

Article

Iterative Dual CNNs for Image Deblurring

Jinbin Wang ^{1,*} , Ziqi Wang ² and Aiping Yang ^{2,*}

¹ The College of Intelligence and Computing, Tianjin University, Tianjin 300072, China

² School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

* Correspondence: wjb@tju.edu.cn (J.W.); yangaiping@tju.edu.cn (A.Y.)

Abstract: Image deblurring attracts research attention in the field of image processing and computer vision. Traditional deblurring methods based on statistical prior largely depend on the selected prior type, which limits their restoring ability. Moreover, the constructed deblurring model is difficult to solve, and the operation is comparatively complicated. Meanwhile, deep learning has become a hotspot in various fields in recent years. End-to-end convolutional neural networks (CNNs) can learn the pixel mapping relationships between degraded images and clear images. In addition, they can also obtain the result of effectively eliminating spatial variable blurring. However, conventional CNNs have some disadvantages in generalization ability and details of the restored image. Therefore, this paper presents an iterative dual CNN called IDC for image deblurring, where the task of image deblurring is divided into two sub-networks: deblurring and detail restoration. The deblurring sub-network adopts a U-Net structure to learn the semantical and structural features of the image, and the detail restoration sub-network utilizes a shallow and wide structure without downsampling, where only the image texture features are extracted. Finally, to obtain the deblurred image, this paper presents a multiscale iterative strategy that effectively improves the robustness and precision of the model. The experimental results showed that the proposed method has an excellent effect of deblurring on a real blurred image dataset and is suitable for various real application scenes.

Keywords: image deblurring; deep learning; dual CNNs; iterative strategy; U-net

MSC: 68T07; 68T45



Citation: Wang, J.; Wang, Z.; Yang, A. Iterative Dual CNNs for Image Deblurring. *Mathematics* **2022**, *10*, 3891. <https://doi.org/10.3390/math10203891>

Academic Editors: Xiangtao Zheng, Jinchang Ren and Ling Wang

Received: 20 September 2022

Accepted: 18 October 2022

Published: 20 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

The process of image blurring degradation can be described as a mathematical model where the input is a clear image, and the output is a blurred image. Meanwhile, the mathematical operation includes a blurring convolution kernel and noise addition. For a blurred image, common blurred types include defocus, motion, Gaussian, etc. Therefore, it can be seen that the image blurring degradation process is complicated, and the reasons for the blurred images can vary. The blurring degradation process is shown in Figure 1, where x is the original clear image, y means the blurred image, k stands for the blurring kernel, w indicates the additive noises, \otimes represents the convolution operation, and \oplus denotes the add operation.

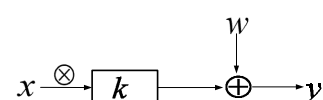


Figure 1. Image blurring degeneration model.

On the contrary, the image-deblurring problem is the inverse process, where the input is the blurred image, and the output is the clear image. Meanwhile, the blur kernel and the noise are unknown. Thus, the aim of deblurring is to simultaneously estimate the blur kernel and original image from the degraded image.

The traditional theory of image deblurring is based on statistical prior, which is relatively mature and leads researchers to design image priori knowledge artificially. The statistical priors proposed by these researchers are based on limited observation and statistics of image features. Thus, it cannot express the internal features of the image well, and the blurred image is not always consistent with the image priori condition selected by the algorithm. As a result, the traditional algorithm based on statistical prior has some drawbacks, including limited detail recovery ability and obvious artificial effect. Therefore, the effect of blur removal is poor in real-scene applications.

Due to the quick advancement of the graphics processing unit (GPU) parallel computing technology, deep learning has recently become a new research hotspot in the field of computer vision (such as super-resolution [1–6] and denoising [7–10], etc.). By learning blurred image features, end-to-end convolutional neural networks are designed to represent and model the image degradation process. When training convolutional neural networks, the clear image that corresponds to the blur degenerate image is regarded as the ground truth. Then, the intrinsic mapping relationship between the degraded image and the targeted clear image is continuously learned through the networks. The parameters of network learning represent the relationship between the blurred image and the clear image, and convolutional neural networks update the network parameters using an error back propagation method to implement the learning process. Meanwhile, the deep learning method is data-driven, where the generated image prior can describe the internal features of images more accurately than the manually designed image prior so that it can achieve superior recovery performance. However, the majority of existing deep learning-based methods do not take the recovery of image details into account, which causes the loss of detail texture after image deblurring. Moreover, it is challenging to accomplish multitype and multiscene image deblurring tasks due to its poor generalization and adaptability abilities.

In this paper, we propose an iterative dual CNNs for image deblurring as well as IDC that is composed of a deblurring sub-networks (Sub-Deblur), a detail recovery sub-networks (Sub-Texture), and an information fusion module.

The main contributions of this work can be summarized as follows:

- Specifically, the deblurring sub-networks based on U-Net structure are used to increase the receptive field through the multiscale inputs. That can acquire multiple-scale structural information on the blurred images and improve both the deblurring performance and the generalization ability of the networks.
- Then, the detail recovery sub-network with a shallow and wide network structure extract the texture information of the blurred images. It keeps the feature maps and the input image at the same resolution without downsampling.
- Finally, the fusion module is used to construct the clean image by fuse and enhance the output of the deblurring and texture recovery sub-networks.
- Additionally, extensive experiments illustrate that our IDC outperforms most of the state-of-the-art deblurring in terms of both quantitative and qualitative analysis.

The remainder of this paper is organized as follows. Section 2 offers the related work of deep learning methods on image deblurring, end-to-end U-net structure for image applications. Section 3 provides the proposed method. Section 4 shows the extensive experiments and results of the proposed method for image deblurring. Section 5 presents the conclusion.

2. Related Work

2.1. Priors-Based Methods

Image deblurring is an inverse problem, and researchers have proposed a large number of methods whose general idea is to constrain the solution space of the problem through the regularization method. The traditional deblurring methods are mostly based on variational models, and the quality of image recovery depends on the choice of image priors, such as the dark channel prior [11] that helps to recover the natural image, the low-rank prior [12] that helps to obtain significant edges, and the gradient prior [13] that can effectively remove

artifacts. In addition, Xu et al. [14] proposed an L0-based image regularizer for motion deblurring, which seeks gradient sparsity close to L0-regularized to remove pernicious small-amplitude structures. Sun et al. [15] proposed a deep learning approach to predicting the probabilistic distribution of motion blur at the patch level using a convolutional neural network (CNN). They further extend the candidate set of motion kernels predicted by the CNN using carefully designed image rotations. A Markov random field model is then used to infer a dense nonuniform motion blur field enforcing motion smoothness. Finally, motion blur is removed by a non-uniform deblurring model through patch-level image prior. The above methods rely heavily on a priori assumptions about image information, leading to poor model generalization. At the same time, most such methods have too-simple assumptions on blur kernels, but the actual blur types are often complex and nonlinear, so it is difficult to get practical applications.

2.2. Deep Learning Methods

Initially, Jain and Seung et al. [16] adopted a four-layer convolutional network structure for image blur degradation, but experiments showed that a more complicated neural network architecture was required for the image deblurring task. Hradi et al. [17] proposed an image deblurring algorithm for text images that was based on deep convolutional neural networks to realize the deblurring task. Su et al. [18] presented a video deblurring method based on deep convolutional neural networks, but this method required learning multiframe image features and was not suitable for a single image. In order to further improve the generalization ability and the performance of the model, some researchers tried to design new network structures and learning mechanisms. To realize the deblurring framework based on the sparse representation, Nimisha et al. [19] combined the convolutional natural network autocoding structure and the survival adversarial network, which effectively reduced network parameters, introduced gradient loss, and enhanced detail recovery capability. However, its training process was complicated and required building a good dictionary that relies on encoding-decoding the network's training results. Moreover, the method based on sparse representation still results in an overly smooth restored image and the loss of details. Subeesh Vasu et al. [20] considered that the information is fused and optimized under the condition of inaccurate blur kernel estimation and generated better deblurring images using different methods as the input and regarding the optimized fusion deblurring image as the output. Since the network structure of this technique is straightforward, training is simple and requires a smaller set of training data. It can handle a variety of blur kinds when used with conventional algorithms, but this process is too expensive and impractical for practical use, different deblurring intermediate images from various strategies are produced. Patrick Wieschollek et al. [21] designed a U-typed network structure, which introduced a residual module with skip connection at different layers. Although this approach can more effectively utilize valuable temporal information and share learning parameters between frames, it is only suitable for multi-frame images and has poor real-time performance. Orest Kupyn et al. [22] employed conditional adversarial networks for image deblurring. However, this method adopts the cross-entropy loss function, which reduces the network training efficiency. In order to categorize blurred images from clear images, Lerenhan et al. [23] trained a binary classification network. They then introduced the learned priori from the feature maps of the middle layer into the traditional statistical prior-based deblurring model. This method, which is simple to train and can handle a variety of blur kinds, combines deep learning techniques with conventional optimization techniques. However, when there is significant noise, it easily fails to deblur. The method is also time-consuming and unable to handle spatial variable blur directly. Jiawei Zhang et al. [24] built a spatially-varying recurrent neural network (RNN) for deblurring. The CNNs in this method follow the rules of RNN and are not directly involved in the deblurring process, thus reducing the number of parameters and the running time of the model. However, the four RNN modules use the same weights generated by CNNs, and only four directions are considered in each RNN, resulting in

the detail loss of the restored images. Nah et al. [25] described a multiscale convolutional neural network deblurring algorithm for dynamic scenes, which used an image pyramid structure to train the networks from coarse to fine and applied the multiscale iterative strategy to effectively learn image information of each scale level. This method efficiently improves the network performance in image detail recovery and avoids artifacts. However, the network has 120 layers in total, which presents the issue of large parameter quantities and excessive running time. In order to extract image features from coarse to fine resolution, Gao et al. [26] published the selective parameter sharing skip connection deblur network and applied nested skip connection and selective parameter sharing to reduce the difficulty of network training and improve the deblurring performance. Based on the generative adversarial network, [19,27] successively proposed DeblurGAN and DeblurGAN-v2 to recover more realistic clear images. Besides, it utilized the cross-entropy loss function to easily appear gradient saturation phenomenon but reduce the training efficiency. Suin et al. [28] used the global attention mechanism and the adaptive local filter for feature extraction to improve the performance of image deblurring at a fast speed. Tao et al. [29] proposed a scale-recurrent network to perform deblurring in different scales by shared used Long short-term memory (LSTM) cells to aggregate feature maps from coarse-to-fine scales. This parameter sharing scheme neglects scale variant properties of features, which are crucial for respective restoration in each scale.

2.3. End to End U-Net Structure

End-to-end CNNs are widely used in image restoration because it eliminates laborious manual feature extraction and preprocessing. In the research of image deblurring, CNNs learn the mapping function between clear images and blurred degenerated images. Its mathematical principle can be presented as:

$$\hat{x} = f(y, \theta) \quad (1)$$

This equation represents the mapping function $f(y, \theta)$ between the blurred image y and the deblurred image \hat{x} obtained by CNNs, where θ represents the network weight parameters. The input of the convolutional neural networks is the blurred image and the output is the estimation of the original clear image.

At present, many end-to-end convolutional neural networks based on the U-Net structure are applied in image restoration fields, such as image noising and image super-resolution reconstruction [30–32]. The method in [33] demonstrates that this network structure can increase the receptive field and acquire more image features at different levels under the same network depth, which has strong applicability and effectiveness for image restoration tasks. However, U-Net uses a pooling layer for downsampling from feature maps, which retains the structure information of the feature maps but loses its high frame details, leading to blurred image details. As one of the most important structures in traditional CNNs, the pooling layer mainly realizes the downsampling of feature maps and has been widely used in image detection, object recognition, image classification, image semantic segmentation, and so on. The main purpose of downsampling is to remove overly detailed textures in the feature maps or original images while maintaining image structure information so as to better realize the feature classification and recognition. It is very important for the task of detection and classification. In the aspect of image deblurring, the method in [34] further notes that structural information is helpful to blur kernel estimation in the deblurring algorithm. Too many fine textures are not beneficial for blur kernel estimation, and the application of pooling layers increases the generalization ability of networks and makes the networks more robust.

Figure 2 shows the structure of the general U-Net. In the coding stage, the input blurred images are downsampled twice in the pooling layer to obtain a series of feature maps with progressively lower sizes. In the decoding stage, these feature maps are sampled on two deconvolution layers, and finally, deblurred images are generated by information fusion. The network achieves excellent image deblurring performance, but

because it adopts the U-Net structure, the texture information of the feature maps and the original image is lost in the learning process, which leaves the deblurred image with insufficient restoration.

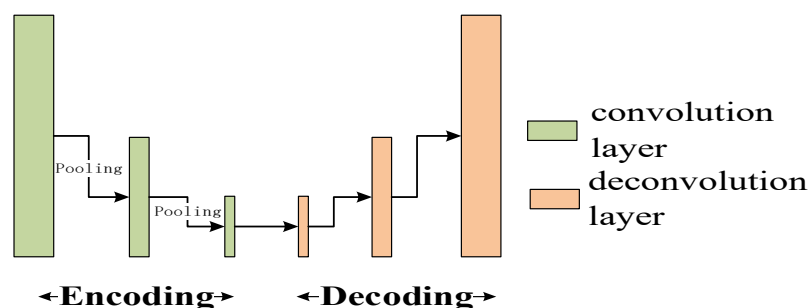


Figure 2. U-Net Structure.

3. The Proposed Method

In order to more effectively address the problem of the limited generalization ability and texture information loss after image deblurring, this paper proposes an iterative dual-CNN (IDC) for image deblurring. The network broadens its receptive field by receiving information from several scales, obtaining multiscale hierarchical information about the blurred image, and designing a detail recovery network that can completely extract and maintain the blurred image's detail features. At the same time, the network extracts and fuses the detail recovery sub-networks output and the deblurring sub-networks output to generate the final deblurred image. The produced deblurred images are rich in detail textures and have a wide range of applications in spatially variable blurring scenes.

3.1. Network Structure

As described in the above section, although the structure of U-Net has a strong generalization ability and can be better applied to the blur image degradation process, the downsampling layer in the structure also causes detail loss and is excessively flat in the final blurred image. Inspired by [35], who realize image deblurring and super-resolution reconstruction through two parallel convolutional neural networks, the proposed method divides the image deblur task into two subtasks, deblurring task and detail recovery subtask, to enhance the ability of detail recovery while considering the generalization performance of the networks. The deblurring sub-network based on the U-Net structure is adopted to learn semantics and structure information of the blurred image, while the detail recovery sub-network only contains a convolution feature extraction layer with a residual structure, without any downsampling layer. The size of feature maps remains the same as the input size, and only texture features are extracted.

The iterative dual-CNN image deblurring network designed in this paper consists of four different convolution modules and two sub-networks. The four convolution modules are:

- ResBlock module: The residual module consists of two convolution layers and activation functions, and the convolution layers are in skip connection.
- Eblock module: The encoding module is connected with two ResBlock modules through an activation function by a convolution layer responsible for blurred image feature extraction.
- Dblock module: The decoding module serially connects two ResBlock modules and one deconvolution layer via an activation function.
- Fusion module: The information integration module combines the deblurring sub-network and the detail recovery sub-network to output the final deblurred image.

The two sub-networks are:

- Sub-Deblur: The deblurring sub-network consists of Eblock modules E1, E2, E3 and Dblock modules D1, D2, D3. It is based on a U-Net structure to complete the basic deblurring task.
- Sub-Texture: The detail recovery sub-network consists of two Eblock modules connected in series for image detail recovery and retention.

The structure of the iterative dual CNNs image deblurring network designed in this paper is shown in Figure 3. Specific network parameters are shown in Table 1 below. Stride is the convolution step size, and Padding refers to the filling mode.

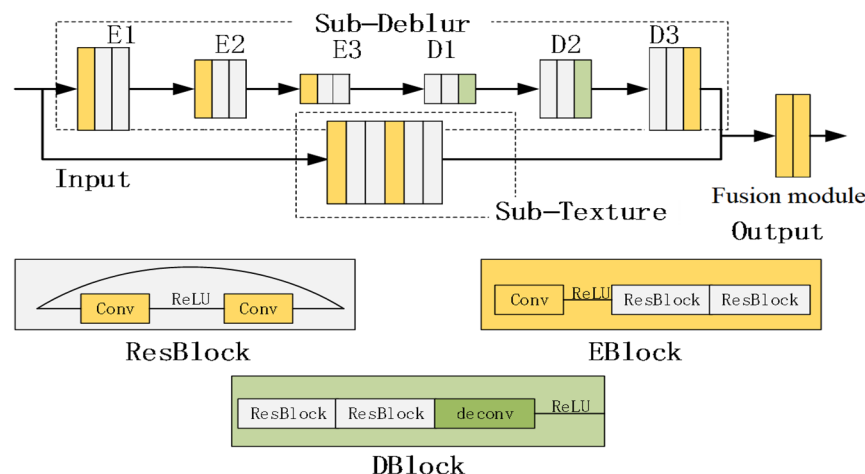


Figure 3. An overview of IDC structure.

Table 1. Network Parameters.

Block	Kernel Number	Kernel Size	Stride	Padding
E2	32	Conv: 5×5	1	SAME
	32	ResBlock: 5×5	1	SAME
E2	64	Conv: 5×5	2	SAME
	64	ResBlock: 5×5	1	SAME
E3	128	Conv: 5×5	2	SAME
	128	ResBlock: 5×5	1	SAME
D1	128	deconv: 4×4	2	SAME
	128	ResBlock: 5×5	1	SAME
D2	64	deconv: 4×4	2	SAME
	64	ResBlock: 5×5	1	SAME
D3	32	deconv: 4×4	1	SAME
	32	ResBlock: 5×5	1	SAME
Sub-Texture	64	Eblock: 5×5	1	SAME
Fusion module	32	Conv: 5×5	1	SAME
	3	Conv: 5×5	1	SAME

3.2. Multi-Scale Iterative Strategy

The process of image blur degradation is very complicated. Taking motion blur as an example, its motion path is random and varied, and the resulting blur effect is different. Reference [33] points out that the network generalization capability can be improved by increasing the diversity of training data and the perception field. In the aspect of image deblurring, only when the effective perception field of the network should theoretically be

larger than the blurring kernel size can the image deblurring task be completed. Inspired by the multilayer pyramid strategy adopted by the model-based deblurring algorithm, a multiscale iteration strategy is adopted to learn the multiscale information of the blurred images through the multiscale iterative input. This strategy can increase the receptive field of the network without changing the network structure or parameters, thus improving the generalization ability and robustness of the model.

In the image preprocessing stage, the multiscale iterative strategy is used by downsampling the original blurred image several times and regarding them as the input of the IDC. Meanwhile, the multiscale clear image obtained by multiple downsampling of the corresponding original clear image is used as the ground truth. As shown in Figure 4, three blurred images (y_1, y_2, y_3) with different sizes are obtained by two downsamplings of the original blurred image as input of the networks, and three corresponding clear images (x_1, x_2, x_3) of size are obtained in the same way.

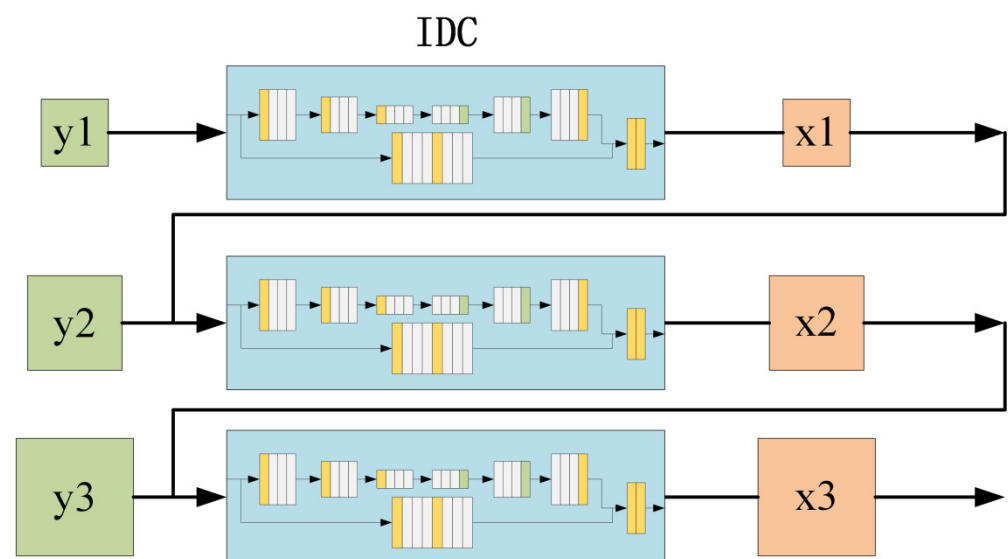


Figure 4. Multi-scale iterative strategy.

3.3. Mixed Loss Function

Using a single mean square error function as the loss function will cause some problems. For example, the reconstructed image output by the network tends to be flat, the contrast of the restored image is insufficient, and the detail textures are lost. This paper aims to design an end-to-end convolutional neural network to complete the deblurring task. On the one hand, the basic shape and structure of the object in the blurred image should be restored. On the other hand, the detailed texture of the image should be kept and restored as much as possible to obtain a high-quality restoration image. Considering both objectives, the final mixed loss function of this paper combines the \mathcal{L}_2 loss function, gradient, and structural similarity loss function. The \mathcal{L}_2 loss function can be expressed as:

$$\mathcal{L}_2 = \frac{1}{N} \sum_{i=1}^N \| f(y_i; \theta) - x_i \|^2 \quad (2)$$

In Equation (2), N represents the number of blur–clear image pairs for training. $f(y_i; \theta)$ represents the network output of the number i training image, where y_i is blurred image in training datasets, θ represents the network weight parameters, and x_i is the corresponding ground truth.

The SSIM takes human visual perception into consideration. A loss function based on SSIM can be expressed as:

$$\mathcal{L}_{SSIM} = \frac{1}{N} \sum_{i=1}^N (1 - \text{SSIM}(f(y_i; \theta), x_i)) \quad (3)$$

In Equation (3), $\text{SSIM}(f(y_i; \theta), x_i)$ represents the value of the SSIM between the output and the ground truth. If the gradient operator is represented by ∇ , then the gradient loss can be expressed as:

$$\mathcal{L}_{grad} = \frac{1}{N} \sum_{i=1}^N (\nabla f, \nabla x_i) \quad (4)$$

Therefore, the mixed loss function can be denoted as:

$$\mathcal{L}_{mix} = \mathcal{L}_2 + \alpha \mathcal{L}_{SSIM} + \beta \mathcal{L}_{grad} \quad (5)$$

α and β indicate the weighting factors. In this paper, both α and β take the values of 0.1. With three scales, the total mixed loss function is the sum of the mixed loss functions of the three scales:

$$\mathcal{L} = \mathcal{L}_{mix}^1 + \mathcal{L}_{mix}^2 + \mathcal{L}_{mix}^3 \quad (6)$$

The superscript in the above equation represents the corresponding order number of scales. To maintain the convergence consistency of the network at different scales, the mixed loss function was used as the unique loss function during training to update all the network parameters instead of being phased.

4. Experimental Results and Discussion

In order to compare the methods fairly, in this paper, the model was trained and tested using the open-source GOPRO dataset [25]. The dataset contains 3214 blur-clear image pairs (with a size of 1280×720). According to the same strategy, 2103 image pairs were used for training, while the remaining 1111 image pairs were used as test sets to evaluate performance. During the test, the images of the test dataset were fed into the trained model to generate the deblurring result. At last, the recovery results were evaluated using two objective indexes: peak signal-to-noise ratio (PSNR) and structural similarity (SSIM).

The experiment environment was based on Python, TensorFlow, Linux 16.04, graphic processing unit Nvidia 1080ti (NVIDIA Corporation, Santa Clara, CA, USA), and 11 G video RAM. During network training, the maximum number of iterative training was 1000, the learning rate was $1 \times 10^{-4} \sim 1 \times 10^{-6}$, the batch size was 16, and the parameter initialization mode was the Xavier method. Furthermore, the Adam optimization algorithm was adopted, and relevant parameters were set as default settings. To avoid overfitting, Gaussian random noise (with different standard deviation randomly sampled from Gaussian distribution, $N(0, (2/255)^2)$) was added to blurry images.

4.1. Objective Assessment

This section compared the proposed algorithm with several advanced image deblurring algorithms based on statistical prior or deep learning. Peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) were used to evaluate the image quality recovered by the algorithm, and the running time was used to evaluate the efficiency of the algorithm. The experimental results of each method are shown in Table 2, where the best results are in bold and the second-best results are underlined.

Table 2. Comparison of Different Algorithms.

Method	PSNR	SSIM	Runtime (per img)
Kim et al. [36]	23.60	0.823	1 h
Sun et al. [15]	24.60	0.842	20 min
Xu et al. [14]	25.10	0.890	12 s
Nah S et al. [25]	29.10	0.913	3 s
Kupyn et al. [33]	29.55	0.934	0.7 s
Tao et al. [29]	30.10	0.932	1.6 s
Gao et al. [26]	30.92/31.58	0.942/0.948	1.7 s
Our (proposed)	31.11	0.948	0.4 s

The compared algorithm includes method proposed by Kim et al. [36], Tao et al. [29], Xu et al. [14], Sun et al. [15], Nah S et al. [25], Gao et al. [26], and DeblurGAN-v2 proposed by Kupyn et al. [33]. Table 2 also shows the comparison of the mean SSIM, PSNR, and runtime (the average inference time for each test image) of the above methods on the GOPRO test set.

From Table 2, we can see that the IDC proposed in this paper obtains the highest SSIM on the GOPRO test set compared with the current advanced deblurring algorithms, including traditional model-based deblurring methods ([14,15,36]) and convolutional neural networks-based deblurring methods ([25,26,29,33]), and PSNR was only slightly lower than in [26], who point out that there exist flaws in some of the ground truth sharp images in the GoPro training set, including severe noise, large smooth region, and significant image blur. To improve the training performance, ref. [26] established a new dataset using GoPro Hero6 (GoPro Inc., San Mateo, CA, USA) and iPhone7 (Apple Inc., Cupertino, USA) at 240 fps. To be noted, the previous result (PSNR = 30.92, SSIM = 0.942) for [26] in Table 2 is obtained on the Gopro training dataset, while the result ((PSNR = 31.58, SSIM = 0.948) is obtained on the mixed training dataset, the Gopro dataset, and [26]’s new dataset). Additionally, benefitting from the relatively shallow network depth and the parallel computation of the dual-CNNs subnetworks, the proposed method is significantly faster than the other methods. Concluded from the above analysis and compared with the existing mainstream algorithm, the proposed IDC achieves remarkable image deblurring performance.

4.2. Subjective Assessment

Further, in order to more intuitively explain the performance of the proposed algorithm, four blurred images were selected in the GOPRO dataset to demonstrate the performance. The blur effect of each method is shown in Figure 5.

As shown in Figure 5c, the method proposed by Zhang et al. [24] still has the ringing phenomenon with a poor visual effect. Moreover, it is clear from Figure 5d that the results of DeblurGAN-v2 [33] are good, but the detail recovery is insufficient (the detail loss of the text area of the slogan on the van in the third image is serious). As can be seen from Figure 5e, when the image is seriously blurred, the deblurred results of [29] have artifacts and insufficient detail recovery (the billboard in the second image, the car outline in the first image, and the slogan on the van in the third image). We can see from Figure 5f,g that both [26] and the proposed method can achieve a better deblurred visual effect.

As shown in Figure 5, for the fast-moving car in the first image, the proposed method (IDC) can recover a better car contour. In the second image, the proposed method (IDC) recovers the license plate part more clearer. In addition, the removal of artifacts in the third image (the billboard part) recovered by IDC is even more obvious. In conclusion, the proposed deblurring network model has better visual quality in image recovery.



Figure 5. Experimental results on the GOPRO test dataset. (a) ground truths; (b) blurred images; (c) deblurring results of [24]; (d) deblurring results of [33]; (e) deblurring results of [29]; (f) deblurring results of [26]; (g) deblurring results of IDC (proposed).

4.3. Ablation Experiment

In order to verify the effectiveness of the multiscale iterative strategy adopted by the proposed algorithm, an ablation experiment is carried out. Four sets of deblurring results obtained by single-scale input and multiscale iterative input are compared as shown in Figure 6. It can be seen from Figure 6 that when there is only a single-scale blurred image as input, the deblurred image obtained by the networks has serious artifact problems, and the effect of deblurring is not satisfied. It is obvious that the deblurred image obtained by multiscale input is better than the single scale. The comparison and analysis of the above experiments demonstrate the multiscale iterative strategy's effectiveness.



Figure 6. Ablation experiment of multiscale iterative strategy. (a) ground truths; (b) blurred images; (c) deblurring results of single-scale; (d) deblurring results of multi-scale.

It can be seen from Figure 7 that the IDC (without Sub-Texture) achieves good deblurring performance, but in the case of a severe blurry degree, the deblurred image details are seriously lost. It can be seen from the partial enlargement of the deblurred results that IDC (with Sub-Texture) proposed in this paper achieves a better deblurring recovery effect. Particularly in the aspect of image detail restoration, the parallel Texture sub-networks can effectively enhance the performance of the networks in image detail preservation.



Figure 7. Ablation Experiment of Texture Sub-Network. (a) ground truths; (b) blurred images; (c) deblurring results of IDC without Sub-Texture; (d) deblurring results of IDC (proposed).

5. Conclusions

This paper presents an iterative dual-CNN procedure for image deblurring. The network consists of deblurring sub-networks, detail recovery sub-networks, and an information fusion module. Different from the previous approaches, we focus on the independence and complementarity between images semantic content and detail textures. In addition to the regular deblur branch based on U-net architecture, it is worth noting that we develop a textures extraction branch to ensure extracting complete detail textures features. The iterative dual CNNs can compensate for the inevitable information loss in traditional U-net architecture. Using the U-Net structure, the deblurring sub-network increases the receptive field through the input of multiscale networks, and it acquires multiple-scale information on the blurred image, which greatly improves the generalization ability of the network. The detail recovery sub-network adopts a shallow and wide network structure without downsampling. It keeps the feature maps and the input image at the same resolution, which can extract and retain the texture information of the blurred image to the maximum extent. The final deblurred image is obtained by merging the output of the deblurring and texture recovery sub-networks through the fusion module. The deblurred image obtained by the proposed models is rich in detail textures and has remarkable deblurring performance. Thus, it can be widely used in various spatial variable blur scenes.

Author Contributions: Conceptualization, J.W.; methodology, J.W. and Z.W.; software, J.W.; validation, J.W.; formal analysis, J.W. and A.Y.; investigation, J.W. and A.Y.; resources, J.W. and A.Y.; data curation, J.W. and Z.W.; writing—original draft preparation, J.W.; writing—review and editing, J.W. and A.Y.; visualization, J.W.; supervision, J.W. and A.Y.; project administration, A.Y.; funding acquisition, A.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (No. 62071323, 62176178).

Data Availability Statement: The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Tian, C.; Zhang, X.; Lin, J.C.W.; Zuo, W.; Zhang, Y.; Lin, C.W. Generative Adversarial Networks for Image Super-Resolution: A Survey. *arXiv* **2022**, preprint. arXiv:2204.13620.
2. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 295–307. [[CrossRef](#)] [[PubMed](#)]
3. Van Ouwerkerk, J.D. Image super-resolution survey. *Image Vis. Comput.* **2006**, *24*, 1039–1052. [[CrossRef](#)]
4. Zhang, Y.; Tian, Y.; Kong, Y.; Zhong, B.; Fu, Y. Residual dense network for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2472–2481.
5. Yang, C.Y.; Ma, C.; Yang, M.H. Single-Image Super-Resolution: A Benchmark. In *European Conference on Computer Vision*; Springer: Cham, Germany, 2014; pp. 372–386.
6. Tian, C.; Yuan, Y.; Zhang, S.; Lin, C.W.; Zuo, W.; Zhang, D. Image Super-resolution with An Enhanced Group Convolutional Neural Network. *arxiv* **2022**, preprint. arXiv:2205.14548. [[CrossRef](#)] [[PubMed](#)]
7. Goyal, B.; Dogra, A.; Agrawal, S.; Sohi, B.S.; Sharma, A. Image denoising review: From classical to state-of-the-art approaches. *Inf. Fusion* **2020**, *55*, 220–244. [[CrossRef](#)]
8. Zhang, Q.; Xiao, J.; Tian, C.; Chun-Wei Lin, J.; Zhang, S. A robust deformed convolutional neural network (CNN) for image denoising. *CAAI Trans. Intell. Technol.* **2022**. [[CrossRef](#)]
9. Zhang, K.; Zuo, W.; Zhang, L. FFDNet: Toward a fast and flexible solution for CNN-based image denoising. *IEEE Trans. Image Process.* **2018**, *27*, 4608–4622. [[CrossRef](#)] [[PubMed](#)]
10. Zheng, M.; Zhi, K.; Zeng, J.; Tian, C.; You, L. A Hybrid CNN for Image Denoising. *J. Artif. Intell. Technol.* **2022**, *2*, 93–99. [[CrossRef](#)]
11. Pan, J.; Sun, D.; Pfister, H.; Yang, M.H. Blind image deblurring using dark channel prior. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1628–1636.
12. Ren, W.; Cao, X.; Pan, J.; Guo, X.; Zuo, W.; Yang, M.H. Image deblurring via enhanced low-rank prior. *IEEE Trans. Image Process.* **2016**, *25*, 3426–3437. [[CrossRef](#)] [[PubMed](#)]
13. Chen, L.; Fang, F.; Wang, T.; Zhang, G. Blind image deblurring with local maximum gradient prior. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 1742–1750.
14. Xu, L.; Zheng, S.; Jia, J. Unnatural l0 sparse representation for natural image deblurring. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, 23–28 June 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 1107–1114.
15. Sun, J.; Cao, W.; Xu, Z.; Ponce, J. Learning a convolutional neural network for non-uniform motion blur removal. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 769–777.
16. Jain, V.; Seung, H.S. Natural image denoising with convolutional networks. In Proceedings of the International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–11 December 2008; pp. 769–776.
17. Hradis, M.; Kotera, J.; Zemčík, P.; Šroubek, F. Convolutional neural networks for direct text deblurring. In Proceedings of the British Machine Vision Conference, Swansea, UK, 7–10 September 2015.
18. Su, S.; Delbracio, M.; Wang, J.; Sapiro, G.; Heidrich, W.; Wang, O. Deep video deblurring for hand-held cameras. In Proceedings of the Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1279–1288.
19. Nimisha, T.M.; Singh, A.K.; Rajagopalan, A.N. Blur-invariant deep learning for blind-deblurring. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 4762–4770.
20. Vasu, S.; Maligireddy, V.R.; Rajagopalan, A.N. Non-blind deblurring: Handling kernel uncertainty with CNNs. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3272–3281.
21. Wieschollek, P.; Hirsch, M.; Schölkopf, B.; Lensch, H. Learning blind motion deblurring. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 231–240.
22. Kupyn, O.; Budzan, V.; Mykhailych, M.; Mishkin, D.; Matas, J. DeblurGAN: Blind motion deblurring using conditional adversarial networks. *arXiv* **2017**, preprint. arXiv:1711.07064.
23. Li, L.; Pan, J.; Lai, W.S.; Gao, C.; Sang, N.; Yang, M.H. Learning a discriminative prior for blind image deblurring. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 6616–6625.
24. Zhang, J.; Pan, J.; Ren, J.; Song, Y.; Bao, L.; Lau, R.W.; Yang, M.H. Dynamic scene deblurring using spatially variant recurrent neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2521–2529.
25. Nah, S.; Kim, T.H.; Lee, K.M. Deep multi-scale convolutional neural network for dynamic scene deblurring. In Proceedings of the Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3883–3891.
26. Gao, H.; Tao, X.; Shen, X.; Jia, J. Dynamic scene deblurring with parameter selective sharing and nested skip connections. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seoul, Korea, 27–28 October 2019; pp. 3848–3856.
27. Kupyn, O.; Martyniuk, T.; Wu, J.; Wang, Z. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27–28 October 2019; pp. 8878–8887.
28. Suin, M.; Purohit, K.; Rajagopalan, A.N. Spatially-attentive patch-hierarchical network for adaptive motion deblurring. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 3606–3615.

29. Tao, X.; Gao, H.; Shen, X.; Wang, J.; Jia, J. Scale-recurrent network for deep image deblurring. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8174–8182.
30. Zhang, X.; Wang, F.; Dong, H.; Guo, Y. A Deep Encoder-Decoder Networks for Joint Deblurring and Super-Resolution. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018.
31. Guo, T.; Kim, Y.; Zhang, H.; Qian, D.; Yoo, B.; Xu, J.; Zou, D.; Han, J.-J.; Choi, C. *Residual Encoder Decoder Network and Adaptive Prior for Face Parsing*; AAAI: Menlo Park, CA, USA, 2018.
32. Mao, X.; Shen, C.; Yang, Y.B. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. *Adv. Neural Inf. Process. Syst.* **2016**, *29*, 2802–2810.
33. Zeiler, M.D.; Krishnan, D.; Taylor, G.W.; Fergus, R. Deconvolutional networks. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13–18 June 2010.
34. Pan, J.; Liu, R.; Su, Z.; Gu, X. Kernel estimation from salient structure for robust motion deblurring. *Signal Process. Image Commun.* **2012**, *28*, 1156–1170. [[CrossRef](#)]
35. Zhang, X.; Dong, H.; Hu, Z.; Lai, W.S.; Wang, F.; Yang, M.H. Gated fusion network for joint image deblurring and super-resolution. *arxiv* **2018**, arXiv:1807.10806.
36. Hyun Kim, T.; Ahn, B.; Mu Lee, K. Dynamic scene deblurring. In Proceedings of the IEEE International Conference on Computer Vision, Sydney, Australia, 1–8 December 2013; IEEE Computer Society: Washington, DC, USA, 2013; pp. 3160–3167.