).

[39]Zhang, J., Guo, X. and Kang, X., 2025, August. Performance Analysis of Pan-sharpening Algorithms Under Different Imaging Conditions. In IGARSS 2025-2025 IEEE International Geoscience and Remote Sensing Symposium (pp. 1900-1904). IEEE.

[40]Wang, J., Lin, Y., Chen, C., Huang, X., Zhang, R., Wang, Y. and Lu, T., 2025. From forgotten to pan-sharpening. Pattern Recognition, p.112653.

.