



Multi Focus Image Fusion using Image Enhancement Methods

Mian Muhammad Danyal¹, Omar Bin Samin¹, Sarwar Shah Khan^{2,*} and Sajid Khan²

¹Center for Excellence in Information Technology, Institute of Management Sciences, Peshawar 25000, Pakistan

²Department of Computer Science, University of Engineering and Technology Mardan, Mardan 23200, Pakistan

Abstract

The challenge with multifocus images lies in different regions being in focus across various shots, resulting in some areas appearing blurry while others are sharp. This issue is prevalent in fields such as medical imaging, remote sensing, and photography, where clear and detailed images are essential. This project introduces a novel approach to multifocus image fusion by integrating the Marr–Hildreth edge detection technique with Discrete Cosine Transform (DCT), Stationary Wavelet Transform (SWT), and Discrete Wavelet Transform (DWT). The Marr–Hildreth algorithm detects edges by identifying zero-crossings in the Laplacian of a Gaussian-blurred image, effectively highlighting areas with significant intensity changes. The proposed method was evaluated across four datasets—clocks, leaves, balloons, and bottles. Performance metrics such as Peak Signal-to-Noise Ratio (PSNR), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Entropy indicated that this integrated approach significantly enhances image quality. By combining Marr–Hildreth edge detection with DCT, SWT, and

DWT, the method improves image clarity and detail, offering a promising advancement in image fusion techniques. This innovative approach not only addresses multifocus image challenges but also paves the way for further developments in image fusion processes. The study suggests that enhanced image processing techniques can be applied across various fields requiring high-quality image fusion.

Keywords: image fusion, multifocus images, discrete cosine transform, high boost filter, discrete wavelet transform, marr-hildreth edge detection.

1 Introduction

Capturing images in everyday life is that it helps remember memories. Taking images allows you to capture special moments, everyday occurrences, and the evolution of people and places over time [1]. An ideal image should perfectly capture every detail of a scene, providing all the necessary information. However, since cameras have limitations, this isn't always possible [2]. One of the most common photographic challenges is to capture an image with varying focal depths, where other parts of the scene have become focused, as well as parts that have been blurred or unclear. This is a result of the fact that the lens of a camera cannot focus at a different depth.



Submitted: 11 April 2025

Accepted: 19 May 2025

Published: 26 June 2025

Vol. 1, No. 2, 2025.

doi:10.62762/JIAP.2025.772403

*Corresponding author:

✉ Sarwar Shah Khan

sskhan0092@gmail.com

Citation

Danyal, M. M., Samin, O. B., Khan, S. S., & Khan, S. (2025). Multi Focus Image Fusion using Image Enhancement Methods. *ICCK Journal of Image Analysis and Processing*, 1(2), 57–72.



© 2025 by the Authors. Published by Institute of Central Computation and Knowledge. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Image fusion techniques are employed to overcome this problem [3].

Image fusion combines elements from multiple photos into a single picture that captures every detail of the scene clearly. This technique is particularly useful when a camera's low depth of field avoids it from capturing everything in focus at once. For example, when taking a photo, the camera may focus well on either a nearby or a distant object, but not both. Image fusion addresses this problem by combining images taken at different distances [4]. Image fusion is used in many different kinds of applications to improve image clarity and information content. In medical imaging, it combines data from various scans such as MRI and CT to improve diagnostic accuracy. Remote sensing uses images from multiple spectral bands to improve monitoring of the environment. It is also used in monitoring to combine feeds from multiple cameras for complete monitoring, and in automotive technology to help advanced driver-assistance systems by consolidating sensor data to improve safety features [5].

Multifocus image fusion is an approach for combining a set of images that are differently focused, to form a single fused composite image and clear representation such that all useful information in every image is well-focused ensuring the efficiency on further analysis or visualization not limited by any loss occurred from defocusing [6]. Multifocus image fusion is majorly used in medical imaging for sharp diagnostic visuals, and improved surveillance so clearer image and photography to achieve high depth of field [7].

In this article, we explore the application of image fusion techniques consisting of Marr–Hildreth edge detection, Discrete Cosine Transform (DCT), Stationary Wavelet Transform (SWT), and Discrete Wavelet Transform (DWT) to perform image fusion, especially multi-focus one. The Marr–Hildreth technique enhances image features by applying a Gaussian blur followed by the Laplacian operator, which helps in detecting and preserving edges. The Discrete Cosine Transform (DCT) is used for separating the image into parts of varying importance, increasing efficiency by focusing on the more important components [8]. Furthermore, the Stationary Wavelet Transform (SWT) helps in maintaining shift-invariance, ensuring better structural preservation in images, while the Discrete Wavelet Transform (DWT) provides efficient multi-resolution decomposition, allowing images to be analyzed

at different levels of detail, which is important for effectively combining data from multiple sources [9].

The first step involves collecting two multi-focused images and applying Marr–Hildreth edge detection as a preprocessing technique. This method enhances image sharpness by detecting and preserving edges, which contributes to better fusion results. After preprocessing, the images are finally combined using image fusion techniques such as DCT, SWT, and DWT to get a single clear, enhanced composite image. The objective is to produce a final image that enhances significant information from the original elements, improving the quality of the total image.

The dataset for this experiment contains four groups of grayscale photos: "Balloons," "Leaves," "Clock," and "Clocks and Leaves." These collections are based on the "Lytro multi-focus datasets" [10]. They have been selected specifically to improve the evaluation and testing of multi-focus images using modern techniques. This study made the following contributions:

- The images were pre-processed using the Marr–Hildreth edge detection technique to enhance image clarity and sharpness. This method involves applying a Gaussian blur to the image to reduce noise, followed by computing the Laplacian to detect areas with rapid intensity changes, effectively highlighting edges and fine details.
- The proposed method used advanced image fusion methods DCT, SWT and DWT that combines multiple images captured at different focal depths to improve image clarity and detail. The experiments also discuss the fusion results using different datasets.
- The experiments were designed to check the efficiency of multi-focus image fusion on a dataset that includes four sets of grayscale images ("Clocks" "Balloons," "Leaves," and "Clock and Leaves"). This testing is important for calculating the effectiveness of the image fusion techniques used.

The paper is structured as follows: Section 2 covers previous research. In Section 3, we describe our research methodology. Section 4 describes the datasets, and Section 5 presents the results and discussions. Finally, Section 6 summarizes the findings and outlines future directions of the study.

2 Literature Review

Multifocus image fusion implies combining images taken at various focus levels to achieve one focused image, where details are well-defined all over. The development has taken place from simple pixel-based techniques to better synthesis algorithms, including wavelet transforms and neural networks. Applications have been noted in various imaging fields, including medical images, remote sensing, and photography, for clarity and detail to help with better analysis and interpretation. Different researchers worked in the area of multifocus image fusion such as Liu et al. [11] suggested method for a new Algorithm of Multi-Focus Color Image Fusion Based on Low-Vision Image Reconstruction and Focused Feature Extraction. They proposed an architectural gradient-based method to determine the focused feature and how it can be extracted to increase the image resolution in the multi-focused point scenarios. One of the main novelties in their approach was a deep residual network, together with a degradation model designed to further improve the level of detail in the fused images. The algorithm accurately detected areas of focus and realized a huge progress in keeping the image's detail and edge structures well preserved, compared to the existing techniques.

Li et al. [12] introduced a swift and efficient image fusion approach that strategically merges multiple images to generate a single, highly informative composite. This method hinges on a dual-scale decomposition framework, where an input image is separated into a base layer—capturing large-scale intensity variations—and a detail layer—responsible for finer structural elements. To achieve effective fusion, a novel guided filtering-based weighted average algorithm is employed, ensuring robust spatial consistency throughout the fusion process. Extensive experiments involving multispectral, multifocus, multimodal, and multiexposure image datasets underscore the superior performance of this technique, as it consistently delivers fused images of remarkable visual clarity and detail preservation across diverse application scenarios.

Zhang et al. [2] presented a comprehensive and detailed survey of multi-focus image fusion (MFIF), specifically addressing the advancements and limitations of deep learning-based methods that have emerged in this field since 2017. Recognizing the gap in the existing literature—namely, the absence of an exhaustive review focusing on deep learning for MFIF—this paper thoroughly reviews a broad

array of algorithms, datasets, and evaluation metrics employed in contemporary MFIF research. As the first of its kind, this survey not only charts the progress of deep learning-driven MFIF approaches but also systematically compares them with traditional fusion techniques, highlighting significant improvements in performance as evidenced by both qualitative and quantitative metrics. Furthermore, the authors provide insightful commentary on the present state of MFIF and outline promising directions for future investigations, thereby offering a valuable roadmap for further research and applications in high-detail imaging domains.

Turgut et al. [13] used a gradient transform-based novel method to combine the focused and unfocused pixels. This aimed to counter the common problems witnessed within the performance-hindering already in use methods, such as shift variance and misregistration. It starts with generating halftone images using the Halftoning-Inverse Halftoning (H-IH) transform. The focus measurement is calculated for these gradient images through the Energy of Gradient (EOG) and Standard Deviation functions, forming a composite image. This composite was further improved with a decision fusion strategy using majority voting for the combination. The method was at the time very unique to most common robustness issues in fusion and, at the same time, captures the details which are normally unobservable by other methods. This new innovative approach has shown excellent results compared to 17 different modern and traditional techniques, both visually and objectively. In a more objective manner, the assessments were using six quantitative metrics, where with the new method doing better in five out of the six, pointing to a huge improvement over other technique.

Li et al. [6] proposed a robust and effective deep learning-based image fusion approach designed to address the challenges of integrating infrared and visible images into a single composite image with all essential features preserved. The technique initiates with a decomposition of the input images into base and detail components, enabling a structured and balanced fusion. Specifically, the base parts of the images are merged through a weighted averaging process to maintain global structure, while the detail content is refined using a deep learning network to extract multi-layer features that encapsulate rich spatial and spectral information. To further improve the fused detail quality, a combination of ℓ_1 -norm and

weighted-average strategies is utilized to generate multiple candidate fused detail contents, followed by a max selection step to finalize the fusion process. Ultimately, the fused image is reconstructed by integrating the refined base and detail components. Experimental evaluations underscore the approach's superior performance in both visual quality and quantitative metrics, positioning it as a promising tool for high-detail image fusion applications.

Zhou et al. [3] presented a comprehensive review of multi-focus image fusion (MFIF), a critical technique in overcoming the depth of field limitations inherent in optical lenses by merging multiple partially focused images into a single, fully focused composite. The authors systematically introduced a new classification system for MFIF algorithms, categorizing them into four distinct groups: transform domain methods, boundary segmentation methods, deep learning-based methods, and hybrid fusion approaches. Within this framework, they elaborated on the defining characteristics of boundary segmentation methods, highlighting their effectiveness in preserving edge information and maintaining image integrity compared to conventional methods. The paper further delineates both subjective and objective evaluation metrics, offering an in-depth discussion of eight widely recognized objective indicators used to assess the quality of fusion results. Drawing on an extensive body of literature, the authors compared and summarized a variety of representative algorithms, providing a well-rounded overview of the field. The work concludes by identifying the primary challenges faced by current MFIF research and outlining future directions for development, underscoring the method's potential for advanced applications requiring high levels of detail and visual clarity, such as medical imaging and remote sensing.

Kaur et al. [14] explored the intrinsic limitations of optical lenses, which, constrained by finite depth-of-field (DOF), struggle to produce fully focused images in a single capture. In practice, only elements within the DOF appear sharp, leaving other regions blurred. To address this challenge, researchers have proposed a suite of algorithms aimed at fusing multiple partially focused images of the same scene into a single, fully focused composite. The authors systematically reviewed these algorithms, classifying them into conventional and deep learning-based approaches, while emphasizing the unique challenges of multi-focus image fusion (MFIF). Applications of MFIF are diverse and impactful, spanning optical

microscopy, micro-image fusion, digital photography, and remote sensing networks. Beyond outlining the conceptual and algorithmic underpinnings, this paper performed a comprehensive qualitative and quantitative evaluation of seven representative MFIF methods using the Real-MFF dataset. The findings highlighted both the advancements and current limitations of existing approaches, culminating in a forward-looking discussion on promising directions for future research and development in the field. Pan et al. [15] suggested an innovative methodology that is DDMF, as denoted by the Feature Difference Network for the betterment of Multi-Focus Image Fusion MFIF. The technique introduces noise to highlight the difference in the information loss between the focused and defocused regions of an image. The method applies Gaussian noise to the picture to improve the representation of features at the denoising stage while using a Denoising Diffusion Probabilistic Model diffusion process. This enables DDMF to detect and label each pixel much more accurately. Their experiments show that this new approach outperforms existing methods in terms of visual quality and objective performance measures.

You et al. [16] introduced new multi-focus image fusion technique by applying the local standard deviations of Laplacian images from the input images that are further enhanced with a guided filter. The key innovation is that locally computed standard deviation indicates the sharpness of the images; indeed, if the pixel is sharper, then the local variance is higher. This technique includes the application of Laplacian filters to source images, followed by calculation and enhancement of local standard deviations by a guided filter, and then the creating of decision maps for the selection of pixels based on the data that is enhanced. Following this, a variant method is applied to refine this decision map further using guided filtering. The approach resulted in much better performance in comparison with those of the prior advanced techniques. Farid et al. [17] have proposed a novel multifocus image fusion scheme to enhance the depth of field and fuse in-focused regions from multiple partially focused images. Blur is induced in another less-in-focus region of the selected single-in-focus image. These authors have applied the CAB algorithm with selective image blurring based on the local content of the image. With this approach, it can be ensured that meaningful blur is induced over focus regions, and blur regions are left relatively unchanged. The initial segmentation map, extracted from the difference

between the original and CAB-blurred images, is further refined for increased accuracy. The algorithm has been tested and outperforms all the current methods using two public datasets.

Bhat et al. [18] further enhanced MFIF by overcoming the problems of low resolution, high noise, and blurriness in traditional techniques. The method used the combination of NSWT in a Neutrosophic Set and Stationary Wavelet Transform to produce more transparent and more informative fused images. NSWT improves image quality and Depth-of-Field better than some advanced methods, as shown during two dataset tests. Wang et al. [19] introduce a new multifocus image fusion algorithm using CNN's in the DWT domain. Their method blends spatial and transform domain techniques, with the CNN creating decision maps for frequency sub bands, replacing traditional fusion rules. The steps include decomposing images into sub bands, using the CNN to generate weight maps, refining these maps with postprocessing, and fusing the sub bands to produce the final image.

Compared to previous studies, our research introduces a distinctive methodology by integrating DCT, SWT, and Discrete Wavelet Transform (DWT) for image fusion, complemented by Marr-Hildreth edge detection as a pre-processing step. This approach emphasizes enhancing edge details and overall image clarity prior to the fusion process. In contrast, prior works such as Li et al. [12] have explored the combination of Laplacian filters with Discrete Fourier Transforms (DFT) for image sharpening, while Liu et al. [11] have provided comprehensive surveys on multi-focus image fusion techniques, focusing on various transformation domains. Other methodologies, including those by Turgut et al. [13], have utilized gradient transform-based methods, and You et al. [16] have employed local standard deviations enhanced by guided filters. Unlike these, our structured framework synergizes frequency domain transformations with edge enhancement techniques, leading to superior image quality. This is evidenced by higher PSNR and RMSE values, underscoring the effectiveness and applicability of our approach. The summary of literature review is shown in Table 1.

3 Methodology

With the increased demand over the years in the field of digital image processing, the popular techniques enhance the level of precision and detail in an image. In this article, a two-step image fusion methodology

is proposed as shown in Figure 1. The first step sharpens the input images to a better-quality level of sharpness by using Marr-Hildreth edge detection. The more informative the detailed image would be, the higher the contrast in the fused image. In the second step, the enhanced images were fused together using DCT fusion, SWT fusion, and DWT fusion methods. These enhanced images are then combined to generate a single highly informative image. The proposed method is designed to improve the most important details of the original input images, since this would increase the overall quality of the fused images.

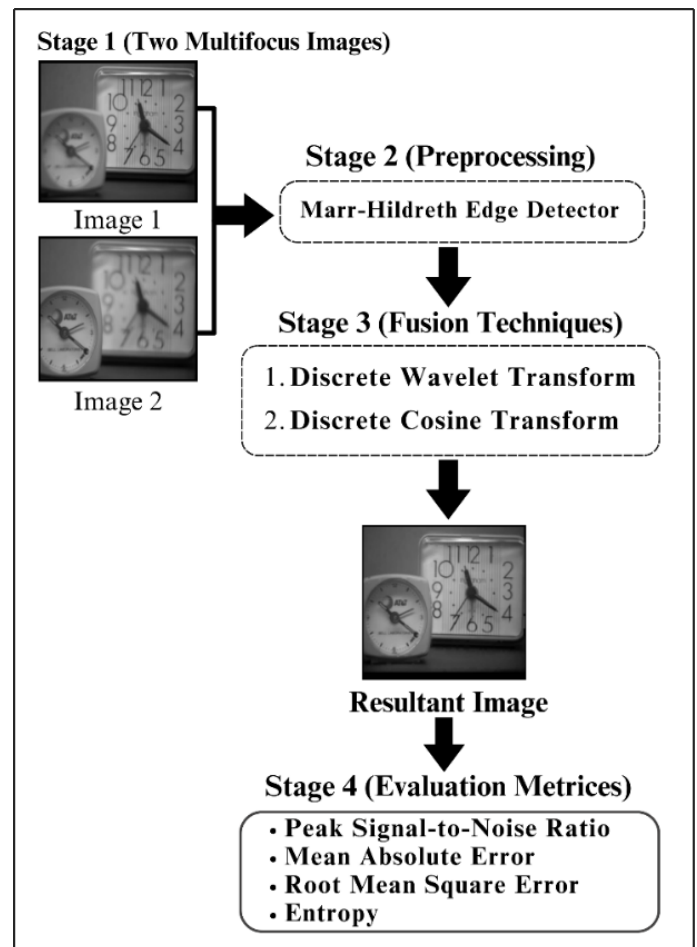


Figure 1. Proposed methodology.

3.1 Preprocessing

Pre-processing in image fusion refers to processes and methods applied to input images before these images are combined or merged into one for a single, enhanced image with higher visual quality and information content [20]. The objective of image fusion pre-processing is to prepare input images for effective fusion, where the most important and usable information will be contained in the result of the fused image [21]. A few of the techniques of pre-processing that the framework uses are described:

Table 1. Summary of literature review.

Reference	Technique(s)	Preprocessing Method(s)	Dataset(s)	Evaluation Measures	Pros & Cons
Liu et al. [11]	Low-vision reconstruction + focused feature extraction with deep residual network	Low-vision image reconstruction	Multi-focus color images (clock, leaves)	PSNR, SSIM, edge preservation rate	+ Accurate focus detection and edge preservation– High computational cost of deep networks
Li et al. [12]	SWT & DWT fusion	Laplacian filter + DFT (LF+DFT)	Medical (breast CT, MRI) & non-medical	RMSE, PSNR, Entropy	+ Sharpens structures effectively– DFT may introduce ringing; extra steps increase runtime
Zhang et al. [2]	SWT fusion	LF+DFT + Unsharp masking	"Planes" & "Clocks"	RMSE, Entropy	+ Combined sharpening methods improve detail– Complex preprocessing pipeline
Turgut et al. [13]	Halftoning-inverse + gradient-based fusion	—	Standard multifocus benchmarks	EOG, Std Dev, visual inspection	+ Robust to shifts/misregistration– Multi-stage; relatively slow
Li et al. [6]	DCT fusion	Unsharp masking	Clock, Leaves	PSNR, SSIM	+ Simple and fast– Less effective on very high-frequency details
Zhou et al. [3]	Two-level SWT fusion	Laplacian filter + DFT	Multi-focus color	Std Dev, Mean, Entropy	+ Reduces distortion and preserves detail– Two-level approach adds latency
Kaur et al. [14]	Statistical/CNN-based fusion metric validation	—	Large-scale multimodal with ground truth	Correlation measures, CNN-based metric	+ Validates metrics against CNN features– Focused on evaluation rather than new fusion algorithm
Pan et al. [15]	Denoising Diffusion + feature-difference network (DDMF)	Gaussian noise injection	Multimodal test sets	Visual IQA, PSNR, RMSE	+ Superior visual fidelity– Computationally heavy diffusion process
You et al. [16]	Guided-filter decision maps via local standard deviation	Laplacian → local std → guided filter	Custom multifocus sets	PSNR, visual comparison	+ Precise pixel selection– Sensitive to noise; filter parameters critical
Farid et al. [17]	CAB selective blur fusion	Content-adaptive blurring	Public multifocus datasets	IQA, segmentation accuracy	+ Preserves depth cues– Depends on accurate blur map
Bhat et al. [18]	Neutrosophic set + NSWT fusion	—	Two medical datasets	PSNR, Depth-of-field measure	+ Excellent noise handling– Complex neutrosophic processing

3.1.1 Marr-Hildreth Edge Detector (MHED)

The Marr–Hildreth technique is a widely used edge detection method in image processing that enhances the clarity of object boundaries by detecting regions of rapid intensity change. It applies Gaussian smoothing to reduce noise and then computes the Laplacian to identify zero-crossings, which represent edges in the image [22]. This approach effectively highlights fine details and structural features, making it particularly useful for applications requiring precise edge detection, such as medical imaging, remote sensing, and object recognition [23]. In the context of multi-focus image fusion, the Marr–Hildreth technique helps refine image details before fusion, ensuring that only the most relevant and well-defined features contribute to the final composite image [24]. Although Marr–Hildreth Edge Detection (MHED) is

traditionally utilized to detect edges in an image, it is adapted here as part of the preprocessing step. Instead of using it for edge detection, MHED is employed to enhance the overall sharpness and edge clarity of the input images. This enhancement enables the fusion algorithms to better combine the information from multiple sources, resulting in a higher-quality fused image. The modified MHED process prepares the images by sharpening them and making edge details more prominent, which is essential for effective fusion.

3.2 Fusion Techniques

After pre-processing, fusion combines enhanced images into a single, informative representation. This amalgamation enhances visual quality, detail richness, and overall scene representation. Fusion maximizes improvement potential by implementing diverse

enhancement techniques, resulting in a final image that effectively communicates desired information. Techniques include:

3.2.1 Stationary Wavelet Transform (SWT)

The Stationary Wavelet Transform (SWT) is an advanced multi-resolution analysis technique used in image processing to enhance and preserve important image details. Unlike the traditional Discrete Wavelet Transform (DWT), SWT is shift-invariant, meaning it does not suffer from information loss due to down-sampling. This property makes SWT particularly useful for applications such as image denoising, enhancement, and fusion, where maintaining spatial consistency is crucial. In multi-focus image fusion, SWT effectively decomposes images into different frequency components while retaining spatial alignment, ensuring that critical details from each source image are accurately preserved and combined into a single, sharp, and well-focused composite image [25].

3.2.2 Discrete Wavelet Transform (DWT)

This approach applies wavelets to get uncertain signals because of noise, and it has proven very useful for processing X-ray images, magnetic resonance images, and applications in medical science. In this approach, it is possible to enhance the image without distortion or blurring of finer features. The DWT applied in this method gives adaptive spatial frequency resolution; that is, it improves the frequency resolution at low frequencies and enhances the spatial resolution at high frequencies. The DWT does this compactly by representing the frequency components of the signal so that the resolution is in real-time and gives better data about frequency. A better image is generated using DWT through non-duplicative graphic representation features, which further improves the spectral and spatial localization [25]. Based on the concept of multiresolution analysis, a fast technique under the title of Wavelet Transform was developed; it tends to have both temporal and frequency discrete properties. The Wavelet Transform has the property of providing a complete representation by covering both time and frequency features in an image quickly. This method decomposes the data into two sets: high-frequency and low-frequency components. Decomposition is the division process that helps to derive statistics related to low and high frequencies [26].

The wavelet transform has been one of the important logical tools developed for describing images at different resolutions. The basic idea behind what

has made wavelet transform theory possible is the mother wavelet. This idea of the mother wavelet has found applications, and many more are likely to be realized as technology advances in texture analysis, image compression, digital image watermarking, and facial recognition. The classification of wavelets comes under two kinds: continuous and discrete. This new tool influences many analytical domains, not necessarily image processing [27]. The discrete wavelet transforms (DWT) use wavelet coefficients, scale factors, and functions. The DWT may be represented as a discrete signal or image, $x[n]$, as shown by Eq. 1:

$$W(a, b) = \frac{1}{\sqrt{a}} \sum_{n=-\infty}^{\infty} x[n] \cdot \psi\left(\frac{n-b}{a}\right) \quad (1)$$

$W(a, b)$ in Eq. 1 are wavelet coefficients at scale a and translation b . $x[n]$ is the input signal or image. ψ is the symbol of the wavelet function. The frequency resolution is determined by the scale factor denoted by a . A factor b defines the position of the wavelet function and, in turn, is a translation or shift factor.

The DWT shows the input signal or image in different levels of expressions, which is done in an iterative process by applying this expression to the decomposition layers. The Wavelet Transform is a function receiving two inputs. When the scale increases, the signal's peaks become more similar, but greater translation values are necessary for the tails of the signal to resemble one another [28]. The Discrete Wavelet Transform can extract hidden changes by analysing signals and visuals from a measurable time-frequency point of view. Wavelets explore sharp changes in signals or visual [29]. In this wavelet transform method, the signal is projected onto a set of wavelet functions, which decomposes the signal into a sequence of wavelet functions. To put it another way, this transformation is a complete conversion [30].

The Discrete Wavelet Transform (DWT) includes advantages such as analyzing signals and images at various sizes and capturing localized and global information. DWT also offers multiscale analysis, effective data compression, feature extraction, and pattern recognition. One of the downsides of the Discrete Wavelet Transform (DWT) is that the fused pictures lose responsiveness to directional features [31].

3.2.3 Discrete Cosine Transform (DCT)

Discrete Cosine Transform (DCT) image fusion method is commonly used to fuse images because of its effectiveness and good properties. The utilization of DCT is needed to transform the original images into frequency domain followed by fusing the obtained DCT coefficients. Finally, the parameters are synthesized in spatial domain to obtain the fused image. One of the good properties of DCT-based fusion methods is their capability to keep high frequency information, which is important for retaining image features [32]. While it is an important transformation in various problems of science and engineering, the key thing where this concept is needed for our focus is for picture compression. DCT offers an efficient transformation, converting a picture from the spatial to frequency domain increasing its applicative area [33]. The overall framework of the proposed multi-focus image fusion method is illustrated in Figure 2, highlighting the main components and processing pipeline.

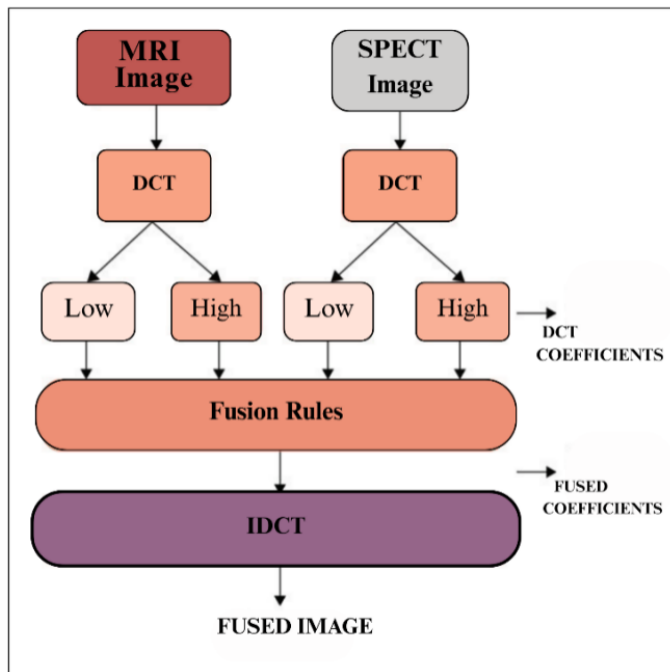


Figure 2. DCT block diagram for image fusion.

DCT is known for its computational efficiency and low memory consumption, which makes it better for image fusion applications. DCT-based fusion has a significant drawback: its sensitivity to image misalignment. When the source images aren't properly aligned, the fused image can exhibit undesirable artifacts such as ringing or blurring. For this purpose, DCT based fusion combined with image fusion techniques can be used. Before carrying out fusion, it is therefore

critical to register the source images as perfectly aligned results in a marked reduction of artefacts and hence significantly improved precision. This results in outputs of higher quality when the strengths and weaknesses of DCT-based fusion are combined properly [32].

4 Experimental Setup

This research experiment is done on Core i5 10th generation system with a ram of 16GB which uses MATLAB for implementation. MATLAB tool was employed for pre-processing, image fusion and performance evaluation.

4.1 Dataset Description

In this research experiment, we used four different sets of grayscale multifocus images. Figure 1 shows the Lytro Multifocus dataset [20]. This dataset is popular among researchers worldwide for multifocus image experiments. It includes images of clocks, balloons, leaves, and bottles.

4.2 Performance Metrics

The performance measurements employed in this study included entropy, peak signal-to-noise ratio (PSNR), mean absolute error (MAE) and root mean square error. Without these metrics, it is impossible to calculate the weight of how effectively those techniques are working in an unbiased way. These are the widely accepted metrics that gives a formal and systematic way for evaluating performance and efficiency of novel techniques.

4.2.1 Root Mean Squared Error (RMSE)

RMSE is frequently used to compare the pixel differences between original and generated images. It provides a good measure of the final image's quality by indicating how accurately the colours and brightness are represented [34]. Equation 2 shows the RMSE formula.

$$RMSE = \sqrt{\frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (F(i, j) - R(i, j))^2} \quad (2)$$

4.2.2 Peak Signal to Noise Ratio (PSNR)

The Peak Signal-to-Noise Ratio (PSNR) measures the quality of an image by comparing the original and fused images. It measures how similar the pixels are by comparing the peak signal strength to the noise

by comparing the number of matching pixels to the total number of gray levels. Equation 3 shows PSNR formula.

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right) \quad (3)$$

4.2.3 Mean Absolute Error (MAE)

The mean absolute error measures how much the pixels differ between original and resulting images. A lower MAE score suggests better image quality and a score of zero means the images match perfectly [35]. Equation 4 shows the formula for calculating the entropy.

$$\text{MAE} = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (4)$$

4.2.4 Entropy (E)

This metric separates an image's texture, appearance, and information content. It is crucial for evaluating the amount of information in the final image [2]. Equation 5 shows the formula for calculating the entropy.

$$E = - \sum_{i=0}^{L-1} p_i \log_2 (p_i) \quad (5)$$

5 Results and Discussion

This section presents the results of all the research experiments carried out in this study. We performed both qualitative and quantitative evaluation of various image fusion and enhancement techniques to provide a comprehensive interpretation of the results.

5.1 Experiment Results



Figure 3. (a) and (b) are two multifocus source images of "Clock image set".

The proposed methodology was tested on different datasets such as clocks, balloons, leaves, and bottles. The results achieved by the proposed method were

compared with image fusion techniques using the same datasets. To provide a comprehensive comparison, the results were analyzed both with and without the application of the Marr–Hildreth technique for image enhancement. Both quantitative and quantitative methodologies were used to compare the results with the proposed novel framework. This main goal of this extensive testing is to evaluate the methodology's effectiveness for multi-focus images. For quantitative analysis, evaluation metrics such as PSNR, MAE, RMSE and entropy are used in this study. All of these metrics led to a clear view of the level of the fused images that came from the proposed methodology. Furthermore, the qualitative analysis allowed visual checking of the changes, improvements, and corrections made in the fused images. The results of these performance measures will now provide a clear insight into the functionality and effectiveness of the applied image fusion techniques of the present study. The output has clearly shown benefits and efficiency for the suggested methodology in image fusion.

Figure 3 shows a set of two multifocus images, where (a) and (b) represent the original images used for the study. The fused enhanced images in Figure 4 were created by combining them using a variety of methods, including DWT, SWT, DCT, DCT+MHED, DWT+MHED, and SWT+MHED.

Table 2. Results of multifocus image fusion methods using Clock and Leaves image set.

Fusion Techniques	PSNR	RMSE	MAE	E
DWT	39.6314	2.6605	24.8771	7.4158
SWT	38.5665	3.0075	28.6275	7.4215
DWT+MHED	39.7077	2.6372	17.2962	7.4311
DCT	39.6266	2.6620	24.8633	7.4155
SWT+MHED	39.7074	2.6373	17.2950	7.4313
DCT+MHED	39.7990	2.6302	17.2633	7.4320

Table 2 compares different multi-focus image fusion methods using the Clock and Leaves image set, based on PSNR, RMSE, MAE, and entropy (E). The DCT+MHED method achieved the highest PSNR of 39.7990, meaning it produced the clearest and most detailed image. It also had the lowest MAE of 17.2633, showing minimal error, and the highest entropy of 7.4320, indicating better detail preservation. Similarly, DWT+MHED and SWT+MHED also performed well, proving that adding the Marr–Hildreth Edge Detection (MHED) technique improves the overall fusion quality.

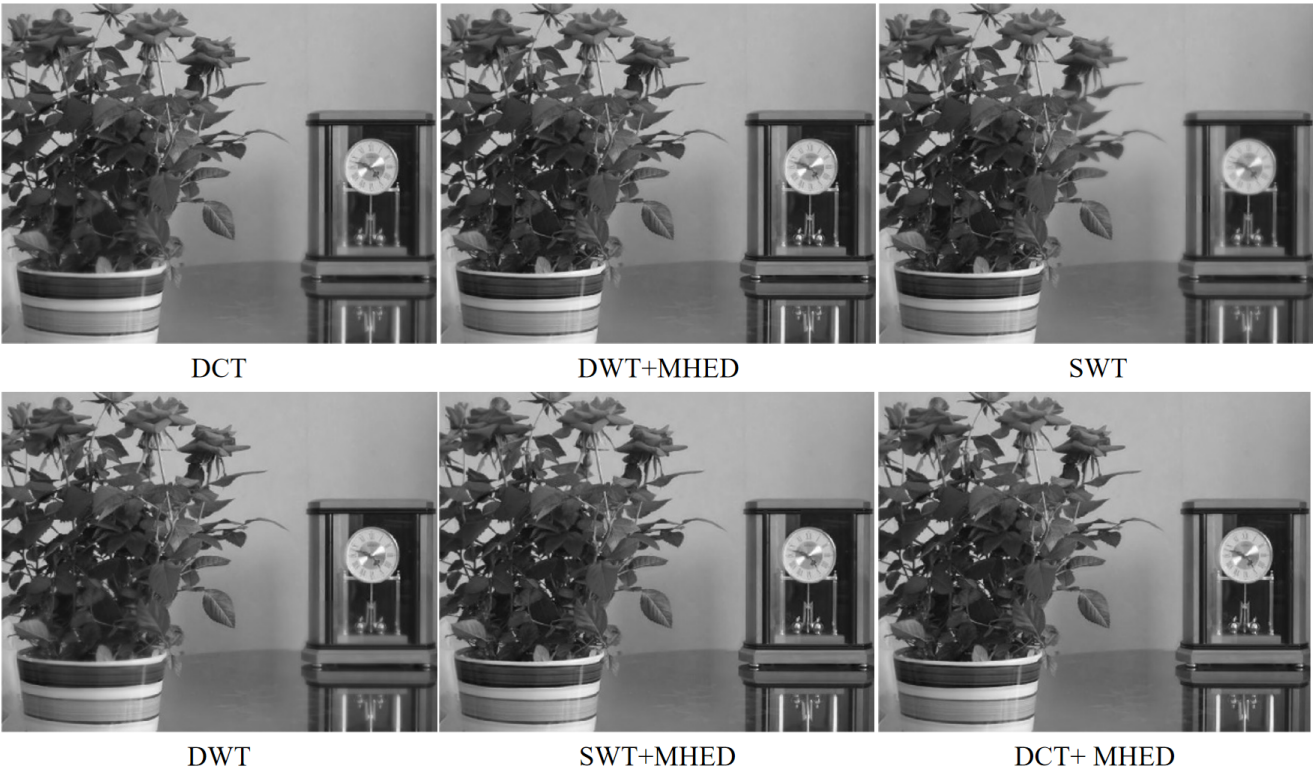


Figure 4. Fusion results using different techniques.

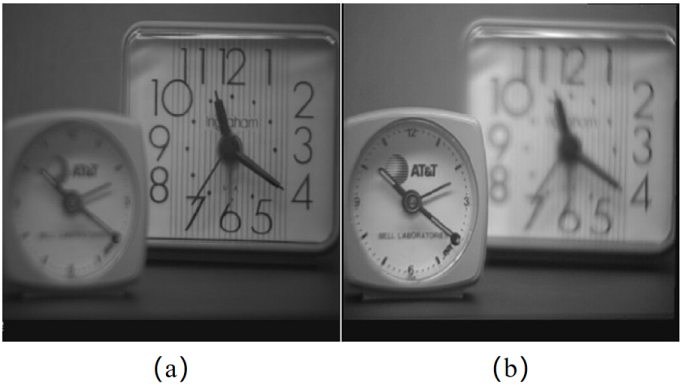


Figure 5. (a) and (b) are two multifocus source images from the another "Clock image set".

In contrast, the SWT method had the lowest PSNR of 38.5665 and the highest MAE of 28.6275, meaning it did not preserve details as effectively. Likewise, DWT and DCT without enhancement had higher RMSE values of 2.6605 and 2.6620, indicating more errors. These results clearly show that using enhancement techniques like MHED helps create sharper and more detailed fused images.

Figure 5 shows a set of two multifocus images, where (a) and (b) represent the original images used for the study. The fused enhanced images in Figure 6 were created by combining them using a variety of methods, including DWT, SWT, DCT, DCT+MHED,

DWT+MHED, and SWT+MHED.

Table 3. Results of multifocus image fusion methods using Clocks image set.

Fusion Techniques	PSNR	RMSE	MAE	E
DWT	40.7087	2.3501	19.4622	6.9863
SWT	41.4077	2.1682	18.9728	6.9942
DWT+MHED	39.7810	2.6150	15.7361	7.0041
DCT	41.4014	2.1700	18.9641	6.9898
SWT + MHED	39.7828	2.6145	15.7375	7.0085
DCT+MHED	39.7724	2.6176	15.7114	7.0041

Table 3 presents the results of different multi-focus image fusion methods using the Clocks image set, evaluated using PSNR, RMSE, MAE, and entropy (E). The SWT method achieved the highest PSNR of 41.4077, indicating that it produced the sharpest and most detailed image. Similarly, DCT closely followed with a PSNR of 41.4014, showing that it effectively preserved image quality. However, the addition of MHED (Marr–Hildreth Edge Detection) in DWT+MHED, SWT+MHED, and DCT+MHED resulted in improved entropy values, highlighting better detail retention.

On the other hand, DWT+MHED and DCT+MHED showed lower MAE values of 15.7361 and 15.7114, respectively, indicating fewer errors and better accuracy. However, DWT alone had a relatively lower

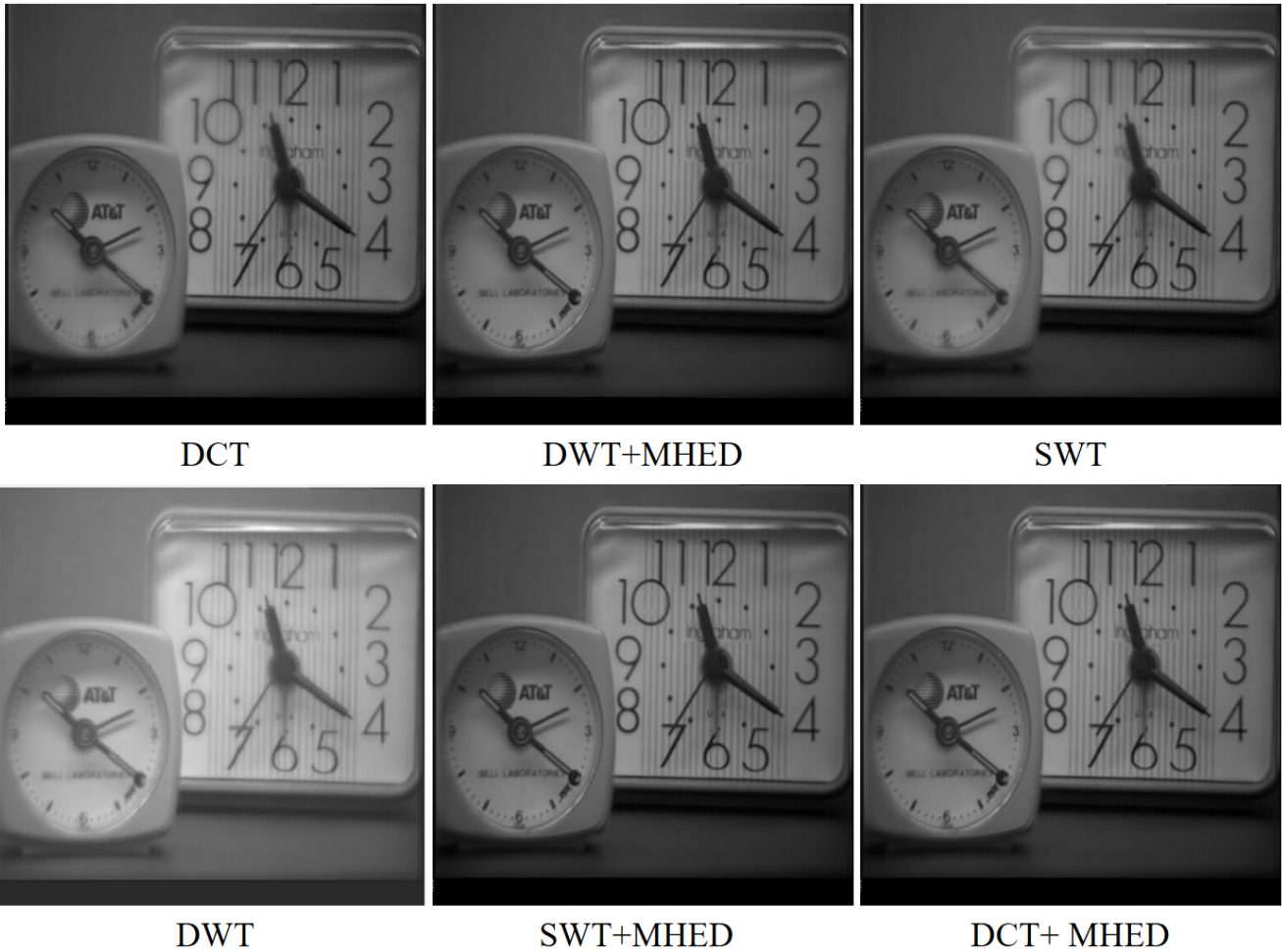


Figure 6. Fusion results through different techniques.

performance with a PSNR of 40.7087 and a higher MAE of 19.4622, suggesting that without enhancement techniques, it may not retain fine details as effectively. These results emphasize that applying enhancement techniques like MHED improves image fusion quality by refining details and reducing errors.

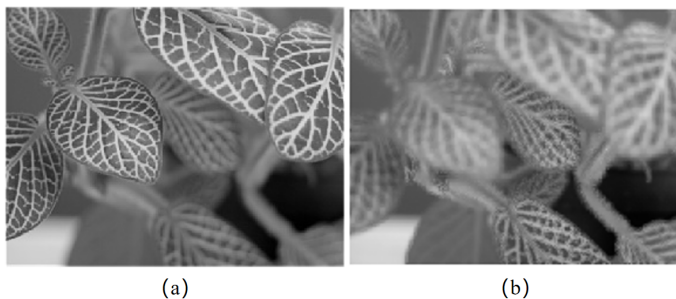


Figure 7. (a) and (b) are two multifocus source images from the "Leaves image set".

Figure 7 shows a set of two multifocus images, where (a) and (b) represent the original images used for the study. The fused enhanced images in Figure 8 were created by combining them using a variety of

methods, including DWT, SWT, DCT, DCT+MHED, DWT+MHED, and SWT+MHED.

Table 4. Results of multifocus image fusion methods using leaves image set.

Fusion Techniques	PSNR	RMSE	MAE	E
DWT	33.7621	5.2291	43.3788	7.2993
SWT	38.2289	3.1267	23.2605	7.3099
DWT+MHED	40.2055	7.3447	8.7347	7.3447
DCT	33.7615	5.2295	43.3852	7.2999
SWT + MHED	40.2078	2.4896	8.7319	7.3447
DCT+MHED	40.1937	2.4937	8.7438	7.3454

Table 4 presents the performance of different multi-focus image fusion methods using the Leaves image set, evaluated using PSNR, RMSE, MAE, and entropy (E). The SWT+MHED method achieved the highest PSNR of 40.2078, indicating superior image quality with well-preserved details. Similarly, DWT+MHED and DCT+MHED produced high PSNR values above 40, demonstrating that applying Marr–Hildreth Edge Detection (MHED) significantly improved image sharpness and clarity.

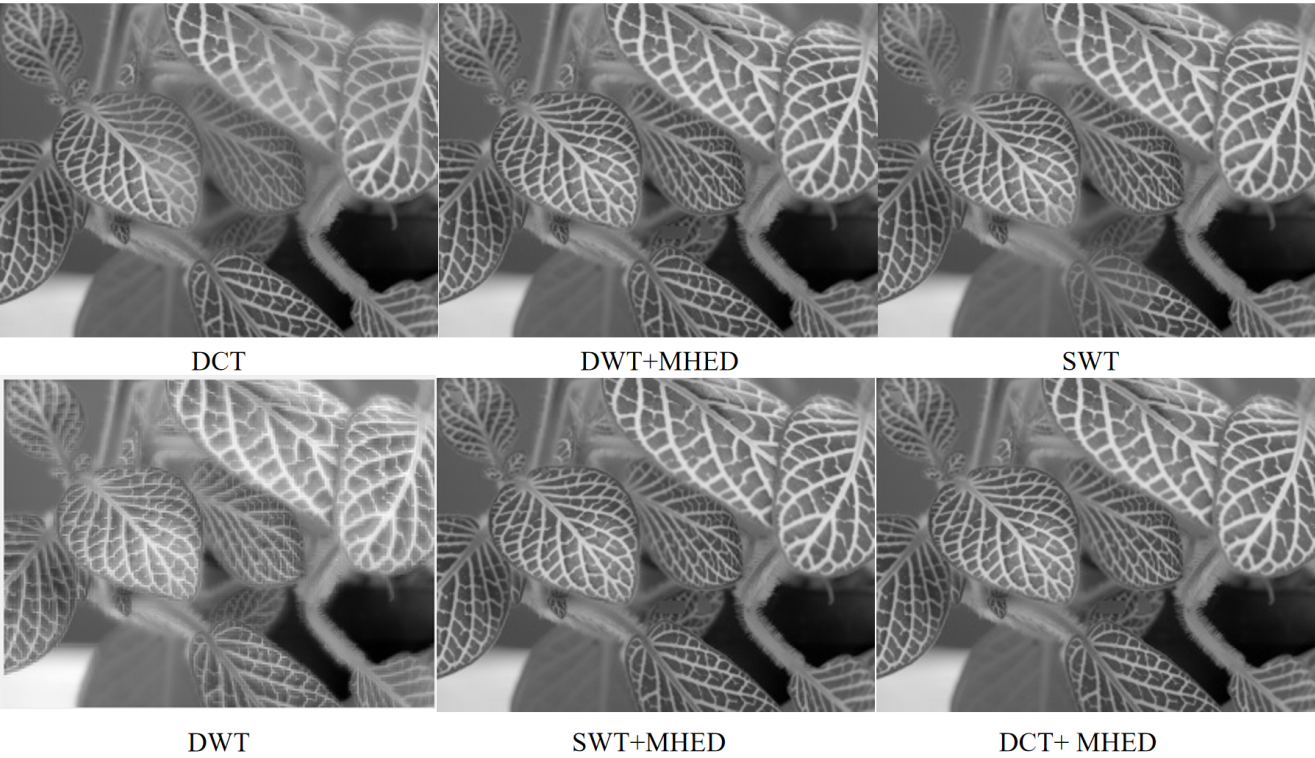


Figure 8. Fusion results through different techniques.



Figure 9. (a) and (b) multifocus images are taken from the "Balloons image set".

On the other hand, DWT and DCT without enhancement yielded lower PSNR values of 33.7621 and 33.7615, respectively, along with higher MAE values of over 43, indicating weaker detail preservation. The SWT method alone also performed well, with a PSNR of 38.2289, but integrating MHED further reduced error and enhanced fine details.

Overall, these results highlight that using enhancement techniques like MHED in combination with transformation methods leads to better fusion quality by minimizing errors and improving edge definition.

Figure 9 shows a set of two multifocus images, where (a) and (b) represent the original images used for the study. The fused enhanced images in Figure 10 were created by combining them using a variety of

methods, including DWT, SWT, DCT, DCT+MHED, DWT+MHED, and SWT+MHED.

Table 5. Results of multifocus image fusion methods using Balloons image set.

Fusion Techniques	PSNR	RMSE	MAE	E
DWT	37.5477	3.3818	31.7169	7.4465
SWT	41.1553	2.2834	12.2743	7.1480
DWT+ MHED	41.3035	2.1946	9.9225	7.4472
DCT	37.5437	3.3833	31.71515	7.4474
SWT + MHED	41.3010	2.1952	9.9227	7.4482
DCT+MHED	41.2940	2.1970	9.9274	7.4498

Table 5 presents the results of different multi-focus image fusion techniques applied to the Balloons image set, evaluated using PSNR, RMSE, MAE, and entropy (E). The DWT+MHED method achieved the highest PSNR of 41.3035, indicating superior image quality with well-preserved details. Similarly, SWT+MHED and DCT+MHED also performed well, achieving PSNR values above 41, demonstrating that integrating Marr–Hildreth Edge Detection (MHED) significantly improved sharpness and clarity.

On the other hand, DWT and DCT without enhancement yielded lower PSNR values of 37.5477 and 37.5437, respectively, with higher MAE values exceeding 31, suggesting weaker edge preservation. The SWT method alone also performed well, but

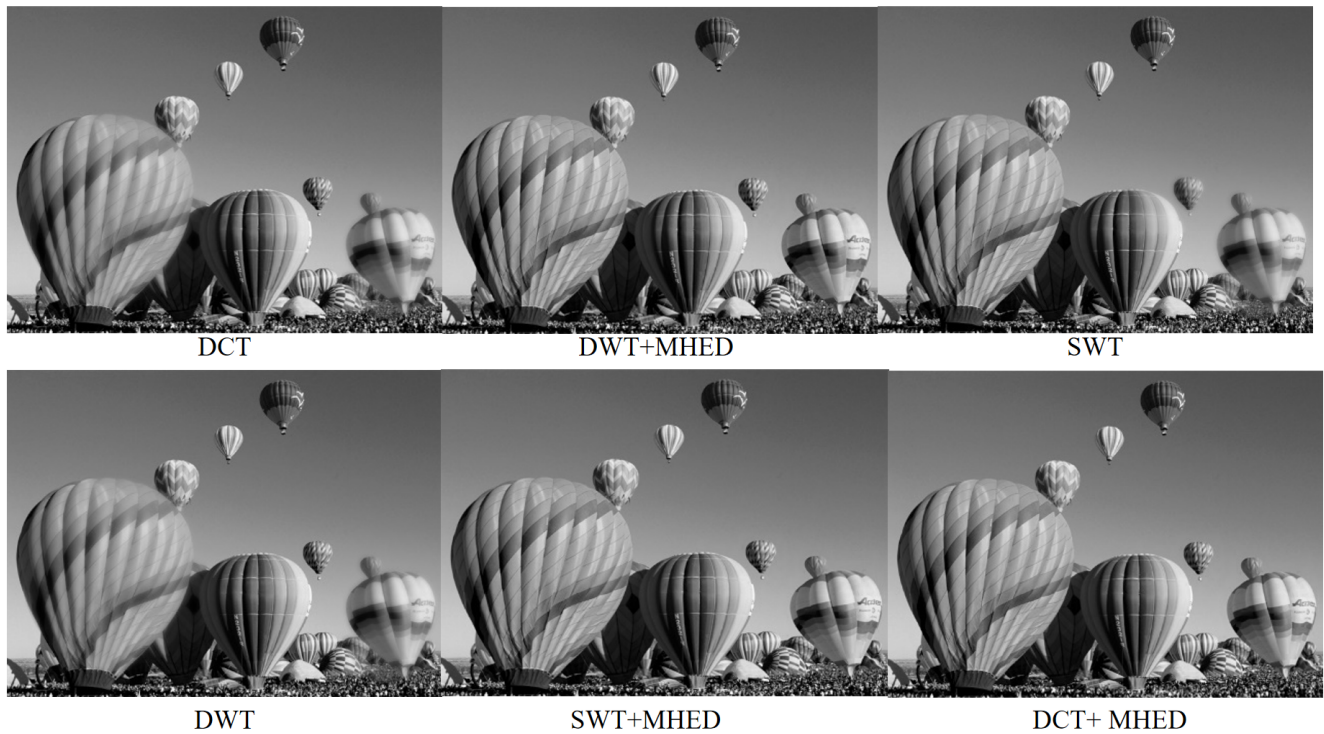


Figure 10. Fusion results through different techniques.

incorporating MHED further reduced errors and enhanced fine details.

These results highlight the effectiveness of combining transformation techniques with edge enhancement methods like MHED, which significantly improve the quality of fused images by reducing errors and enhancing clarity.

5.2 Significance of the Study

This paper presents a significant development in image processing, demonstrating the advantages of integrating multi-focus image fusion techniques based on the Marr–Hildreth technique, DCT, SWT, and DWT. The resulting clarity and detail can enhance results in medical imaging, remote sensing, and photography. The structured methodology provides a reliable way to produce high-quality fused images, providing further analysis and visualization. These findings lay a foundation for future studies to refine and expand these methods, potentially making a pioneering contribution to image fusion technology.

5.3 Limitations

This work demonstrates the performance of versatile image fusion technology based on the Marr–Hildreth technique, DCT, SWT, and DWT with enhancement and without prior masking. It is limited to grayscale images from the Lytro dataset. The application to other types of images or datasets has not been

explored. Additionally, while DCT, SWT, and DWT have shown promising results, other pre-processing and fusion methods were not investigated in this research. The experiments were conducted in a controlled MATLAB environment, which may differ from real-world applications. Future research could expand on these areas to further validate and enhance the findings.

6 Conclusion

This research study suggested a novel technique for multi-focus image fusion using the Marr–Hildreth technique, Discrete Cosine Transform, Stationary Wavelet Transform, and Discrete Wavelet Transform. The experimental results show that both qualitative and quantitative evaluations have significantly improved. Furthermore, when applied to the multi-focus image datasets, the proposed technique works well. The results are significant for image fusion and enhancement in the field, revealing the multiplicity of ways researchers approach solving multi-focus image challenges. It also provides a better understanding of how to develop image fusion and processing research in the new horizon of advanced scientific investigation, an emerging domain.

The research limitations are that the study focuses only on grayscale images from the Lytro dataset using DCT, SWT, and DWT-based multi-focus image fusion

techniques in MATLAB, without considering other types of images, datasets, or different pre-processing and fusion methods. Future work will expand on this by testing the proposed techniques on a broader range of image types and datasets, including color images. Additionally, exploring other pre-processing and fusion methods could further enhance image quality. Implementing and evaluating these techniques in real-world scenarios would help validate their practical applications. Integrating advanced machine learning algorithms could also offer new ways to optimize the image fusion process, leading to greater improvements in clarity and detail. In conclusion, while Marr-Hildreth Edge Detection (MHED) is typically used for edge detection, in this study, it served a unique role as a pre-processing step. Its primary function was to enhance the sharpness and clarity of the input images before fusion, which proved crucial in improving the overall quality of the fused images. This demonstrates the versatility of MHED in image processing beyond its traditional use.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

Ethical Approval and Consent to Participate

Not applicable.

References

- [1] Luhtakallio, E., Meriluoto, T., & Malafaia, C. (2024). Visual politicization and youth challenges to an unequal public sphere: Conceptual and methodological perspectives. In *Handbook on youth activism* (pp. 140-153). Edward Elgar Publishing. [\[Crossref\]](#)
- [2] Zhang, X. (2021). Deep learning-based multi-focus image fusion: A survey and a comparative study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9), 4819-4838. [\[Crossref\]](#)
- [3] Zhou, Y., Yu, L., Zhi, C., Huang, C., Wang, S., Zhu, M., ... & Fu, S. (2022). A survey of multi-focus image fusion methods. *Applied Sciences*, 12(12), 6281. [\[Crossref\]](#)
- [4] Zhou, Z., Dong, M., Xie, X., & Gao, Z. (2016). Fusion of infrared and visible images for night-vision context enhancement. *Applied optics*, 55(23), 6480-6490. [\[Crossref\]](#)
- [5] Singh, S., Singh, H., Bueno, G., Deniz, O., Singh, S., Monga, H., ... & Pedraza, A. (2023). A review of image fusion: Methods, applications and performance metrics. *Digital Signal Processing*, 137, 104020. [\[Crossref\]](#)
- [6] Li, H., Wu, X. J., & Kittler, J. (2018, August). Infrared and visible image fusion using a deep learning framework. In *2018 24th international conference on pattern recognition (ICPR)* (pp. 2705-2710). IEEE. [\[Crossref\]](#)
- [7] Liu, Y., Wang, L., Cheng, J., Li, C., & Chen, X. (2020). Multi-focus image fusion: A survey of the state of the art. *Information Fusion*, 64, 71-91. [\[Crossref\]](#)
- [8] Begum, M., Ferdush, J., & Uddin, M. S. (2022). A Hybrid robust watermarking system based on discrete cosine transform, discrete wavelet transform, and singular value decomposition. *Journal of King Saud University-Computer and Information Sciences*, 34(8), 5856-5867. [\[Crossref\]](#)
- [9] Basu, S., Singhal, S., & Singh, D. (2025). Multi-Focus Image Fusion: A Systematic Literature Review. *SN Computer Science*, 6(2), 150. [\[Crossref\]](#)
- [10] Mishra, P. K., Patni, J. C., Saini, S., Kishore, J., Baloni, D., & Verma, S. (2024, March). Adaptive Gamma Correction With Unsharp Masking and Gaussian Filter for Image Contrast Enhancement. In *2024 International Conference on Automation and Computation (AUTOCOM)* (pp. 439-444). IEEE. [\[Crossref\]](#)
- [11] Liu, S., Ma, J., Yang, Y., Qiu, T., Li, H., Hu, S., & Zhang, Y. D. (2022). A multi-focus color image fusion algorithm based on low vision image reconstruction and focused feature extraction. *Signal Processing: Image Communication*, 100, 116533. [\[Crossref\]](#)
- [12] Li, S., Kang, X., & Hu, J. (2013). Image fusion with guided filtering. *IEEE Transactions on Image processing*, 22(7), 2864-2875. [\[Crossref\]](#)
- [13] Turgut, S. S., & Oral, M. (2022). Multi-focus image fusion based on gradient transform. *arXiv preprint arXiv:2204.09777*. [\[Crossref\]](#)
- [14] Kaur, R., & Singh, S. (2023, March). Multi-focus image fusion methods: A review. In *International Conference on Advanced Computing, Machine Learning, Robotics and Internet Technologies* (pp. 112-125). Cham: Springer Nature Switzerland. [\[Crossref\]](#)
- [15] Pan, X., Zhai, H., Yang, Y., Chen, L., & Li, A. (2024). Improving multi-focus image fusion through Noisy image and feature difference network. *Image and Vision Computing*, 142, 104891. [\[Crossref\]](#)
- [16] You, C. S., & Yang, S. Y. (2022). A simple and effective multi-focus image fusion method based on local standard deviations enhanced by the guided filter. *Displays*, 72, 102146. [\[Crossref\]](#)
- [17] Farid, M. S., Mahmood, A., & Al-Maadeed, S.

- A. (2019). Multi-focus image fusion using content adaptive blurring. *Information fusion*, 45, 96-112. [Crossref]
- [18] Bhat, S., & Koundal, D. (2021). Multi-focus image fusion using neutrosophic based wavelet transform. *Applied Soft Computing*, 106, 107307. [Crossref]
- [19] Wang, Z., Li, X., Duan, H., Zhang, X., & Wang, H. (2019). Multifocus image fusion using convolutional neural networks in the discrete wavelet transform domain. *Multimedia Tools and Applications*, 78, 34483-34512. [Crossref]
- [20] Nejati, M., Samavi, S., & Shirani, S. (2015). Multi-focus image fusion using dictionary-based sparse representation. *Information fusion*, 25, 72-84. [Crossref]
- [21] Tribuana, D., & Arda, A. L. (2024). Image preprocessing approaches toward better learning performance with cnn. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 8(1), 1-9. [Crossref]
- [22] Shah, P., Merchant, S. N., & Desai, U. B. (2013). Multifocus and multispectral image fusion based on pixel significance using multiresolution decomposition. *Signal, Image and Video Processing*, 7, 95-109. [Crossref]
- [23] Fu, Y., Wu, X. J., & Durrani, T. (2021). Image fusion based on generative adversarial network consistent with perception. *Information Fusion*, 72, 110-125. [Crossref]
- [24] Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 207(1167), 187-217. [Crossref]
- [25] Danyal, M. M., Khan, S., Khan, R. S., Jan, S., & Rahman, N. (2024). Enhancing Multi-Modality Medical Imaging: A Novel Approach with Laplacian Filter+ Discrete Fourier Transform Pre-Processing and Stationary Wavelet Transform Fusion. *J. Intell. Med. Healthc*, 2, 35-53.
- [26] Joshi, K., Kirola, M., Chaudhary, S., Diwakar, M., & Joshi, N. K. (2019, March). Multi-focus image fusion using discrete wavelet transform method. In International conference on advances in engineering science management & technology (ICAESMT)-2019, Uttaranchal University, Dehradun, India.
- [27] Haribabu, M., Bindu, C.H.: Multi focus image fusion based on discrete wavelet transform with statistical measurements (2022).
- [28] Alwan, I. M. (2018). Multi-Focus Image Fusion Using Discrete Cosine Harmonic Wavelet Transform. *Int. J. Sci. Res.*, 7(1), 835-841.
- [29] Avci, D., Sert, E., Özyurt, F., & Avci, E. (2024). MFIF-DWT-CNN: Multi-focus image fusion based on discrete wavelet transform with deep convolutional neural network. *Multimedia Tools and Applications*, 83(4), 10951-10968. [Crossref]
- [30] Tan, L., Chen, Y., & Zhang, W. (2019, August). Multi-focus Image Fusion Method based on Wavelet Transform. In *Journal of Physics: Conference Series* (Vol. 1284, No. 1, p. 012068). IOP Publishing. [Crossref]
- [31] Nanmaran, R., Srimathi, S., Yamuna, G., Thanigaivel, S., Vickram, A. S., Priya, A. K., ... & Muhibbullah, M. (2022). Investigating the role of image fusion in brain tumor classification models based on machine learning algorithm for personalized medicine. *Computational and Mathematical Methods in Medicine*, 2022(1), 7137524. [Crossref]
- [32] Singh, J., Banga, V.K.: An integrated approach for image fusion using pca and dct. *Journal Name* (2019).
- [33] Bhat, S., & Koundal, D. (2021). Multi-focus image fusion techniques: a survey. *Artificial Intelligence Review*, 54(8), 5735-5787. [Crossref]
- [34] Khan, S. S., Ran, Q., Khan, M., & Ji, Z. (2019, December). Pan-sharpening framework based on laplacian sharpening with Brovey. In *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)* (pp. 1-5). IEEE. [Crossref]
- [35] Rajaei, A., Abiri, E., & Helfroush, M. S. (2024). Balanced spatio-spectral feature extraction for hyperspectral and multispectral image fusion. *Computers and Electrical Engineering*, 118, 109391. [Crossref]



Mian Muhammad Danyal received his Bachelor's degree in Software Engineering from City University of Science and Technology, Peshawar, Pakistan, in 2023. He is currently pursuing an MS in Computer Science from the Institute of Management Sciences, Peshawar. His research interests include sentiment analysis, image processing, opinion mining, text summarization, machine learning, and deep learning. (Email: miandaniyal502@gmail.com)



Omar Bin Samin I hold a Master's degree in Engineering (Electronics) from Nottingham Trent University, UK, and I am currently pursuing a PhD in Information and Communication Technology at Al-Madinah International University, Kuala Lumpur, Malaysia. With approximately 14 years of experience in both industry and academia, I have developed expertise in Artificial Intelligence (AI), Internet of Things (IoT) automation, and the development of smart systems. (Email: omar.samin@imsiences.edu.pk)



Dr. Sarwar Shah Khan is a Lecturer in the Department of Computer Science at the University of Engineering and Technology. He completed his Ph.D. in December 2024, specializing in the field of medical image processing. With seven years of experience in teaching and research, Dr. Sarwar has published over 38 research articles in reputable journals and conferences. His research interests extend to machine learning

applications in image processing, and he has developed frameworks that integrate artificial intelligence and natural language processing for various applications. Currently, he is actively pursuing research in two major areas: Data Science (including text processing, natural language processing, and machine learning) and Digital Image Processing. His contributions have significantly enriched these domains. (Email: sskhan0092@gmail.com)



Sajid Khan is a dedicated academican, researcher, entrepreneur, and lecturer at the Department of Computer Science, UET Mardan, with over a decade of industrial experience. He combines a passion for advancing technology with a commitment to teaching and innovation. His research interests include machine learning (ML), Image processing, sentiment analysis, text summarization, and agile development

methodologies.

As a software developer, Sajid has developed expertise in building scalable solutions and has contributed to numerous impactful projects. His entrepreneurial spirit drives him to explore emerging technologies, mentor future leaders in the tech industry, and align academic efforts with industry needs.