

Article

# WIG-Net: Wavelet-Based Defocus Deblurring with IFA and GCN

Yi Li , Nan Wang, Jinlong Li and Yu Zhang \*

School of Physical Science and Technology, Southwest Jiaotong University, Chengdu 611756, China; y\_li@my.swjtu.edu.cn (Y.L.); wangnan\_phy@swjtu.edu.cn (N.W.); jinlong\_lee@swjtu.edu.cn (J.L.)

\* Correspondence: zhangyu2@swjtu.edu.cn

**Abstract:** Although the existing deblurring methods for defocused images are capable of approximately recovering clear images, they still exhibit certain limitations, such as ringing artifacts and remaining blur. Along these lines, in this work, a novel deep-learning-based method for image defocus deblurring was proposed, which can be applied to medical images, traffic monitoring, and other fields. The developed approach is equipped with wavelet transform, an iterative filter adaptive module, and graph neural network and was specifically designed for handling defocused blur. Our network exhibits excellent properties in preserving the original information during the restoration of clear images, thereby enhancing its capability to spatially address varying blurriness and improving the quality of deblurring. From the acquired experimental results, the superiority of the introduced method in the context of image defocus deblurring compared to the majority of the existing algorithms was clearly demonstrated.

**Keywords:** image deblurring; deep learning; wavelet transform; defocus deblurring



**Citation:** Li, Y.; Wang, N.; Li, J.; Zhang, Y. WIG-Net: Wavelet-Based Defocus Deblurring with IFA and GCN. *Appl. Sci.* **2023**, *13*, 12513. <https://doi.org/10.3390/app132212513>

Academic Editor: Yinan Kong

Received: 2 October 2023

Revised: 10 November 2023

Accepted: 16 November 2023

Published: 20 November 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

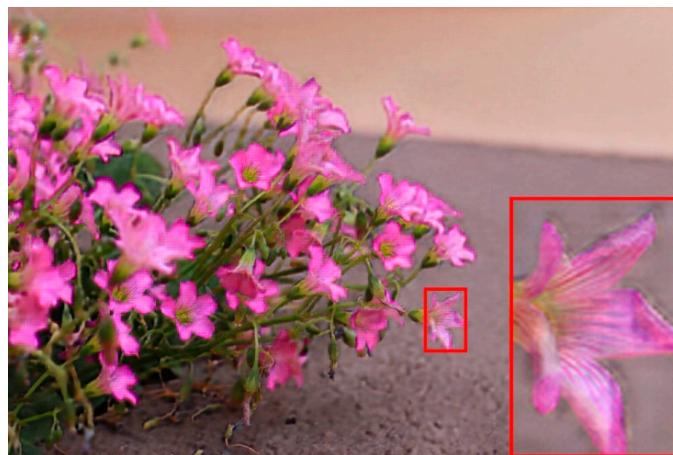
## 1. Introduction

When an object resides on the focal plane of a lens, the light emitted from a point on the object can be projected onto the image plane as a single point. However, for points that are not positioned on the focal plane, a blurred circle composed of numerous image points is formed, resulting in defocus blur. In real-world applications, errors in the camera's focal length or depth of field settings often lead to image defocus blur, significantly impairing the practical utility of the image. Image deblurring, as a way to improve image quality, is widely used in various fields, such as medical imaging and traffic monitoring. Image defocus deblurring is considered a typical ill-posed problem in image processing, with the aim to recover a corresponding clear image from a defocused blurred image. In recent years, this technology has received widespread attention from the scientific community. Nonetheless, the recovery of a clear image from a real blurred image remains challenging due to the complex and variable nature of the blur kernel in real-world scenarios.

In traditional defocus deblurring methods, defocused blurred images are treated as the outcomes of convolving clear images with various blur kernels. Therefore, these methods can restore clear images by predicting these blur kernels and subsequently applying non-blind deconvolution. However, due to the oversight of real-world blur nonlinearity during the process of defocus deblurring, the performance of the traditional approaches in defocus deblurring is suboptimal.

Recently, Abuolaim and Brown [1] introduced the first end-to-end learning-based approach, namely DPNet, which is independent of the specific blur models and directly recovers clear images. By employing an end-to-end learning methodology, DPNet exhibits superior performance in handling real-world defocused images compared to prior methods. The authors introduced the Dual-Pixel Defocus deblurring (DPDD) dataset. However, their deblurred image frequently exhibited ringing artifacts and remaining blur (Figure 1), primarily due to an excessive loss of original information during the image processing

procedure using straightforward UNet [2] architecture. On the other hand, the network's structure is overly simplistic, hindering both the accurate extraction of image information and the reconstruction of clear images.



**Figure 1.** Ringing artifacts and remaining blur in the defocus of the deblurring image.

In this work, an end-to-end network for performing single-image defocus deblurring was proposed. Innovatively, in the encoding phase of the network, the downsampling layers were replaced with wavelet transforms. It was also considered that the lossless image decomposition of images can be achieved using the wavelet transform. Therefore, the size of the feature maps was reduced while ensuring the preservation of image information. Simultaneously, in the decoding phase, wavelet inverse transforms were employed to upsample low-resolution feature maps into high-resolution images, effectively minimizing the image information loss during the process of restoring clear images from feature maps. Additionally, the images exhibiting sparsity in the wavelet domain enabled the network to learn mapping from sparse features to sparse features [3], which significantly enhanced the learning efficiency of the network.

Furthermore, to effectively address extensive defocus blur, the Iterative Filter Adaptive module (IFA) [4] and Graph Convolution Network modules (GCN) [5] were utilized. Particularly, IFA, a module proficient in handling spatial variations and significant defocus blur, was used. To address spatial variations, IFA utilizes an adaptive filter prediction scheme. More specifically, IFA does not directly predict pixel values; instead, it adaptively generates per-pixel defocus-deblurring filters and applies them to feature maps. Additionally, the GCN module was employed to exploit the characteristics of graph structures, connecting feature maps and employing graph convolution to recover lost details from the encoded feature maps of different channels.

In summary, the contributions of this work are as follows:

- (1) An end-to-end network for single-image de-defocusing is presented. Wavelet transforms are also incorporated into the encoding stage of the proposed network, reducing the feature map size while ensuring a wide receptive field.
- (2) IFA and GCN modules are introduced to increase the network's depth, thereby enhancing the ability to reconstruct clear images.
- (3) A proprietary dataset is curated. In contrast to previous datasets, the proposed collection included a higher proportion of images with extensive defocus blur, alongside their corresponding all-in-focus images.

## 2. Related Works

The process of image deblurring can be regarded as the pursuit of an optimal solution in the solution space. Traditional methods employ various natural image priors to constrain the solution space by estimating the maximum a posteriori mode [6]. However,

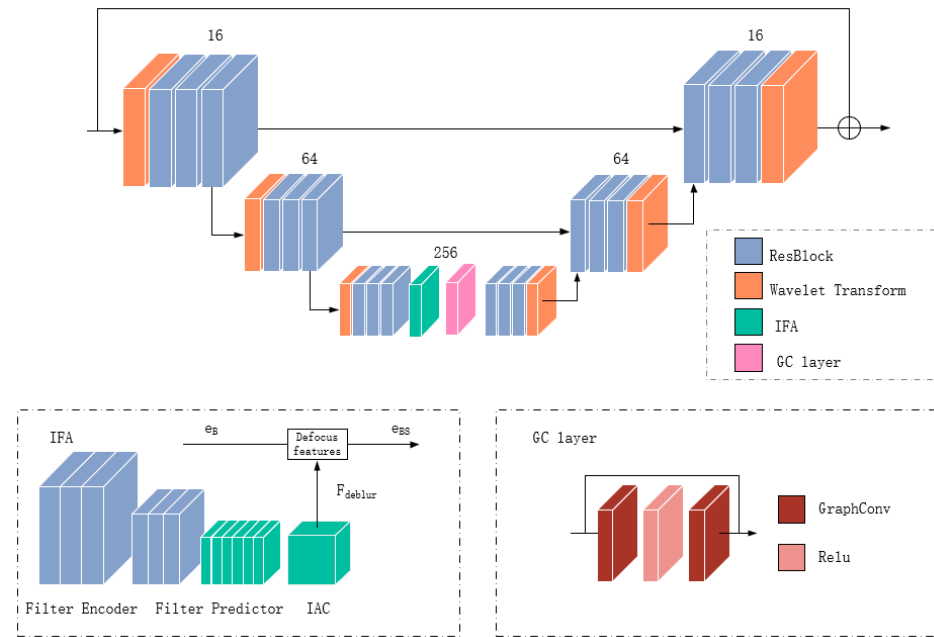
the conventional optimization approaches involve intricate iterative computations and lack real-time capabilities. Moreover, simplistic model assumptions can lead to inaccurate blur kernel estimation, thereby reducing the algorithm's accuracy. In recent years, with the advancements of deep learning, Convolutional Neural Networks (CNNs) have been extensively applied in the field of image deblurring [7–9].

Recent research has primarily focused on enhancing network architectures, introducing multi-scale and increasing receptive fields to enhance the performance of image restoration algorithms. Ronneberger et al. [2] proposed an encoder–decoder structure network (U-net), which effectively exploits contextual information to achieve superior performance in image semantic segmentation. Nah et al. [9] applied multiple scales to defuzzing networks, gradually removing differences in the degree of ambiguity. Chen et al. [10] embedded smooth dilated convolution to the network, while keeping the number of parameters in receptive fields constant to improve network performance. Zuozheng Lian et al. [11] introduced an enhanced U-Net, incorporating depth-wise separable convolutions, residual depth-wise separable convolutions, and wavelet transform. This approach enables the extraction of finer image details while simultaneously reducing computational complexity. Qian Ye et al. [12] introduced a network designed to establish direct mapping from a blurred input image to a clear image, leveraging the estimated defocus map to condition this mapping process. Li et al. [13] introduced the blind text image deblurring method to obtain a clean text image from the given blurry text image without knowing the blur kernel. However, their algorithm still lags behind mainstream deblurring algorithms in terms of processing time. Additionally, it does not achieve satisfactory results in the task of image deblurring caused by defocusing. Joan Bruna et al. [14] cascaded wavelet transform convolution and nonlinear modulus calculations to compute translation-invariant image features, preserving high-frequency information for classification. In another interesting work, Bae et al. [15] discovered that wavelet transform within CNNs is beneficial for single-image super-resolution and introduced wavelet residual networks. Yanyun Wu et al. [16] presented a deblurring method that employs a two-level wavelet-based convolutional neural network (CNN). This network incorporates discrete wavelet transform (DWT) to distinguish image context from texture information, consequently reducing computational complexity. Additionally, Junyong Lee et al. [4] introduced an Iterative Adaptive Network to address spatially varying blur. Their approach enhances the network's deblurring capabilities by predicting a unique deblurring filter for every pixel within the image.

Nevertheless, current defocus deblurring methods still exhibit certain drawbacks, such as image ringing artifacts and remaining blur. Furthermore, most networks are constrained when dealing with pronounced severe defocus blur, and are characterized by a large number of model parameters, long training times, and significant practical limitations.

### 3. Methods

The network architecture presented in this work is illustrated in Figure 2. Our approach was built upon U-net [2] by integrating wavelet transform into the encoding stage, replacing the conventional pooling layer. The sparse nature of wavelet coefficients simplifies the deblurring process while providing a larger receptive field to the network by decreasing the size of the feature maps. In the decoding stage, wavelet inverse transform, instead of the upsampling layer, was used to generate high-resolution feature maps from low-resolution ones, effectively mitigating the information loss due to pooling layers. The Iterative Filter Adaptive module (IFA) [4] and Graph Convolutional Network between the encoding and decoding stages were also introduced. The network employs a filter size of 3\*3 and uses Leaky Rectified Linear Unit (Leaky ReLU) [17] as the activation function. The Mean Squared Error (MSE) loss function was also adopted, which is widely used in image deblurring tasks and is the most suitable for our approach.



**Figure 2.** Schematic illustration of the proposed network architecture, which was built upon the U-Net framework, integrating wavelet transforms into the encoding stage and replacing conventional pooling layers. In the decoding phase, wavelet inverse transforms in lieu of upsampling layers were utilized. Furthermore, IFA modules and GCN layers between the encoding and decoding stages were introduced.

### 3.1. Wavelet Transform

In two-dimensional Discrete Wavelet Transform (2D DWT), the input signal is divided into four filter bands: LL, LH, HL, and HH. Initially, a one-dimensional (1D) DWT is applied to each row of the image, resulting in a low-frequency component (L) and a high-frequency component (H) in the horizontal direction. Subsequently, another 1D DWT is performed on each column of the transformed data, generating four subbands:

LL Subband: This subband contains low frequencies in both horizontal and vertical directions.

LH Subband: This subband represents low frequency in the horizontal direction and high frequency in the vertical direction.

HL Subband: This subband denotes high frequency in the horizontal direction and low frequency in the vertical direction.

HH Subband: This subband represents high frequency in both horizontal and vertical directions.

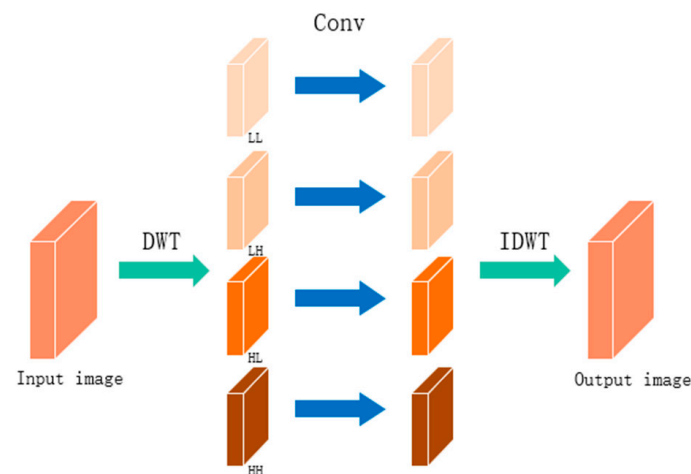
We employed 1D filters  $\phi(x)$  and  $\varphi(x)$  for filtering and horizontal downsampling of each column of the image. Subsequently, we employed two filters for filtering and vertical downsampling of each row. As a result, four sub-images,  $I_{LL}$ ,  $I_{LH}$ ,  $I_{HL}$ , and  $I_{HH}$ , could be computed. The 2D DWT can be represented as follows:

$$\begin{aligned}
 \psi_{LL}(x, y) &= \phi(x)\phi(y) \\
 \psi_{LH}(x, y) &= \phi(x)\varphi(y) \\
 \psi_{HL}(x, y) &= \varphi(x)\phi(y) \\
 \psi_{HH}(x, y) &= \varphi(x)\varphi(y)
 \end{aligned}
 \tag{1}$$

Conversely, during the reconstruction phase of 2D DWT, the inverse one-dimensional discrete wavelet transform was applied first to each column of the transformed result, followed by a similar operation on each row of the transformed data to obtain the reconstructed image. To summarize, the wavelet decomposition process of an image involves

separating the signal into low and high frequencies, and if necessary, the LL subband can be further decomposed until the desired level of detail is achieved.

In this work, the pooling and upsampling layers in the U-net were substituted with discrete wavelet transform (DWT) and inverse discrete wavelet transform (IDWT), respectively. During the encoding phase, a series of convolutional operations and DWT were employed to extract informative features from the input image. Subsequent to DWT processing, the data size was reduced to one-quarter of its original size, whereas the number of channels was quadrupled (Figure 3). Unlike conventional pooling layers that may lead to the loss of some original information due to methods like merging or adjusting convolutional strides to reduce data size, our approach can mitigate the data loss caused by this step.

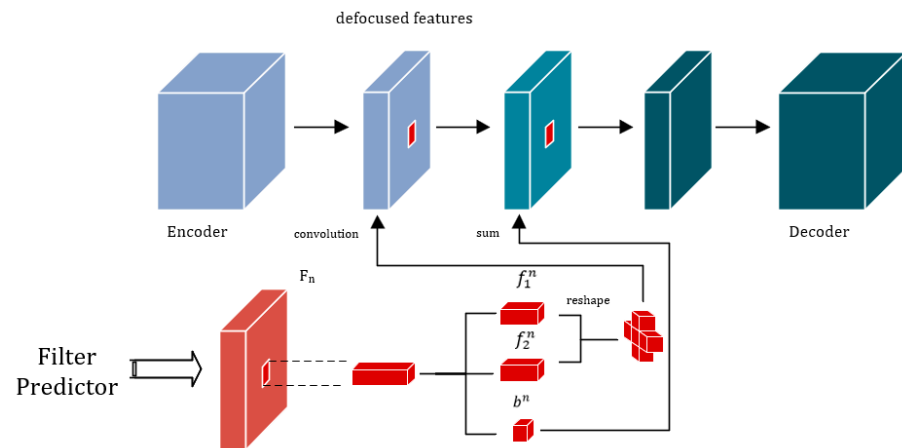


**Figure 3.** Application of wavelet transform in the encoding and decoding process.

### 3.2. Iterative Filter Adaptive Module

The IFA module [4] was added at the bottom of the U-net. IFA takes the defocused feature map extracted by the network as input and outputs the deblurred feature map. In the IFA module,  $I_b$ , which is the same as the main network input, was introduced to predict the deblurring filter. IFA consists of a filter encoder, a filter predictor, and an Iterative Adaptive Convolution layer (IAC). The filter encoder encodes  $I_b$  as  $e^F \in \mathbb{R}^{h \times w \times 256}$ , the filter predictor predicts the deblurring filter map  $F_{\text{deblur}} \in \mathbb{R}^{h \times w \times c_{\text{deblur}}}$ , where  $F_{\text{deblur}}$  is  $c$  ( $2k + 1$ ),  $c$  is 256, and  $k$  is 3, and finally, the IAC layer uses the prediction filter  $F_{\text{deblur}}$  to transform the input feature  $e_B$  to generate the deblurring feature  $e_{BS}$ .

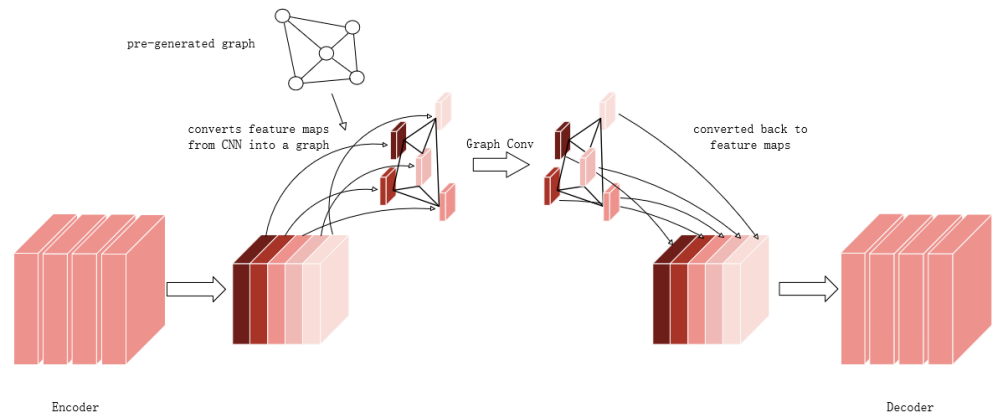
The Iterative Aggregation and Convolution (IAC) layer operates on the input  $F_{\text{deblur}}$ , where each spatial location in  $F_{\text{deblur}}$  is represented as an  $Nc(2k + 1)$ -dim vector, which corresponds to  $N$  sets of filters  $\{F_1, F_2, \dots, F_N\}$ . The  $n$ -th filter set  $F_n$  has two 1-dimensional filters,  $f_1^n$  and  $f_2^n$ , whose size is  $k \times 1$  and  $1 \times k$ , respectively, and a bias vector  $b^n$ . The IAC layer reshapes  $f_1^n$ ,  $f_2^n$ , and  $b^n$  into a filter whose size is  $k \times k$  and the channel is  $c$ . Corresponding to each spatial location, a filter is generated with channel  $c$ . The IAC layer decomposes the vectors in each position  $F$  into filter and bias vectors and iteratively applies them to  $e_B$  in a channel fashion to generate the output feature map. The process is shown in Figure 4.



**Figure 4.** Illustration of the Iterative Adaptive Convolution.

### 3.3. Graph Convolutional Network

The graph convolutional network was incorporated into the proposed network to leverage the distinctive features of graph structures. This permitted the recovery of details that were lost in defocus images. The specific process is illustrated in Figure 5. First, the encoder generates high-dimensional features, which are transformed into independent vertices. Then, these vertices are connected through a pre-generated graph. As a result, the feature maps are converted into a graph network that can be further processed by graph convolutions. Following multiple graph convolution operations, the nodes in the graph structure are restored to feature maps in the same order [18].



**Figure 5.** By reassembling feature maps from different channels into a graph structure and performing graph convolution operations, the lost image information was restored.

In the context of GCN, the propagation process involves the use of aggregators to acquire the hidden states of nodes. Different GCNs employ a variety of aggregators to collect data from neighbouring nodes, along with specific updates for adjusting node weights. Kipf et al. [5] introduced an aggregator designed for spectral GCNs. The aggregator is defined as follows:

$$T = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} X \tag{2}$$

where  $T$  is the aggregator,  $\tilde{A} = A + I_N$  is the adjacency matrix of the undirected graph with added self-connections,  $I_N$  is the identity matrix, and  $\tilde{D}$  is the degree matrix from [5]

$$\text{GraphConv}(X) = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} X \Theta \tag{3}$$

Therefore, we represent the graph convolution utilized in the proposed network using Equation (3), where  $\tilde{D}$  is the degree matrix,  $\tilde{A}$  is the adjacency matrix of the undirected graph,  $X$  is the convolved matrix, and  $\Theta$  is a matrix of graph convolution filter parameters.

Throughout the process of the graph convolution operation, the Watts–Strogatz (WS) model [19] was implemented to transform the feature map and graph structure. During the graph convolution process, the pre-generated graph was utilized based on the Watts–Strogatz model to achieve better results while reducing computational complexity. To further enhance the performance, the Residual Graph Convolutional Networks (ResGCN) [20] were integrated to increase the depth of the network.

### 3.4. Dataset

The DPDD dataset [1] was employed to train our network, which is a commonly used end-to-end defocus dataset. In addition, our own dataset (Figure 6) was created. To this end, a Canon EOS 800D camera was utilized to capture a pair of identical static images at various apertures (F1.8–F22). As the aperture size varied, the depth of field in the captured images differed, resulting in both clear (Ground Truth) and defocus blurred images. The specific capture procedure was as follows: First, the camera was securely positioned on a tripod to maintain stability throughout the entire shooting process while keeping the lens focal length constant. The aperture was adjusted to its minimum value to acquire fully clear images and subsequently switched to its maximum value to obtain blurred images. Throughout the entire capture process, the automatic exposure mode was employed to adjust the exposure time, ensuring consistent brightness across all images. To capture diverse scenes and different types of defocused and blurred images, 100 scenes (indoor and outdoor settings) and varying focal lengths were examined for image acquisition. Similarly, our custom-made dataset comprised a total of 500 pairs of images. These images were divided into training, validation, and test sets, with each set containing 400, 50, and 50 scenes, respectively.

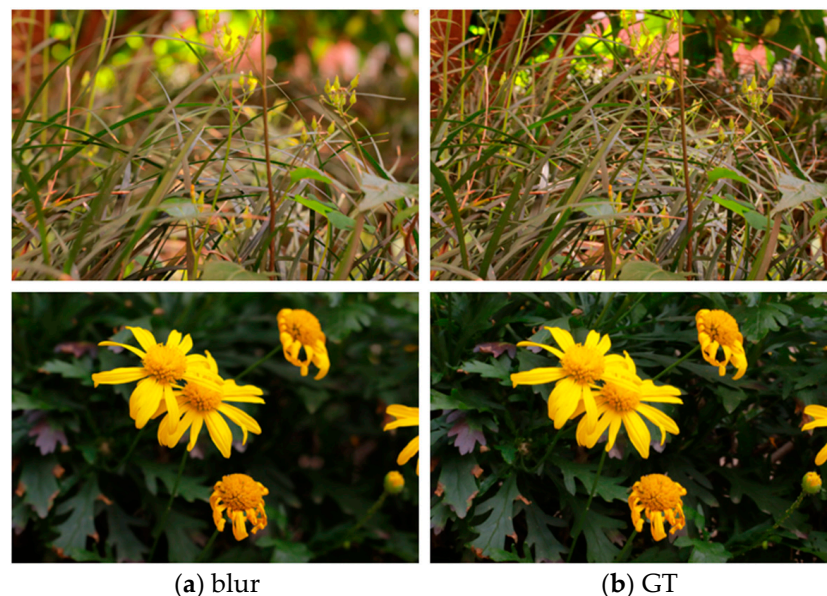


Figure 6. The custom-made defocus-blur dataset.

## 4. Experimental Section

Our model was implemented using PyTorch [21]. The training of the proposed model was performed using an Adam optimizer where  $\beta_1 = 0.9$  and  $\beta_2 = 0.99$ , with a weight decay rate of 0.01. We employed a total of 9 ResBlocks in the encoder and an additional 9 ResBlocks in the decoder. The learning rate was initialized to  $1.0 \times 10^{-4}$ . The number of filters were set to  $N = 17$  for  $F_{\text{deblur}}$ . Moreover, a batch size of 8 was used, and during

training, a  $256 \times 256$  region was randomly cropped from a blurred image and its ground truth image at the same location was used as the training input. MSE loss was also used, as it is the most suitable loss function and is widely used in image deblurring.

For evaluating the performance of defocus deblurring, the following metrics were used: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM) [22], and Mean Absolute Error (MAE). Higher PSNR values indicate better image quality. SSIM values range from  $-1$  to  $1$ , with  $1$  indicating perfect similarity. Lower MAE values signify better deblurring performance.

The developed models and previous ones were evaluated using a PC with an NVIDIA GeForce RTX 3060Ti GPU. The training process of our network using the DDPD dataset required a duration of 12 h. During testing, the average processing time for each image was 0.1 s.

#### 4.1. Comparison with Previous Methods

The proposed approach was compared with previous defocus map-based methods and recent end-to-end learning-based methods: Just Noticeable Blur estimation (JNB) [23], Edge-Based Defocus Blur estimation (EBDB) [24], DPDNet [1], Ye et al.'s method [12], and IFAN [4]. Among these methods, the JNB and EBDB are both defocus map-based approaches. They start by estimating the defocus map and subsequently carry out non-blind deconvolution. Ye et al.'s method attempts to learn direct mapping from the blurry input image to the clean image by utilizing the estimated defocus map to condition the mapping. DPDNet and IFAN belong to the end-to-end learning-based method that restores deblur images directly. A visual comparison is shown in Figure 7, indicating that WIG-Net restored clearer and sharper details.

These networks involved in the experiment were trained using the custom-made dataset. Table 1 presents the quantitative comparison. The results indicate that under the same training conditions, our network exhibits significantly improved defocus deblurring performance compared to the current mainstream methods, with notable enhancements in both PSNR and SSIM metrics.

**Table 1.** Quantitative comparison with previous defocus deblurring methods.

Model	PSNR	SSIM	MAE
Input	22.31	0.614	0.502
Shi et al. [23]	22.39	0.620	0.504
Karaali et al. [24]	22.45	0.632	0.487
Abuolaim et al. [1]	22.73	0.687	0.464
Ye et al. [12]	23.54	0.715	0.428
Lee et al. [4]	23.64	0.723	0.419
Ours	23.71	0.742	0.412

Our network outperforms the other competing networks in defocus deblurring due to several key reasons. Firstly, the introduced wavelet transform enables lossless image decomposition, significantly reducing the loss of valuable information during processing. Secondly, IFA and GCN extend the network's receptive field, enhancing its ability to handle large-scale defocusing blur effectively.





**Figure 7.** The comparative restoration results of the contrastive experimental algorithms on the custom-made dataset's test set [1,4,12,23,24].

#### 4.2. Ablation Study

To analyse the impact of each module in our model on the deblurring effect, an ablation study was performed (Table 2). All models in the ablation study were trained under the same conditions. To evaluate the effectiveness of each module, the baseline model and its four variants were compared. For the baseline model, conventional convolutional layers

and residual blocks were used instead of the key modules designed specifically for our network. In this work, comparison tests were performed on the DPDD dataset [1].

**Table 2.** Ablation study.

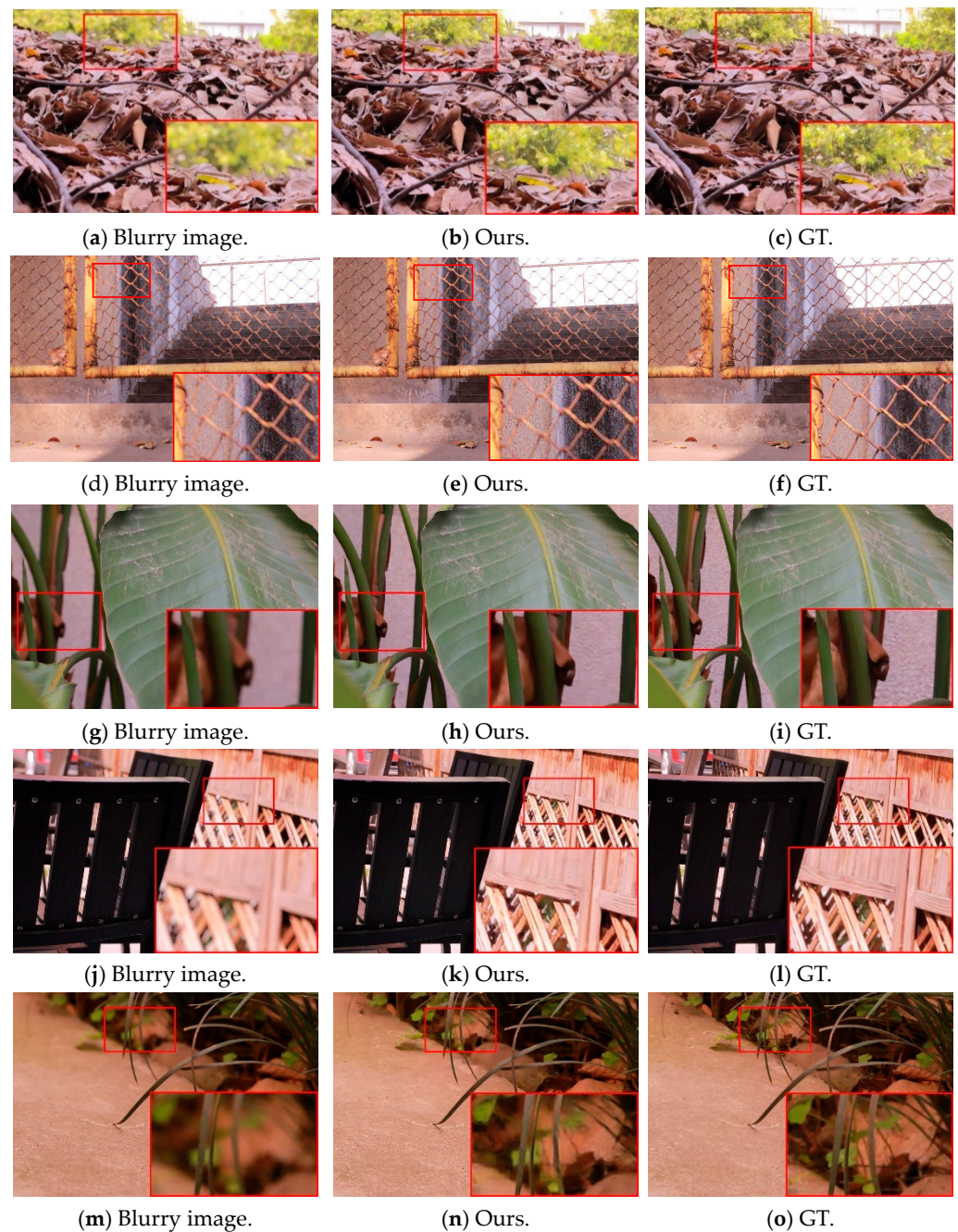
WtT	IFA	GCN	PSNR	SSIM	MAE
			23.67	0.705	0.436
✓			24.65	0.757	0.409
	✓		24.56	0.751	0.414
		✓	24.03	0.736	0.423
✓	✓		25.12	0.765	0.399
✓		✓	25.04	0.763	0.401
	✓	✓	24.77	0.749	0.411
✓	✓	✓	25.37	0.774	0.394

As shown in Table 2, the PSNR was 23.67 dB without introducing WtT, IFA, and GCN, which indicates that the baseline model alone has a limited effect on the defocus deblurring of images. With the introduction of WtT, IFA, and GCN, the sharpness of the deblurred images was improved to different degrees. The introduction of WtT, IFA, and GCN improved the average PSNR by 1.28 dB, 1.45 dB, and 1.37 dB, respectively, and the average SSIM by 0.06, enabling the network to obtain high-quality reconstructed images.

Compared to the baseline model (the first and second rows in Table 2), the introduction of WtT significantly enhanced the network's deblurring capabilities. This confirms that incorporating wavelet transform into the code-and-decode structure increases the perceptual field of the network, and the contextual information of the image can effectively be used to produce clearer results at the edges, and the reversibility of wavelet transform avoids the loss of image information. Meanwhile, the image is sparse in the wavelet domain, and the downsampling of the wavelet transform enhances the sparsity of the image features and improves the learning ability of the network. Furthermore, introducing IFA into the baseline model can also improve the deblurring performance, as evident from the first and third rows in Table 2. This validates the advantage of IFA in adaptively handling spatial variations. The GCN module [18] also demonstrates good performance in deblurring, as shown in the first and fourth rows of Table 2. This substantiates the effectiveness and feasibility of the method that restores lost information by associating features across channels through graph convolution. Nevertheless, the introduction of the GCN module increases the network's parameter count, leading to longer training times. Ensuring reasonable training times can impose limitations on the depth of our network.

#### 4.3. Generalization Ability

Given that our approach was trained using the DPDD training dataset [1], a natural question arises concerning the model's generalization to images from different datasets. To address this, the performance of our method on alternative test sets was assessed. Figure 8 depicts the outcomes of our network trained on the DPDD dataset when applied to our custom-made dataset's test set. The restored image details are prominently discernible, showcasing the efficacy of our model's training on the DPDD dataset. Table 3 illustrates the quantitative comparison of our custom-made dataset's test set. A noticeable enhancement in the image quality achieved by our model can be ascertained, indicating its favourable generalization capacity to images captured by other cameras. The underlying reasons for this effect could lie in the deep learning network utilizing wavelet transforms. Particularly, owing to its inherent low information loss property, the network's transferability is enhanced and the generalization ability of networks with significant depth is also improved.



**Figure 8.** Comparison of deblurring performance on the custom-made dataset's test set.

**Table 3.** Quantitative evaluation on the custom-made dataset's test set.

	PSNR	SSIM	MAE
Blurry image	21.05	0.632	0.513
Ours	23.46	0.708	0.435

## 5. Conclusions

In this work, a deep-learning-based image deblurring algorithm was proposed that employs a forward wavelet transform to replace downsampling and an inverse wavelet transform for upsampling. This method increased the network's receptive field and minimized information loss during transmission. To recover the lost image information in the wavelet domain, a graph neural network was employed that utilizes pre-defined graph

structures to improve image restoration performance, while keeping the computational cost low. An IFA module can handle spatially varying defocus blur by predicting a separable deblurring filter for each pixel. Our experimental results, using the DPDD test set and the custom-made dataset's test set, demonstrated that the proposed method could produce images with better visual quality and stronger robustness in various scenarios. Our network is capable of handling large areas of defocus blurring and can be applied to medical images, traffic monitoring, and other fields.

Our proposed network performed well in addressing defocus blur for smaller objects, as indicated in the labeled section of Figure 8e, but it still exhibited certain limitations. We plan to address this shortfall in future research by extending the dataset and enhancing the network's ability to extract detailed information.

**Author Contributions:** Conceptualization, Y.L.; Methodology, Y.L.; Software, Y.L.; Formal analysis, Y.L.; Investigation, Y.L.; Data curation, Y.L.; Writing—original draft, Y.L.; Writing—review & editing, N.W., J.L. and Y.Z.; Supervision, N.W. and Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Sichuan Province Science and Technology Support Program, grant number 2021ZYZFGY06; the National Natural Science Foundation of China, grant number 62104202; the Sichuan Province Science and Technology Support Program, grants numbers 2022YFG0217 and 2021ZYZFGY06; the Guang'an Science and Technology Innovation Project, grant number 2022CG04.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author. The data are not publicly available due to the data that support this study are proprietary.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Abuolaim, A.; Brown, M.S. Defocus deblurring using dual-pixel data. In *Computer Vision—ECCV 2020, Proceedings of the 16th European Conference, Glasgow, UK, 23–28 August 2020*; Springer: Cham, Switzerland, 2020; pp. 111–126.
2. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Proceedings of the 18th International Conference, Munich, Germany, 5–9 October 2015*; Springer: Cham, Switzerland, 2015; pp. 234–241.
3. Guo, T.; Seyed Mousavi, H.; Huu, V.T.; Monga, V. Deep wavelet prediction for image super-resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017*; pp. 104–113.
4. Lee, J.; Son, H.; Rim, J.; Cho, S.; Lee, S. Iterative filter adaptive network for single image defocus deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021*; pp. 2034–2042.
5. Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv* **2016**, arXiv:1609.02907.
6. Xu, L.; Zheng, S.; Jia, J. Unnatural l0 sparse representation for natural image deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013*; pp. 1107–1114.
7. Wang, R.; Zhang, C.; Zheng, X.; Lv, Y.; Zhao, Y. Joint Defocus deblurring and Superresolution Learning Network for Autonomous Driving. *IEEE Intell. Transp. Syst. Mag.* **2023**. [[CrossRef](#)]
8. Jiang, N.; Zhang, Y.; Yan, F.; Fu, X.; Kong, T. Image blind motion deblurring method with longitudinal channel and wavelet dynamic convolution. *Comput. Graph.* **2023**, *116*, 275–286. [[CrossRef](#)]
9. Nah, S.; Hyun Kim, T.; Mu Lee, K. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017*; pp. 3883–3891.
10. Chen, D.; He, M.; Fan, Q.; Liao, J.; Zhang, L.; Hou, D.; Yuan, L.; Hua, G. Gated context aggregation network for image dehazing and deraining. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 7–11 January 2019*; pp. 1375–1383.
11. Lian, Z.; Wang, H.; Zhang, Q. An Image Deblurring Method Using Improved U-Net Model. *Mob. Inf. Syst.* **2022**, *2022*, 6394788. [[CrossRef](#)]
12. Ye, Q.; Sukanuma, M.; Okatani, T. Accurate Single-Image Defocus deblurring Based on Improved Integration with Defocus Map Estimation. In *Proceedings of the IEEE International Conference on Image Processing (ICIP), Kuala Lumpur, Malaysia, 8–11 October 2023*.
13. Li, Z.; Yang, M.; Cheng, L.; Jia, X. Blind Text Image Deblurring Algorithm Based on Multi-Scale Fusion and Sparse Priors. *IEEE Access* **2023**, *11*, 16042–16055. [[CrossRef](#)]

14. Bruna, J.; Mallat, S. Invariant scattering convolution networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1872–1886. [[CrossRef](#)] [[PubMed](#)]
15. Bae, W.; Yoo, J.; Chul Ye, J. Beyond deep residual learning for image restoration: Persistent homology-guided manifold simplification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 145–153.
16. Wu, Y.; Qian, P.; Zhang, X. Two-level wavelet-based convolutional neural network for image deblurring. *IEEE Access* **2021**, *9*, 45853–45863. [[CrossRef](#)]
17. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. *Proc. ICML* **2013**, *30*, 3.
18. Xu, B.; Yin, H. Graph convolutional networks in feature space for image deblurring and super-resolution. In Proceedings of the International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 18–22 July 2021; pp. 1–8.
19. Watts, D.J.; Strogatz, S.H. Collective dynamics of ‘small-world’ networks. *Nature* **1998**, *393*, 440–442. [[CrossRef](#)] [[PubMed](#)]
20. Li, G.; Müller, M.; Thabet, A.; Ghanem, B. Deepgcns: Can gcns go as deep as cnns? In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 9267–9276.
21. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*; MIT Press: Cambridge, MA, USA, 2019.
22. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)] [[PubMed](#)]
23. Shi, J.; Xu, L.; Jia, J. Just noticeable defocus blur detection and estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 657–665.
24. Karaali, A.; Jung, C.R. Edge-based defocus blur estimation with adaptive scale selection. *IEEE Trans. Image Process.* **2017**, *27*, 1126–1137. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.