



## Research Article

# A new classification-based approach for multi-focus image fusion

Samet AYMAZ<sup>1,\*</sup>, Şeyma AYMAZ<sup>1</sup>, Cemal KÖSE<sup>1</sup>

<sup>1</sup>Department of Computer Engineering, Karadeniz Technical University, Trabzon, 61080, Türkiye

## ARTICLE INFO

### Article history

Received: 13 February 2022

Revised: 25 April 2022

Accepted: 15 June 2022

### Keywords:

Classification; Focus Metrics;

Image Fusion; Scoring; SML;

WAVV; SF

## ABSTRACT

Multi-focus image fusion combines two or more images of the same scene with different focus points to create a single detailed fully-focused image. The primary purpose of multi-focus image fusion methods is to transfer the correct focus information from the source images to the fused image. This study proposes a new classification mechanism based on focus metrics. This mechanism is designed to classify focused, non-focused and ambiguous regions. The most important feature of the proposed mechanism is that it can detect ambiguous areas and transfer these regions to the fused image correctly. Firstly, each source image is split into non-overlapping image patches of specified sizes in this study. Then, the generated image patches are classified using created classification mechanism. After the classification process, a decision map is created for each source image. These decision maps are then refined using morphological operations. In the final stage of the designed study, a dynamic fusion rule is proposed. This fusion rule transfers focused and non-focused pixels to the fused image according to a specific rule. In contrast, ambiguous regions, frequently encountered in transitions from focused to non-focused areas, are transferred to the fused image using the gradient-based fusion rule. In this way, the negative effect of the regions that the classification algorithms classified incorrectly on the fused image is reduced. In addition, in this study, the impact of image size on image fusion success is analyzed by using different image sizes in the classification mechanism. As a result, the proposed study is evaluated using objective and subjective metrics. The evaluations show that the proposed method is suitable for achieving multi-focus image fusion purposes.

**Cite this article as:** Aymaz S, Aymaz Ş, Köse C. A new classification-based approach for multi-focus image fusion. Sigma J Eng Nat Sci 2024;42(1):11–25.

## INTRODUCTION

In digital images, some problems arise since imaging devices can focus at limited depths. These problems complicate the analysis of images. Multi-focus image fusion methods are used to eliminate such issues. These methods

bring together meaningless pieces of the puzzle to create meaningful, easy-to-examine images. For this reason, these methods have an important place these days when digital imaging is becoming widespread. Multi-focus image fusion methods can be used in different areas. Health, combining images taken from space, combining images taken during

### \*Corresponding author.

\*E-mail address: sametaymaz5@gmail.com

This paper was recommended for publication in revised form by Regional Editor Ahmet Selim Dalkilic



the day and night are applications of this area. In short, these methods are widely used in almost every field of digital imaging.

Digital imaging is vital for experts in different fields to make the right decisions. Detailed, easy-to-examine, and near-perfect images are needed to make the right decisions. Multi-focus image fusion methods aim to achieve this perfect image by transferring the correct information from the source images to the fused image and transferring the critical information (edge, corner, etc.) from the source images to the combined images. Many approaches are suggested in the literature to achieve this aim. These approaches are generally studied in two classes: spatial and transformation domains. In recent years, classification-based approaches have been added to these approaches.

Spatial domain methods work by directly addressing pixel brightness levels. These methods can be used as block-based or pixel-based. The block-based methods allow processing by decomposing the image according to specific block size. It is faster than pixel-based methods but weaker at preserving details of source images.

On the other hand, pixel-based methods are based on finding the importance of each pixel for the fused image. These methods are slower but allow for more detailed fused image creation. In the literature, there are different studies in the spatial domain. The first of these is proposed by Li et al. [1]. It is a block-based approach. In this approach, images are split into fixed-size blocks. Then, the spatial frequency metric is calculated for these blocks, and these values are transferred to the fused image according to a specific threshold value. Another approach is applied by Huang et al. [2]. In this approach, the image is divided into blocks, and focus metrics are calculated for these blocks. Finally, the fused image is created with the fusion rule based on the focus metrics. In addition, methods created using the block-based Pulse-Coupled Neural Network (PCNN) method are available in the literature [3,4]. In recent years, pixel-based methods have been preferred more than block-based methods in the spatial domain. These methods are frequently used in the literature to provide more detailed fused image creation. The focus weight of each pixel is found and transferred to the fused image in these methods. Random-walk-based methods [5,6], Guided filtering-based methods [7-11], Bilateral filtering-based methods, Markov-random field-based methods [12], Independent Component Analysis (ICA)-based methods, and Sparse Representation (SR)-based methods [13] are the critical methods used in this field.

In addition to spatial-based methods, transform-based methods are also frequently used for multi-focus image fusion. Since transform-based methods are suitable for separating essential and unimportant parts of the image, flexible approaches can be created easily. The most critical approaches in this field are wavelet-based and pyramid-based approaches. The first pyramid-based approach is applied by Burt et al. [14] in this field. In this study, the

fusion process is made using the Laplace pyramid. Another approach based on the Laplace pyramid is suggested by Jin et al. [15]. In this approach, Laplace components are combined with the help of a PCNN. Again, for the Laplace features, the fusion of images is made with the area tessellation process by Kou et al. [16]. In addition to pyramid-based approaches, methods based on Wavelet transform also find their place in the literature. Tian et al. [17] propose a wavelet-based approach. In this approach, the components formed after the wavelet transform are combined by statistical sharpness. Again, Aymaz et al. [18] combine Stationary Wavelet Transform (SWT) and Principal Component Analysis (PCA) methods to make a compelling fusion. In addition, there are PCNN-based [19], Support Vector Machine (SVM)-based [20,21], Fisher classifier-based, and Fuzzy sets-based [22,23] approaches applied to wavelet components in the literature.

In addition to these studies, approaches based on classification have also been found in the literature in recent years. These approaches are generally learning-based. The important thing here is how accurately the classification can be made. Classification in multi-focus image fusion is usually based on classifying image blocks as focused or non-focused. There are primarily deep learning-based approaches for multi-focus image fusion in the literature. Although learning-based methods are successful in classification, they also have difficulties preparing the training set, creating the network, determining the hyperparameters, etc. There are different approaches based on the classification of image patches in the literature. Some of these approaches are the Two-stage learning-based approach [24], Deep Convolutional Neural Network (CNN)-based approach [25], Generative Adversarial Network (GAN)-based approach [26], CNN-based approach [27], Ensemble CNN-based approach [28], Fuzzy CNN-based approach [29], Connection and dilated CNN-based approach [30], and Self-supervised mask-optimization-based approach [31].

The proposed study creates a scoring-based classification mechanism using focus metrics based on image blocks' spatial and transformative properties. This approach requires no training. In addition, the difficulties found in learning-based approaches such as parameter determination and creating networks are not found in the proposed approach. In this approach, source images are divided into non-overlapping blocks. Focus metrics are calculated for these blocks, and a score is added to the image block with the enormous metric value. As a result, the image block with a significant score is labelled focused, and the small one is labelled non-focused. In addition to these, the transition regions from focused regions to the non-focused areas, which are ignored in the literature and where classification approaches often make mistakes, are also considered in this study. In other words, a three-class classification process is performed. The regions labelled with focus are directly transferred to the fused image. In contrast, areas belonging

to the ambiguous class are transferred to the fused image using the gradient-based fusion rule. Thus, the errors caused by the classifier are minimized. Details of the proposed method will be given in other sections. The features that distinguish the proposed method from the methods in the literature are:

- A new classification approach based on focus metrics
- Addressing ambiguous areas as well as focused and non-focused areas
- Analyzing the effect of different image sizes on the fusion
- Application of two different fusion rules with dynamic fusion rule.

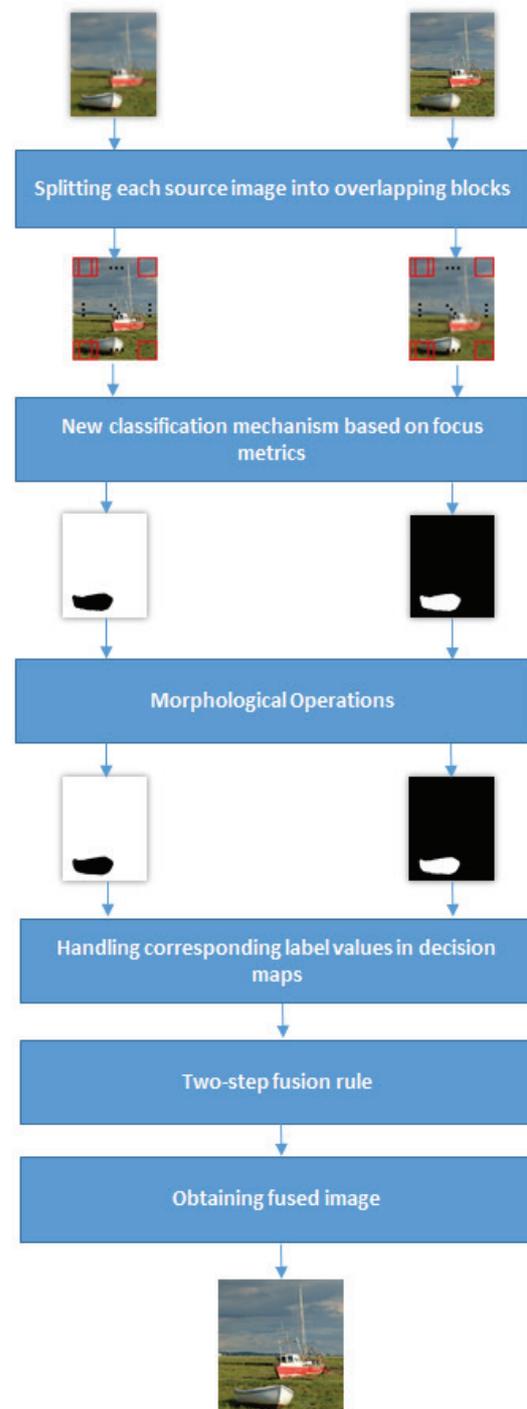
## MATERIALS AND METHODS

This study proposes a new classification-based approach to multi-focus image fusion. Unlike classical learning-based techniques, an innovative classification process based on focus metrics is applied in the proposed system. The Sum-modified Laplacian (SML), Variance of Wavelet (WAVV), and Spatial Frequency (SF) are chosen as focus metrics. As the first step of the method, source images are divided into overlapping image blocks. Classifications are usually carried out using a fixed image block size in the literature. The proposed work is evaluated for different image block sizes, which are 8x8, 16x16, and 32x32. Thus, the effect of block sizes on fusion is shown. Details of the methods designed for the proposed study will be presented in the following sections

In addition, the two-step fusion rule is applied in the proposed study to minimize classification errors. In the regions where the classifiers can clearly distinguish the focused and non-focused areas, the corresponding pixel values in the source images are multiplied by the corresponding label value and transferred directly to the fused image. For indefinite parts, a gradient-based fusion rule is applied. After the proposed study, highly detailed, near-perfectly combined images are obtained. Finally, objective and subjective metrics are used in performance evaluation. Evaluation and comparison results show that the proposed approach applies multi-focus image fusion successfully. Figure 1 shows the flow of the proposed method.

### The New Classification Mechanism

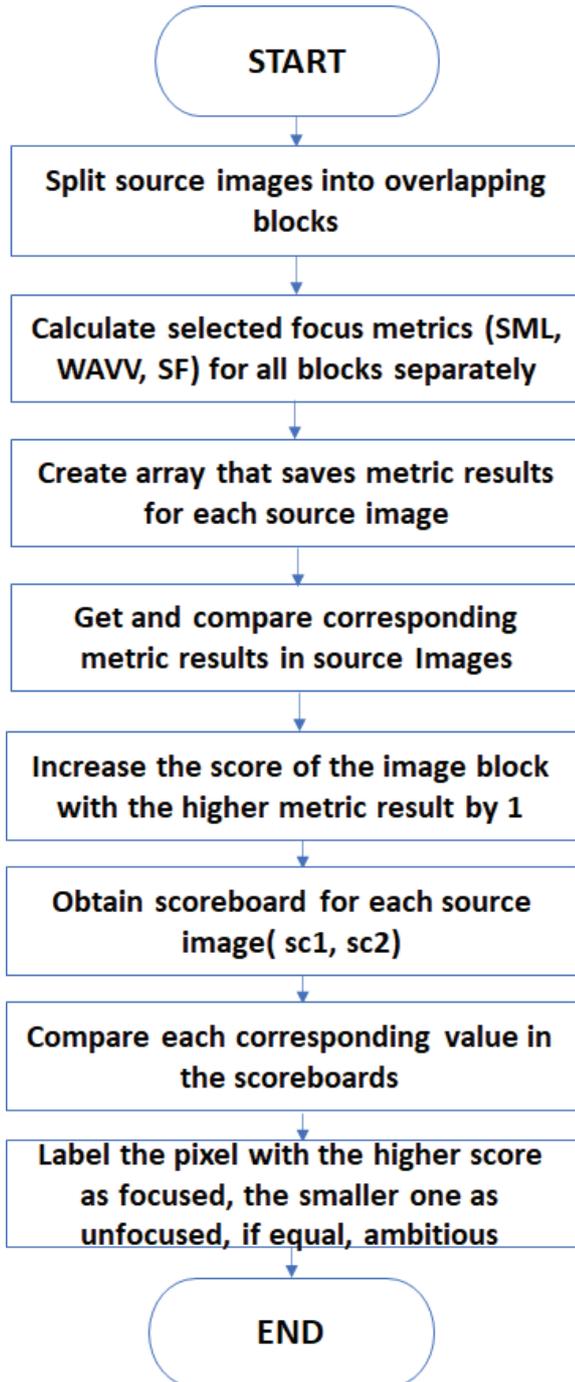
This study proposes a new classification mechanism based on focus metrics that do not require training, unlike the learning-based classification methods in the literature. In this mechanism, three metrics are chosen that highlight the spatial and transformative properties of the image blocks. The number of metrics can be increased, but since this method will adversely affect the working times, it is ensured that operations are performed with the minimum number of metrics. Selected metrics are SML, which brings edge information to the fore, WAVV for calculating frequency components and focus information, and SF for



**Figure 1.** The flow of the proposed method.

extracting important points in images. Since focused regions contain more detail than non-focused regions, they can easily reach focus information using these metrics. In addition, the proposed mechanism also addresses the ambiguous areas that are important for the fused image, which are ignored in the methods in the literature. Classifiers often misclassify (same label value for each block) for the

transition from focused to non-focused regions. This situation negatively affects the success of the methods. In the proposed approach to improve methods in the literature, a third class, the ambiguous class, is put forward. By applying the remarkable fusion rule for these regions, perfectly combined images can be created. The operation of the proposed mechanism for two source images is as follows step by step in Figure 2.



**Figure 2.** The flow chart of the new classification mechanism.

### Focus Measurement Metrics

Focus measurement metrics measure the degree of focus of blocks in images [32]. It also gives a degree of clarity to the images. Therefore, they are vital in classifying an image block as focused or non-focused. For this study, three focus metrics with different properties are selected. These metrics are WAVV, SML, and SF. Selected metrics reveal both spatial and transformational properties of image blocks. In addition, these metrics effectively highlight important parts of images, making them very suitable for multi-focus fusion purposes.

*SML*: Sum-modified Laplacian metric is suggested by Nayar et al.[33]. This metric is based on the Laplace operator with quadratic derivatives and highlights essential parts of images. The metric calculation is illustrated with Equations (1-2).

$$\nabla_{ML}^2 f(x, y) = |2f(x, y) - f(x - step, y) - f(x + step, y)| + |2f(x, y) - f(x, y - step) - f(x, y + step)| \quad (1)$$

$$SML = \sum_{i=x-N}^{i=x+N} \sum_{j=y-N}^{j=y+N} \nabla_{ML}^2 f(i, j) \quad \nabla_{ML}^2 f(i, j) \geq T \quad (2)$$

*SF*: The SF metric is calculated based on the image gradient. Gradient takes higher values in important image parts, such as edges, corners, etc. This analysis makes it easy to find focused regions. Therefore, it is well suited for multi-focus compositing purposes. Equations (3-5) give the calculation of the SF metric. [34].

$$SF = \sqrt{RF^2 + CF^2} \quad (3)$$

In Equations (4-5), RF and CF represent the row and column frequencies. RF shows important information (edge, corner, etc.) in rows of images, while CF shows important information in columns of images.

$$RF = \sqrt{\frac{1}{M \times N} \sum_{x=1}^M \sum_{y=2}^N [f(x, y) - f(x, y - 1)]^2} \quad (4)$$

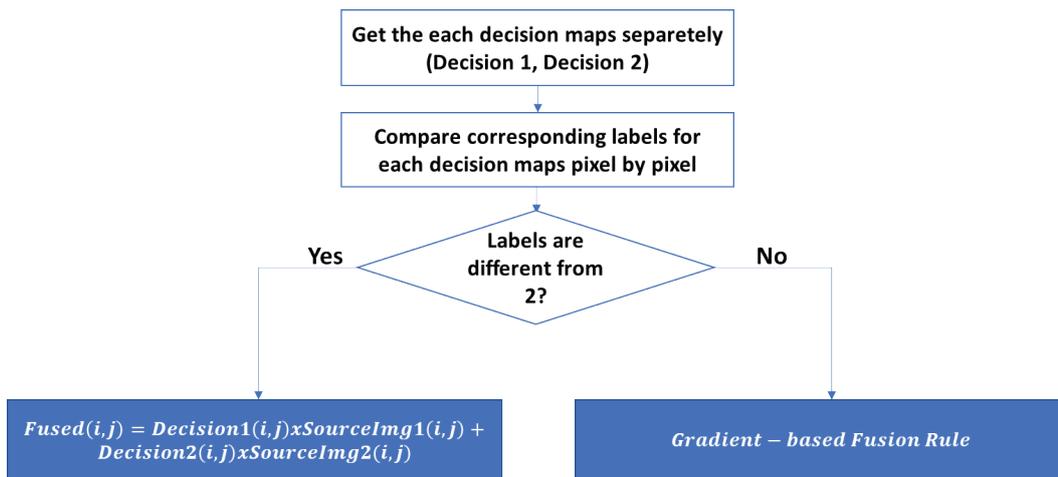
$$CF = \sqrt{\frac{1}{M \times N} \sum_{x=2}^M \sum_{y=1}^N [f(x, y) - f(x - 1, y)]^2} \quad (5)$$

In these equations, M and N give the row and column sizes of source images, respectively. Also, f is the original signal.

*WAVV*: Transformation-based methods facilitate the analysis of images. Wavelet transform is frequently preferred in the literature. This metric tries to reach the focus information by using the frequency components of the image parts.

### Two-Step Fusion Rule

In this study, an innovative fusion approach is used. In classification-based methods for multi-focus image fusion,



**Figure 3.** The flow chart of the two-step fusion rule.

the fused images are created according to the classifier's decision. This situation causes even a tiny error in classification to affect the combined image adversely. A two-step fusion approach is designed in the proposed study to turn this negative situation into a positive one. The classifier classifies the image blocks as focused, non-focused, and ambiguous. In the regions where the classifiers can clearly distinguish between focused and non-focused, the relevant decision map label value and the source image pixel value are multiplied and transferred to the fused image. The main problematic parts of the images are the transition points from the focused regions to the non-focused areas. At these points, classifiers often make mistakes. Therefore, in the proposed study, these parts are ambiguous and transferred to the fused image using the gradient-based fusion rule. The steps of the two-step fusion rule are shown in figure 3.

The most significant feature that distinguishes the proposed study from the studies in the literature is the presence of ambiguous pixels. The images of the ambiguous pixels obtained for the Boat image as a result of the proposed method are given in Figure 4.

When Figure 4 is examined, it is seen that ambiguous pixels appear in the transitions from focused regions to defocused regions in source images. In addition, it can be seen that the number of these pixels increases as the block size increases.

### Gradient-based Fusion Rule

Multi-focus image fusion methods aim to transfer the correct focus and essential information (edge, corner, etc.) from the source images to the fused image. A gradient is a tool that highlights vital information of images that can best accomplish these purposes. This tool takes high values in crucial parts of the image (edge, corner, etc.) and low values in unimportant areas such as the background. In this study, a gradient-based fusion rule is applied, which is based on gradient analysis and determines the importance ratios

of each pixel for the merged image. This rule is used only when the classifier is unstable. Thus, detailed fused images are obtained. The step-by-step implementation of the gradient-based fusion rule is as follows;

**Step 1:** This fusion rule works when the classifier gives label 2. First, the gradient magnitudes of each pixel in the source images are calculated using the 3x3 Sobel filter. The filter size is chosen small because every detail in the source images is significant for the fused image. Then, maps containing gradient magnitudes are obtained for each source image ( $Gmag1$ ,  $Gmag2$ ).

**Step 2:** Then, the gradient magnitudes are proportioned, and each pixel's importance ratios for the fused image are determined. Transferring a pixel directly to the fused image is risky. It is essential to transfer the pixels to the combined image by weighting to reduce this risk. The weights of the pixels are calculated using Equation (6) and Equation (7).

$$ImpRate1(i,j) = \frac{Gmag1(i,j)}{Gmag1(i,j) + Gmag2(i,j)} \quad (6)$$

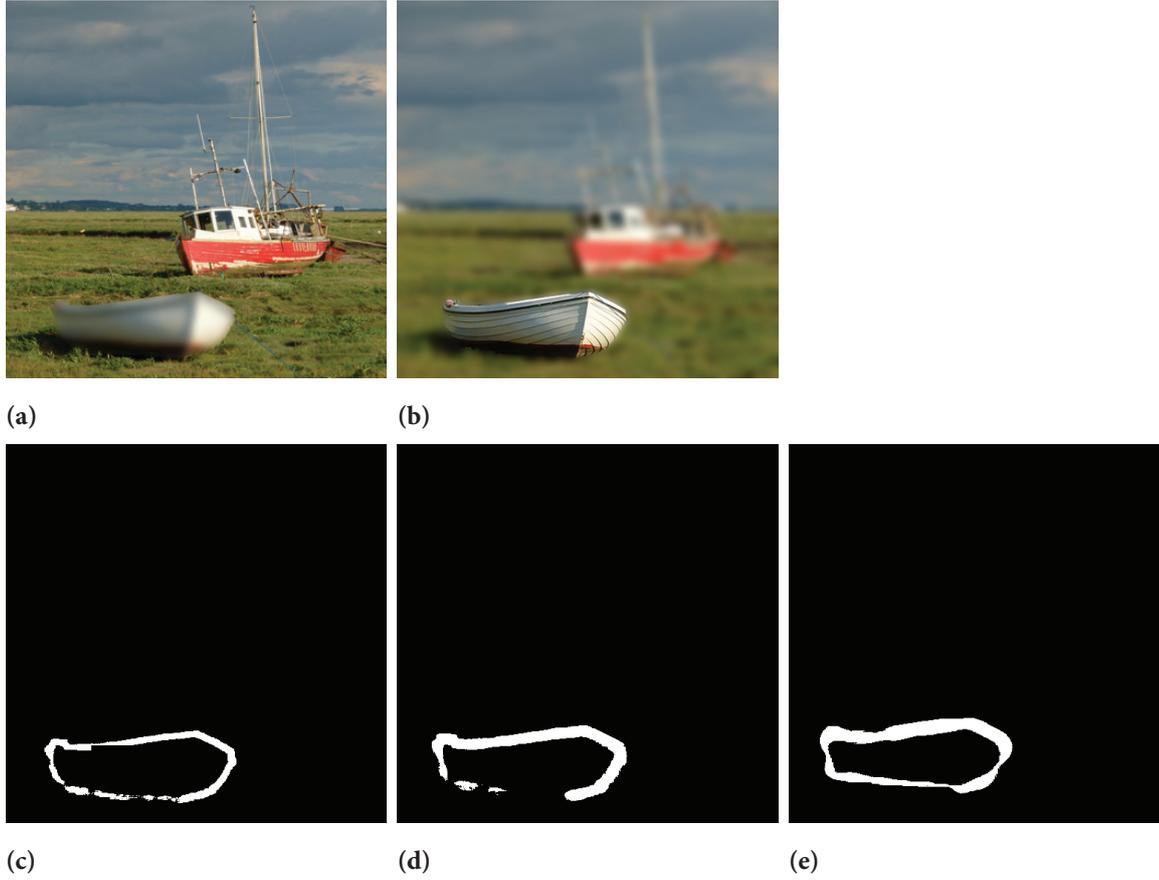
$$ImpRate2(i,j) = 1 - ImpRate1(i,j) \quad (7)$$

**Step 3:** After the importance ratios of pixels in source images are calculated, the ambiguous regions are transferred to the combined image according to their importance ratios using Equation (8).

$$Fused(i,j) = ImpRate1(i,j) \times SourceImg1(i,j) + ImpRate2(i,j) \times SourceImg2(i,j) \quad (8)$$

### Performance Measurement Metrics

Performance measurements are made in two different ways in multi-focus image fusion methods. These are objective and subjective metrics.



**Figure 4.** Multi-focus Boat image, a) Right-focused source image, b) Left-focused source image, Ambiguous pixels after proposed classification method with c) 8x8 block size, d) 16x16 block size, e) 32x32 block size.

### Objective Metrics

QAB/F and Mutual Information (MI) metrics are the most popular objective metrics. These metrics are used when we do not have a reference image available. Since the datasets used in the literature generally do not contain original images, researchers use these metrics more frequently. With the help of these metrics, the correct amount of data transferred from the source images to the combined image can be measured.

*QAB/F Metric:* The QAB/F metric measures the edge information transferred from the source images to the fused image. It is calculated using Equations (9-13) [35].

$G^{AF}(n, m)$  and  $A^{AF}(n, m)$  show the relative strength and orientation values of input images A to F. These values are calculated using Equations (9-10).

$$G^{AF}(n, m) = \begin{cases} \frac{g_F(n, m)}{g_A(n, m)}, & \text{if } g_A(n, m) > g_F(n, m) \\ \frac{g_A(n, m)}{g_F(n, m)}, & \text{otherwise} \end{cases} \quad (9)$$

$$A^{AF}(n, m) = \frac{|\alpha_A(n, m) - \alpha_F(n, m)| - \frac{\pi}{2}}{\frac{\pi}{2}} \quad (10)$$

The model  $Q_g^{AF}(n, m)$  and  $Q_a^{AF}(n, m)$  give the edges between the A and F images. Some constant values are used in these equations. The constants are  $\Gamma_g$ ,  $\kappa_g$ ,  $\sigma_g$  and  $\Gamma_a$ ,  $\kappa_a$ ,  $\sigma_a$ . These constants are the exact form of the sigmoid functions used to generate the edge strength values. The calculations of these values are given in Equations (11-12).

$$Q_g^{AF}(n, m) = \frac{\Gamma_g}{1 + e^{\kappa_g(G^{AF}(n, m) - \sigma_g)}} \quad (11)$$

$$Q_a^{AF}(n, m) = \frac{\Gamma_a}{1 + e^{\kappa_a(A^{AF}(n, m) - \sigma_a)}} \quad (12)$$

As a result, Equation (13) measures the amount of edge transferred from the source images to the fused image.

$$Q^{AB/F} = \frac{\sum_{i=1}^M \sum_{j=1}^N (Q^{AF}(i, j)w^A(i, j) + Q^{BF}(i, j)w^B(i, j))}{\sum_{i=1}^M \sum_{j=1}^N (w^A(i, j) + w^B(i, j))} \quad (13)$$

In Equation (13), the source images and the fused image are represented by A, B, and F, respectively. The edge relationship between the images A and F is represented by  $Q^{AF}$ , While the edge relationship represents the B and F images

$Q^{BF}$ . The weights  $w^A(i, j)$  and  $w^B(i, j)$  represent the required weights for  $Q^{AF}$  and  $Q^{BF}$ . In addition, the higher the value of this metric, the more successful the method.

*MI*: The Mutual Information metric measures the correct amount of information transferred from the source images to the fused image. The metric is calculated using Equations (14-15).

$$I_{AF} = \sum_{a,f} P_{AF}(a, f) \log \frac{P_{AF}(a, f)}{P_A(a)P_F(f)} \quad (14)$$

$$I_{BF} = \sum_{b,f} P_{BF}(b, f) \log \frac{P_{BF}(b, f)}{P_B(b)P_F(f)} \quad (15)$$

In Equations (14-15), the  $I_{AF}$  and  $I_{BF}$  values indicate the amount of information between images A and B and between images B and F, respectively. In addition, the P in the equations represents the entropy information.

$$MI = I_{AF} + I_{BF} \quad (16)$$

As a result, the total information transferred from source images to fused images is represented with MI.

### Subjective Metrics

The most popular subjective metrics are Root Mean Square Error (RMSE) and Peak Signal Noise Ratio (PSNR). These metrics are used when we have a reference image. It is used to measure the similarity between the fused image and the original image.

*RMSE*: The RMSE metric measures the similarity of the fused image resulting from the method and the original image from the dataset. It is calculated using the metric Equation (17) [36]. The smaller the result of this metric, the more successful the method.

$$RMSE = \sqrt{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (I_R(i, j) - I_F(i, j))^2} \quad (17)$$

In Equation (17), M and N show the row and column of images, respectively. In addition, the reference image and fused image are represented with  $I_R$  and  $I_F$ , respectively.

*PSNR*: The PSNR metric shows how noise-resistant the fused image is compared to the original image. The higher the metric result, the more successful the proposed method is [36].

$$PSNR = 10 \log_{10} \left( \frac{L^2}{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (I_R(i, j) - I_F(i, j))^2} \right) \quad (18)$$

In Equation (18), M and N show the row and column of images, respectively. In addition, the reference image and fused image are represented with  $I_R$  and  $I_F$ , respectively. In this Equation, L shows the maximum grey level.

## RESULTS AND DISCUSSION

This study proposes a new approach for multi-focus image fusion. This approach includes a new classification mechanism that does not require training and a flexible two-step fusion rule. In addition, the proposed study with different sizes of blocks is tried to show the effect of image block sizes on fusion success. These dimensions are chosen as 8x8, 16x16, and 32x32. Evaluations for each dimension are made separately. Very few datasets exist for multi-focus image fusion methods in the literature. The most popular of these datasets is the Lytro dataset [37], consisting of 20 color images, and the Durga Prasad dataset [38], grey-level images. In the proposed study, evaluations are made for the samples taken from both datasets. While making the evaluations, objective and subjective metrics are used.

Objective metrics measure how much the correct focus information is transferred from the source images to the fused image, and essential information such as edges and corners of the source images are preserved without the need for a reference image. This paper uses the objective metrics QAB/F and MI, popularly used in the literature. These metrics effectively measure how well multi-focus image fusion goals are achieved. First, evaluations and visual results with images taken by Durga Prasad [38] will be given. The samples taken from this dataset are compared with nine different methods in the literature. The first of these methods is proposed by Chen et al. [39]. They used Multiwavelet transform to divide source images into frequency components. Then, the SML metrics of these components are calculated, and the larger values are transferred to the combined image. The second one is Hua et al.'s [40] method. They applied the random walks method to fuse images. In this study, focus metrics using local and global properties of images are used as a fusion rule. Zhang et al.'s [41] method classifies the focused regions with the Graf-based Visual Saliency method, and then the parts are smoothed by morphological operations. Also, the Shearlet transform is used for defocused areas, while the focused regions are transferred directly to the fused image in this method. Another approach is implemented by Li et al. [42]. First, the source images are divided into frequency components with the Nonsubsampled Counterlet Transform (NSCT) method. Then, the Energy of Gradient (EOG) metric for these components is calculated and combined according to the mean and the maximum rules. Nejati et al. [43] designed a new focus measure that handles the intersections of source images, and focused regions are separated from the defocused areas.

Another approach in this area is proposed by Jiang et al. [23]. They offered a method based on SWT and fuzzy sets. Abdipour et al. [44] first split source images into sub-frequency components. Then, these components' variance and spatial frequency information are calculated and combined. Chaudrary et al. [45] proposed a block-based approach. This approach divides the source images

into two subcomponents with neighbourhood filtering. Then, these subcomponents are combined using spatial frequency metrics, and a fused image is created. The last of the methods is recommended by Yang et al. [46]. They implement a Counterlet transform to divide source images into two subbands. While low subbands are transferred to the fused image with the PCNN method, the maximum SML metric is transferred to the combined image for high-frequency subbands. The proposed approach compared with these methods and comparison results are given in Table 1.

The methods given in Table 1 use either spatial domain or transform domain methods to fuse source images. The proposed approach is superior to these methods as it can transfer the focus information directly to the fused image without any coefficients. The disadvantage of these studies is finding constant coefficients that accurately transfer the focus information to the fused image. When the evaluations are examined, it is seen that the proposed approach is objectively superior to the methods in the literature. In addition, the proposed method results related to these images are given in the figures. The visual results

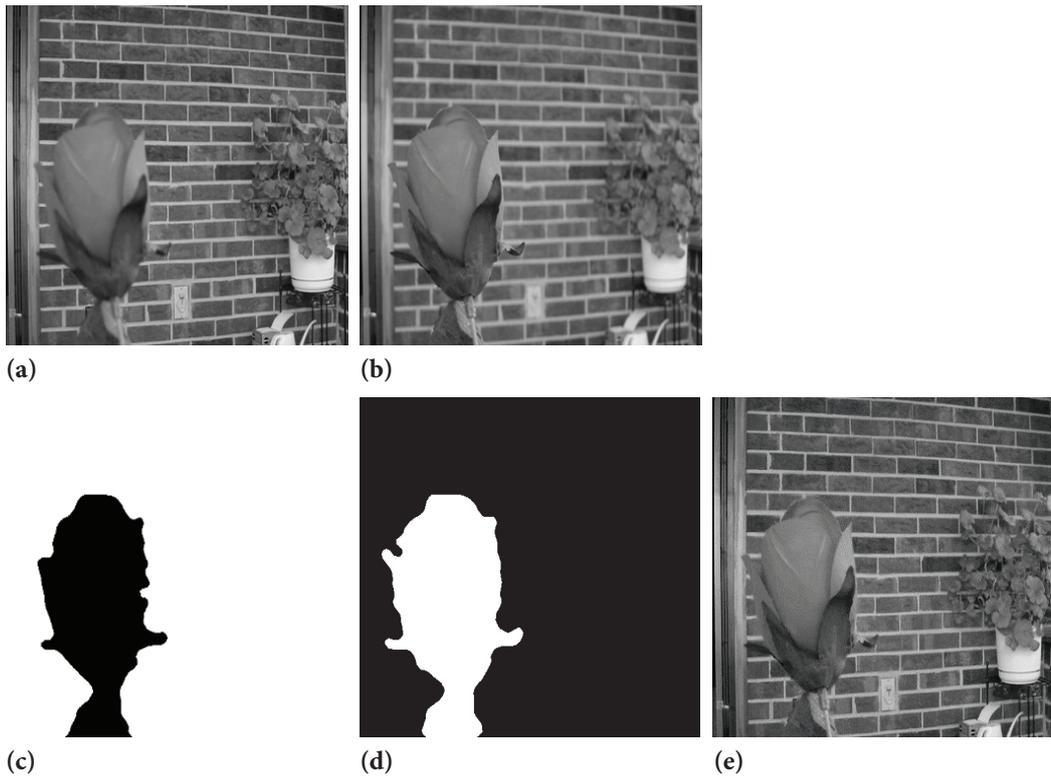
of the proposed method for the dataset prepared by Durga Prasad [38] are presented in this section. Visual results of Flower and Leaf images are given in Figures 5 and 6, respectively.

Figures 5-6 present the visual results of the proposed study on the dataset prepared by Durga Prasad. This dataset consists entirely of grey-level images. Figures 5-6 are created with a similar logic. The a and b images in the figures give source images with defocused regions taken from the relevant dataset. Since the proposed study focuses on classification into focused and non-focused areas, images showing classification success are added to the figures. After the proposed classification mechanism and morphological processes, give decision maps with c and d images. In addition, the fused image formed after the proposed study is shown with the e image. As can be seen from the figures, the proposed approach successfully creates decision maps and perfectly fused images with the help of these decision maps.

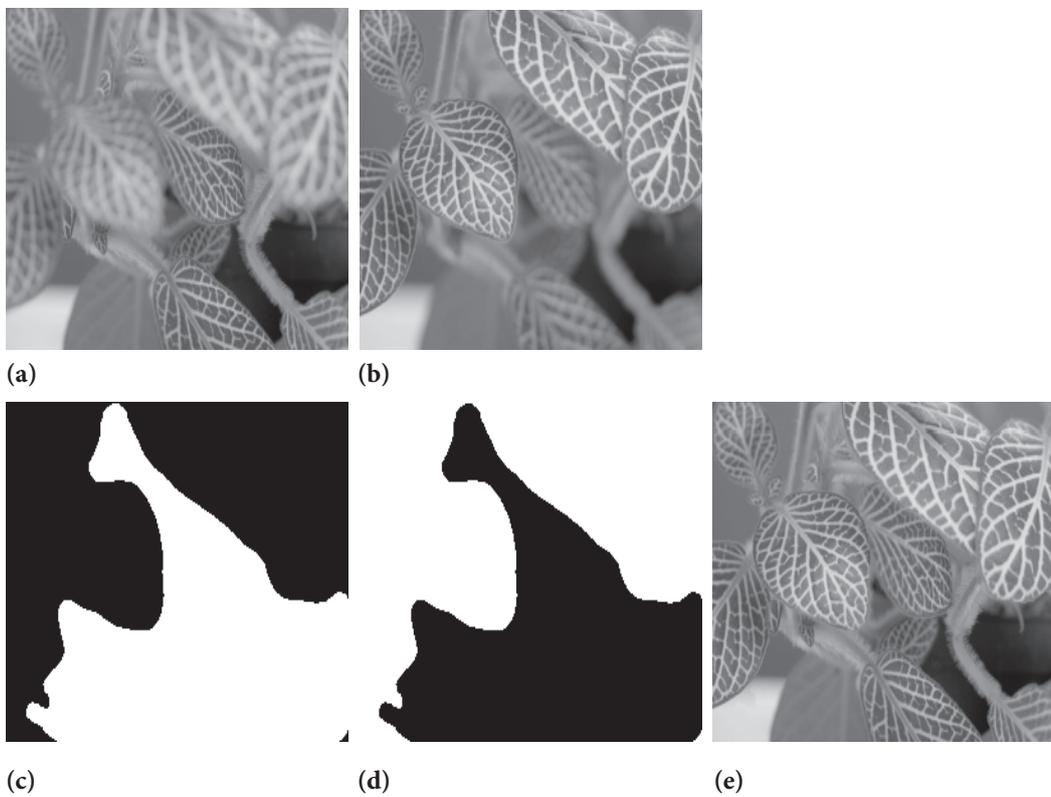
Secondly, subjective evaluations are made for some samples taken from the dataset prepared by Durga Prasad [38]. Since it is difficult to find reference images, such metrics

**Table 1.** Objective Comparison Results for Durga Prasad [38] Dataset

	Images	BOOK	CLOCK	PEPSİ	LAB	DİSK	LEOPARD
	Metrics						
Chen et al. [39]	MI	8,61	X	7,42	8,10	7,66	X
	Q <sup>AB/F</sup>	0.71	X	0.75	0.73	0.71	X
Hua et al. [40]	MI	9,24	8,29	7,80	8,50	8,00	X
	Q <sup>AB/F</sup>	0.73	0.73	0.76	0.74	0.73	X
Zhang et al. [41]	MI	X	7,83	7,09	8,04	6,49	X
	Q <sup>AB/F</sup>	X	0.71	0.78	0.73	0.68	X
Li et al. [42]	MI	9,63	8,52	8,83	8,78	X	X
	Q <sup>AB/F</sup>	0.71	0.68	0.76	0.73	X	X
Nejati et al. [43]	MI	9,44	8,58	8,87	8,70	8,29	X
	Q <sup>AB/F</sup>	0.73	0.72	0.78	0.74	0.73	X
Jiang et al. [23]	MI	8,96	8,38	8,25	8,78	7,75	X
	Q <sup>AB/F</sup>	0.73	0.71	0.78	0.73	0.72	X
Abdipour et al. [44]	MI	9,27	7,16	8,90	8,69	X	X
	Q <sup>AB/F</sup>	0.73	0.65	0.79	0.75	X	X
Chaudhary et al. [45]	MI	X	X	7,07	6,26	6,45	X
	Q <sup>AB/F</sup>	X	X	0.76	0.68	0.71	X
Yang et al. [46]	MI	9,34	8,65	X	X	8,38	10,90
	Q <sup>AB/F</sup>	0.76	0.74	X	X	0.73	0.83
<b>Proposed Method (with 8x8 Block Size)</b>	<b>MI</b>	<b>9.97</b>	<b>10.25</b>	<b>11.48</b>	<b>12.44</b>	<b>8.74</b>	<b>13.03</b>
	<b>Q<sup>AB/F</sup></b>	<b>0.95</b>	<b>0.87</b>	<b>0.91</b>	<b>0.88</b>	<b>0.86</b>	<b>0.92</b>
<b>Proposed Method (with 16x16 Block Size)</b>	<b>MI</b>	9,78	10,09	11,43	12,37	8,21	12,82
	<b>Q<sup>AB/F</sup></b>	0.94	0.88	0.91	0.88	0.86	0.92
<b>Proposed Method (with 32x32 Block Size)</b>	<b>MI</b>	9,29	9,90	11.25	12,10	8,08	12,24
	<b>Q<sup>AB/F</sup></b>	0.93	0.89	0.90	0.87	0.85	0.91



**Figure 5.** Multi-focus Flower image. (a) Right-focused source image, (b) Left-focused source image, (c) Decision map for right-focused image After proposed classification method, (d) Decision map for left-focused image after proposed classification method and (e) Fused image.



**Figure 6.** Multi-focus Leaf image. (a) Right-focused source image, (b) Left-focused source image, (c) Decision map for right-focused image after proposed classification method, (d) Decision map for left-focused image after proposed classification method, and (e) Fused image.

are not preferred in the literature. In this study, RMSE and PSNR metrics, which are subjective, are used. The proposed method is compared with three different studies in the literature using these metrics.

When Table 2 is examined, it is seen that the proposed approach obtains the closest fused images to the original images. In addition, as the image block sizes increase, it is seen that the success of the method decreases as it becomes more challenging to find the transition points from the focused regions to the defocused areas.

Finally, the proposed approach is objectively compared with classification-based techniques. This comparison is meaningful as the proposed approach also includes a classification mechanism. Comparisons are made with nine different studies applied in the literature in recent years. Firstly, Liu et al. [25] proposed a system based on detecting focused regions by deep CNN. Secondly, Ma et al. [26] used GAN architecture to detect focused and defocused areas. Another approach using deep learning is suggested by Amin-Naji et al. [28]. This approach uses an ensemble-based CNN. Zhang et al. [27] proposed a different approach based on CNN. The first two convolution

layers are used to extract features, and these features are combined according to the max, min rule. Liu et al. [49] applied an adaptive sparse-based approach. This approach creates dictionaries for source images, and coefficients are made for each source image with the K-SVD algorithm. If the corresponding coefficient is higher, it is transferred to the fused vector, and the fused image is obtained. Gai et al. [24] proposed a two-stage CNN structure. In the first stage, images are separated as focused or non-focused with Densenet. In Phase 2, EDGAN architecture is applied to protect the edges in source images. In Bai et al.'s [52] method, source images are divided into subcomponents with a Quadtree structure, and these subcomponents are classified as focused or non-focused using the SML metric. Then, the results are improved by morphological operations, and a fused image is obtained. Zhang et al. [53] found the transition between the focused and non-focused regions and combined them. Gradient information is used when finding a boundary. Finally, Liu et al. [51] applied a transform and sparse-based approach. First, the source images are separated into low and high-frequency components using the transform-based method.

**Table 2.** Subjective Comparison Results for Durga Prasad [38] Dataset

Methods	Book Image		Clock Image		Flower Image		Saras Image	
	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR	RMSE	PSNR
Moushmi et al. [47]	7.04	31.17	4.51	35.04	X	X	2.85	39.03
Li et al. [48]	X	X	X	X	7.84	30.24	X	X
Aymaz et al. [49]	5.28	33.67	<b>0.95</b>	<b>48.57</b>	4.99	34.16	2.76	39.31
<b>Proposed Method (with 8x8 Block Size)</b>	<b>0.41</b>	<b>55.87</b>	2.47	40.27	<b>1.16</b>	<b>46.84</b>	<b>0.70</b>	<b>51.22</b>
<b>Proposed Method (with 16x16 Block Size)</b>	0.62	52.28	2.63	39.73	1.58	44.15	0.73	50.86
<b>Proposed Method (with 32x32 Block Size)</b>	1.02	47.95	2.74	39.37	1.68	43.62	0.76	50.51

**Table 3.** Objective Comparison Results for Lytro Dataset [37]

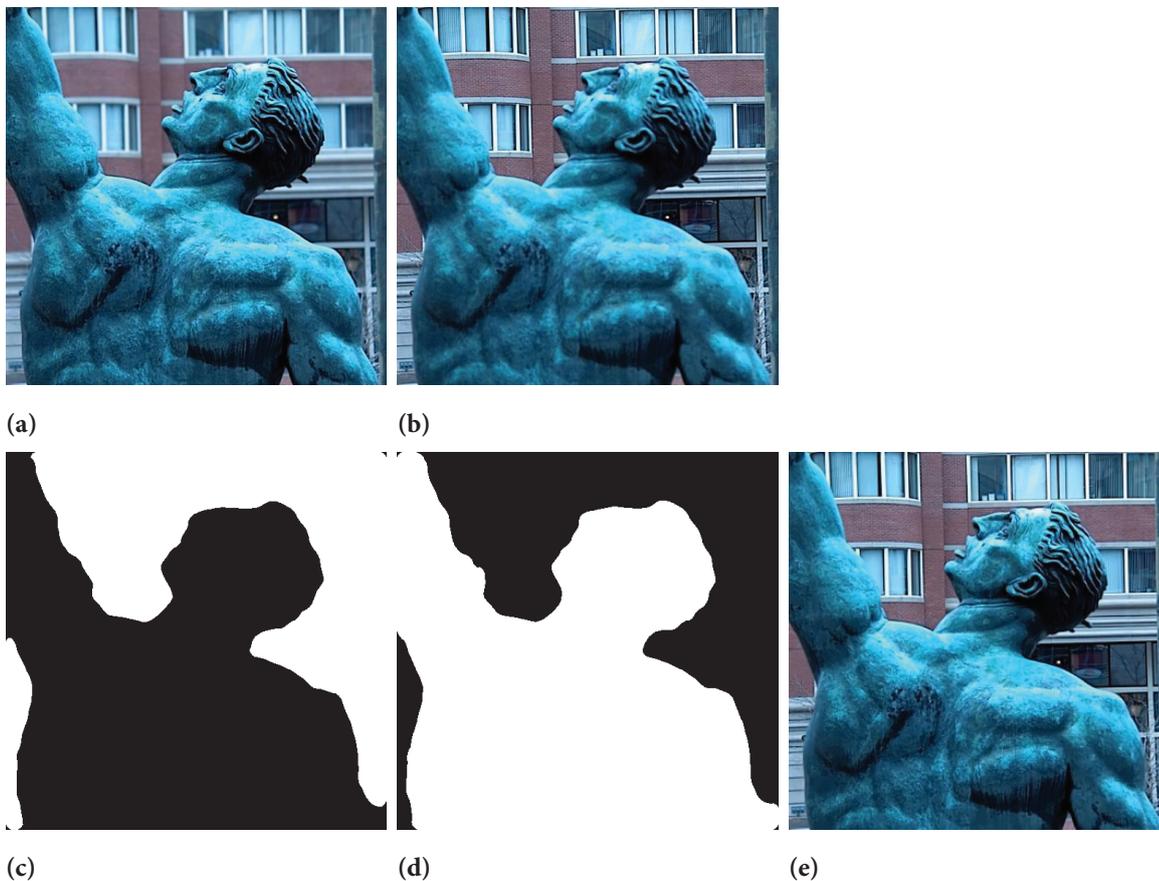
Method	Mean for Lytro Images	MI	$Q^{A/B/F}$	STD
DCNN [25]	Mean	5.96	0.76	57.46
FusionGAN [26]	Mean	3.50	0.37	48.35
ECNN [28]	Mean	6.15	0.75	57.51
ASR [40]	Mean	4.91	0.74	56.84
IFCNN [27]	Mean	4.87	0.7296	57.55
TWOSTAGE [24]	Mean	6.12	0.7621	57.57
MST-SR [51]	Mean	5.15	0.7483	57.42
QB [52]	Mean	5.50	0.7446	57.53
BF [53]	Mean	6.07	0.7583	57.29
<b>Proposed Method (8x8 Block Size)</b>	Mean	<b>9.24</b>	<b>0.898</b>	<b>57.94</b>
<b>Proposed Method (16x16 Block Size)</b>	Mean	9.01	0.896	57.38
<b>Proposed Method (32x32 Block Size)</b>	Mean	8.78	0.885	56.98

Low-frequency information is combined with a sparse-based system, while high-frequency information is combined with the maximum selection rule. The proposed approach compared with these methods and comparison results are given in Table 3.

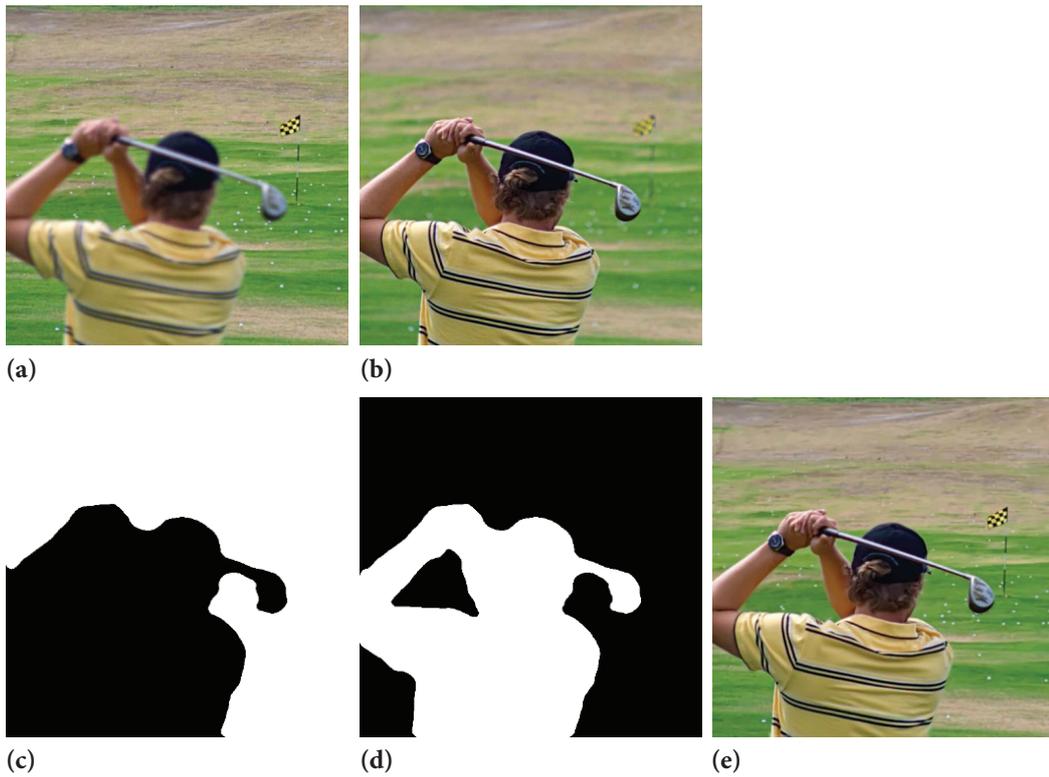
The methods for comparison in Table 3 aim to combine patches in source images by classifying them as focused and non-focused. These approaches generally use deep learning methods as classifiers. Although deep learning approaches produce typically successful results, they include difficulties such as determining parameters, creating a suitable network, and creating a training dataset. The proposed approach in this study makes classification without the need for these parameters. It also includes a two-step fusion rule that can minimize classification errors. Thanks to these features, the proposed approach produces successful results for multi-focus image fusion. Objective metric results for all examples from the Lytro [37] dataset are given in Table 3. When the metric results are examined, it is seen that the proposed approach is superior to the methods in the literature. This superiority arises because the proposed study is designed

following multi-focus image fusion purposes. In addition, visual results of the proposed approach for some images from the Lytro [37] dataset are given in this section. When the visuals are examined, it is seen that the visual results of the proposed approach are as good as the numerical results for fusion. This section gives Lytro-08, Lytro-01, and Lytro-10 from the related dataset.

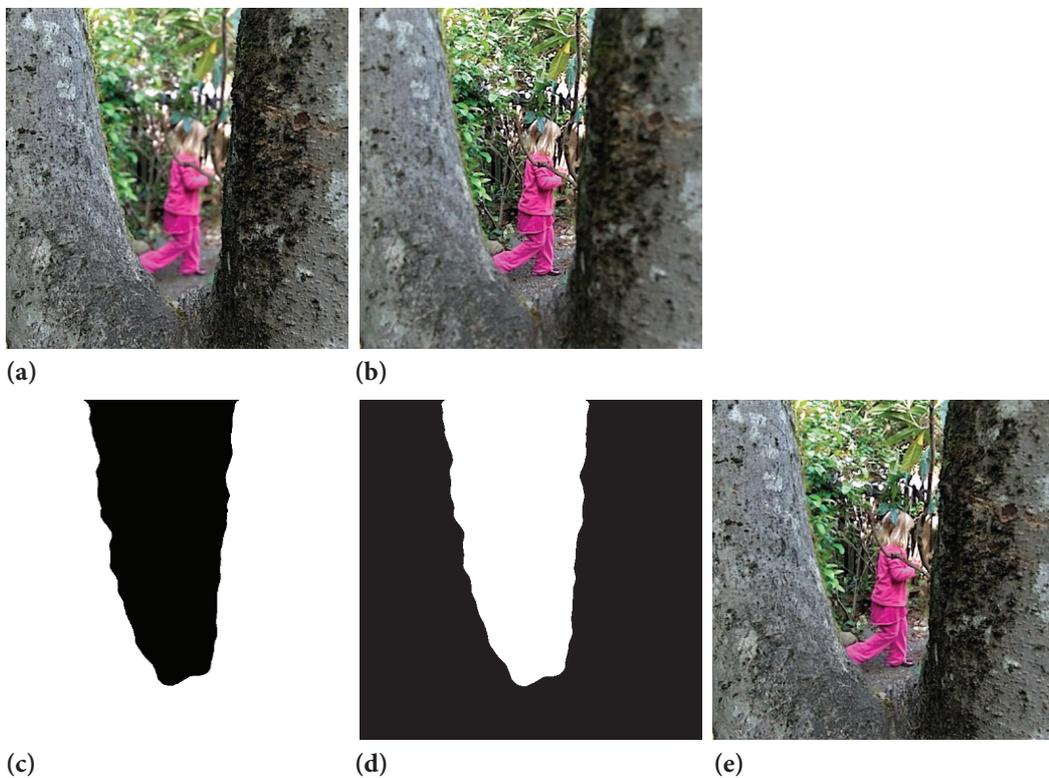
Figures 7-9 provide the visual results of the proposed study. All of these figures are organized in the same way. Images a and b in each figure represent source images with different focus values from the Lytro [37] dataset. Images c and d show the decision maps formed after the proposed classification mechanism and morphological operators. As can be seen from the decision maps, the proposed approaches successfully separated the focused and non-focused regions. Finally, after the proposed study, the f image gives the created fused image. The fused image is rendered perfectly, very close to the original. The Lytro [37] dataset consists of colour images. The proposed approach successfully combines both grey and colour-level images.



**Figure 7.** Multi-focus Lytro-08 image. (a) Right-focused source image, (b) Left-focused source image, (c) Decision map for right-focused image after proposed classification method, (d) Decision map for left-focused image after proposed classification method and (e) Fused image



**Figure 8.** Multi-focus Lytro-01 image. (a) Right-focused source image, (b) Left-focused source image, (c) Decision map for right-focused image after proposed classification method, (d) Decision map for left-focused image after proposed classification method and (e) Fused image.



**Figure 9.** Multi-focus Lytro-10 image. (a) Right-focused source image, (b) Left-focused source image, (c) Decision map for right-focused image after proposed classification method, (d) Decision map for left-focused image after proposed classification method and (e) Fused image.

## CONCLUSION

Multi-focus image fusion methods create a full-focus fused image by combining two images with focusing problems. These methods aim to transfer the correct focus information from the source images to the merged image and to preserve essential details such as edges and corners in the source images. The proposed approach is a new approach prepared for this purpose. In this approach, a new classification mechanism is created based on the classification of image blocks as focused or non-focused. This mechanism works based on focus metrics and does not require any training process. In addition, another feature that makes this mechanism unique is that it can also classify ambiguous regions. The fuzzy areas that appear in the transitions from focused to non-focus areas are essential in the fused image. Therefore, it is considered in the proposed approach. Decision maps created with the classification mechanism are improved with morphological operations, and final decision maps are created.

One of the essential steps in multi-focus image fusion methods is determining the fusion rule. The fusion rule reveals the importance of pixels in the source images for the fused image. In the proposed work, a two-step flexible fusion rule is applied. This fusion rule can directly transfer the classifier's regions to the fused image. In contrast, the ambiguous areas are transferred to the fused image with the gradient-based fusion rule. The methods in the literature ignore classification errors. Therefore, the success of image fusion is limited, while the classifier's success is limited. Thanks to the two-step fusion rule used in the proposed study, classification errors are minimized, and more detailed combined images are obtained compared to the methods in the literature. The proposed method performance is calculated using objective and subjective metrics, and the results show that the proposed study is quite successful in multi-focus image fusion.

## AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

## DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

## CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## ETHICS

There are no ethical issues with the publication of this manuscript.

## REFERENCES

- [1] Li S, Kwok JT, Wang Y. Combination of images with diverse focuses using the spatial frequency. *Information Fusion* 2001;2(3):169–176. [\[CrossRef\]](#)
- [2] Huang W, Jing Z. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters* 2007;28(4):493–500. [\[CrossRef\]](#)
- [3] Agrawal D, Singhai J. Multifocus image fusion using modified pulse coupled neural network for improved image quality. *IET Image Processing* 2010;4(6):443–451. [\[CrossRef\]](#)
- [4] Jin X, Zhou D, Yao S, Nie R, Jiang Q, He K, Wang Q. Multi-focus image fusion method using S-PCNN optimized by particle swarm optimization. *Soft Computing* 2017;22(19):6395–6407. [\[CrossRef\]](#)
- [5] Hua K, Wang H, Rusdi AH, Jiang S. A novel multi-focus image fusion algorithm based on random walks. *Journal of Visual Communication and Image Representation* 2014;25(5):951–962. [\[CrossRef\]](#)
- [6] Ma J, Zhou Z, Wang B, Miao L, Zong H. Multi-focus image fusion using boosted random walks-based algorithm with two-scale focus maps. *Neurocomputing* 2019;335:9–20. [\[CrossRef\]](#)
- [7] Yan X, Qin H, Li J, Zhou H, Yang T. Multi-focus image fusion using a guided-filter-based difference image. *Applied Optics* 2016;55(9):2230–2239. [\[CrossRef\]](#)
- [8] Qiu X, Li M, Zhang L, Yuan X. Guided filter-based multi-focus image fusion through focus region detection. *Signal Processing: Image Communication* 2019;72:35–46. [\[CrossRef\]](#)
- [9] Xiao J, Liu T, Zhang Y, Zou B, Lei J, Li Q. Multi-focus image fusion based on depth extraction with inhomogeneous diffusion equation. *Signal Processing* 2016;125:171–186. [\[CrossRef\]](#)
- [10] Bouzos O, Andreadis I, Mitianoudis N. Conditional random Field model for robust multi-focus image fusion. *IEEE Transactions on Image Processing* 2019;28(11):5636–5648. [\[CrossRef\]](#)
- [11] Li Q, Yang X, Wu W, Liu K, Jeon G. Multi-focus image fusion method for vision sensor systems via dictionary learning with guided filter. *Sensors* 2018;18(7):2143. [\[CrossRef\]](#)
- [12] Nejati M, Samavi S, Shirani S. Multi-focus image fusion using dictionary-based sparse representation. *Information Fusion* 2015;25:72–84. [\[CrossRef\]](#)
- [13] Yang Y, Yang M, Huang S, Ding M, Sun J. Robust sparse representation combined with adaptive PCNN for Multifocus image fusion. *IEEE Access* 2018;6:20138–20151. [\[CrossRef\]](#)
- [14] Burt PJ, Adelson EH. Merging images through pattern decomposition. *SPIE Proceedings* 1985;(0575):173–181. [\[CrossRef\]](#)
- [15] Jin X, Hou J, Nie R, Yao S, Zhou D, Jiang Q, He K. A lightweight scheme for multi-focus image fusion. *Multimedia Tools and Applications* 2018;77(18):23501–23527. [\[CrossRef\]](#)

- [16] Kou L, Zhang L, Zhang K, Sun J, Han Q, Jin Z. A multi-focus image fusion method via region mosaicking on Laplacian pyramids. *PLOS ONE* 2018;13(5):e0191085. [CrossRef]
- [17] Tian J, Chen L. Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure. *Signal Processing* 2012;92(9):2137–2146. [CrossRef]
- [18] Aymaz S, Köse C. A novel image decomposition-based hybrid technique with super-resolution method for multi-focus image fusion. *Information Fusion* 2019;45:113–127. [CrossRef]
- [19] Chai Y, Li H, Guo M. Multifocus image fusion scheme based on features of multiscale products and PCNN in lifting stationary wavelet domain. *Optics Communications* 2011;284(5):1146–1158. [CrossRef]
- [20] Li S, Kwok J, Tsang I, Wang Y. Fusing images with different focuses using support vector machines. *IEEE Transactions on Neural Networks* 2004;15(6):1555–1561. [CrossRef]
- [21] Yu B, Jia B, Ding L, Cai Z, Wu Q, Law R, Huang J, Song L, Fu S. Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion. *Neurocomputing* 2016;182:1–9. [CrossRef]
- [22] Saeedi J, Faez K. A classification and fuzzy-based approach for digital multi-focus image fusion. *Pattern Analysis and Applications* 2011;16(3):365–379. [CrossRef]
- [23] Jiang Q, Jin X, Lee S, Yao S. A novel multi-focus image fusion method based on stationary wavelet transform and local features of fuzzy sets. *IEEE Access* 2017;5:20286–20302. [CrossRef]
- [24] Gai D, Shen X, Chen H, Su P. Multi-focus image fusion method based on two stage of convolutional neural network. *Signal Processing* 2020;176:107681. [CrossRef]
- [25] Liu Y, Chen X, Peng H, Wang Z. Multi-focus image fusion with a deep convolutional neural network. *Information Fusion* 2017;36:191–207. [CrossRef]
- [26] Ma J, Yu W, Liang P, Li C, Jiang J. FusionGAN: A generative adversarial network for infrared and visible image fusion. *Information Fusion* 2019;48:11–26. [CrossRef]
- [27] Zhang Y, Liu Y, Sun P, Yan H, Zhao X, Zhang L. IFCNN: A general image fusion framework based on convolutional neural network. *Information Fusion* 2020;54:99–118. [CrossRef]
- [28] Amin-Naji M, Aghagolzadeh A, Ezoji M. Ensemble of CNN for multi-focus image fusion. *Information Fusion* 2019;51:201–214. [CrossRef]
- [29] Bhalla K, Koundal D, Sharma B, Hu Y-C, Zaguia A. A fuzzy convolutional neural network for enhancing multi-focus image fusion. *Journal of Visual Communication and Image Representation* 2022; 84, 103485. [CrossRef]
- [30] Guan Z, Wang X, Nie R, Yu S, Wang C. NCDCN: Multi-focus Image Fusion via nest connection and dilated convolution network. *Applied Intelligence* 2022. [CrossRef]
- [31] Ma J, Le Z, Tian X, Jiang J. SMFuse: Multi-Focus Image Fusion via self-supervised mask-optimization. *IEEE Transactions on Computational Imaging* 2021;7:309–320. [CrossRef]
- [32] Huang W, Jing Z. Evaluation of focus measures in multi-focus image fusion. *Pattern Recognition Letters* 2007;28(4):493–500. [CrossRef]
- [33] Nayar SK, Nakagawa Y. Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1994;8(16): 824–831. [CrossRef]
- [34] Li S, Kwok JT, Wang Y. Combination of images with diverse focuses using the spatial frequency. *Information Fusion* 2001;2(3):169–176. [CrossRef]
- [35] Zhang X, Li X, Feng Y. A New Multi-Focus Image Fusion Based on Spectrum Comparison. *Signal Processing*, 123(2016) 127–142. [CrossRef]
- [36] Jagalingam P, Hegde AV. A review of quality metrics for fused image. *Aquatic Procedia* 2015;4:133–142. [CrossRef]
- [37] <http://mansournejati.ece.iut.ac.ir/content/lytro-multi-focus-dataset>, 9 March 2015.
- [38] <http://bdps1989.wixsite.com/durgaprasad/datasets>, 3 January 2017.
- [39] Chen C, Gend P, Lu K. Multi-focus Image Fusion Based on Multiwavelet and DFB. *Chemical Engineering Transactions* 2015;(46):277–282. [CrossRef]
- [40] Hua K, Wang H, Rusdi AH, Jiang S. A novel multi-focus image fusion algorithm based on random walks. *Journal of Visual Communication and Image Representation* 2014;25(5):951–962. [CrossRef]
- [41] Zhang B, Lu X, Pei H, Liu H, Zhao Y, Zhou W. Multi-focus image fusion algorithm based on focused region extraction. *Neurocomputing* 2016;174:733–748. [CrossRef]
- [42] Li L, Ma H, Jia Z, Si Y. A novel multiscale transform decomposition based multi-focus image fusion framework. *Multimedia Tools and Applications* 2021;80(8):12389–12409. [CrossRef]
- [43] Nejati M, Samavi S, Karimi N, Reza Soroushmehr S, Shirani S, Roosta I, Najarian K. Surface area-based focus criterion for multi-focus image fusion. *Information Fusion* 2017;36:284–295. [CrossRef]
- [44] Abdipour M, Nooshyar M. Multi-focus image fusion using sharpness criteria for visual sensor networks in wavelet domain. *Computers & Electrical Engineering* 2016;51:74–88. [CrossRef]
- [45] Chaudhary V, Kumar V. Block-based image fusion using multi-scale analysis to enhance depth of field and dynamic range. *Signal, Image and Video Processing* 2017;12(2):271–279. [CrossRef]

- 
- [46] Yang Y, Tong S, Huang S, Lin P, Fang Y. A Hybrid Method for Multi-Focus Image Fusion Based on Fast Discrete Curvelet Transform. *IEEE Access* 2017;5:14898–14913. [\[CrossRef\]](#)
- [47] Moushmi S, Sowmya V, Soman KP. Empirical Wavelet Transform for Multi-focus Image Fusion. *Proceedings of the International Conference on Soft Computing Systems, Advances in Intelligent Systems and Computing*, 2016, p. 257–263. [\[CrossRef\]](#)
- [48] Li H, Chai Y, Yin H, Liu G. Multifocus image fusion and denoising scheme based on homogeneity similarity. *Optics Communications* 2012;285(2):91–10. [\[CrossRef\]](#)
- [49] Aymaz S, Köse C, Aymaz Ş. Multi-focus image fusion for different datasets with super-resolution using gradient-based new fusion rule. *Multimedia Tools and Applications* 2020;79(19-20):13311–13350. [\[CrossRef\]](#)
- [50] Liu Y, Wang Z. Simultaneous image fusion and denoising with adaptive sparse representation. *IET Image Processing* 2015;9(5):347–357. [\[CrossRef\]](#)
- [51] Liu Y, Liu S, Wang Z. A general framework for image fusion based on multi-scale transform and sparse representation. *Information Fusion* 2015;24:147-164. [\[CrossRef\]](#)
- [52] Bai X, Zhang Y, Zhou F, Xue B. Quadtree-based multi-focus image fusion using a weighted focus-measure. *Information Fusion* 2015;22:105–118. [\[CrossRef\]](#)
- [53] Zhang Y, Bai X, Wang T. Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure. *Information Fusion* 2017;35:81–101. [\[CrossRef\]](#)