Image blind deblurring network with back projection feature fusion

Chi Li (lc61900@163.com)
Xi'an University of Posts and Telecommunications
Weiwei Kong
Xi'an University of Posts and Telecommunications
Jiawei Xue
Xi'an University of Posts and Telecommunications
Ze Wang
Xi'an University of Posts and Telecommunications
Liang Chang
Guilin University of Electronic Technology

Research Article

Keywords: image deblurring, u-net network, generative adversarial networks, back projection, self-attention mechanism, relative dis-criminator

Posted Date: May 31st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1684386/v1

Image blind deblurring network with back projection feature fusion

Chi Li[1,2], Weiwei Kong[1,2], Jiawei Xue[1,2], Ze Wang[1,2], Liang Chang[3,4]

1. Xi'an University of Posts and Telecommunications, Xi'an Shaanxi 710121, China
2. Shaanxi Provincial Key Laboratory of Network Data Analysis and Intelligent Processing, Xi'an 710121, China
3. Guilin University of Electronic Technology, Guilin, Guangxi 541004, China
4. Guangxi Key Laboratory of Trusted Software, Guilin, Guangxi 541004, China

Corresponding author: Chi Li (e-mail: lc61900@163.com).

ABSTRACT: Aiming at the problem of image motion blur caused by hand-held camera jitter and object motion in the process of collecting photos, a generative adversarial network (GAN) based on feature fusion of back projection is proposed for blind image deblurring. Firstly, the generator network is established by using U-Net structure, and a feature fusion residual block based on back projection is designed according to the error feedback principle, which solves the problem of saving spatial information in U-Net structure. Secondly, the self-attention module is introduced into the generator network to extract the feature map that pays more attention to detail. Finally, the combination of perceptual loss, mean square error loss and relative generative adversarial loss effectively alleviates the mode collapse problem of traditional GAN and improves the stability of model training. The experimental results show that the peak signal to noise ratio (PSNR) and structural similarity (SSIM) of this method on GoPro data set are 30.183dB and 0.941 respectively, and 26.962 and 0.837 on the Kohler dataset, with the shortest running time, which are better than the existing mainstream methods. The restored image is clearer in subjective vision and richer in texture details, which can effectively improve the image deblurring effect.

Key words: image deblurring; u-net network; generative adversarial networks; back projection; self-attention mechanism; relative discriminator

1 Introduction

Image blur, caused by many factors like camera shake, object movement and so on, is one of the important factors affecting image quality. Blurred images have poor usability and cause great difficulties for subsequent tasks like target detection and image seg-mentation.

The task of image deblurring is to recover the corresponding sharp image from the blurred image. It has been applied to various fields such as medical imaging, criminal case detection, and remote sensing images [1] in recent years. According to whether the blur kernel is known, it can be generally divided into two types: non-blind deblurring and blind deblurring. Early researches [3] mostly focused on non-blind deblurring, such as the Lucy-Richardson algorithm, which is based on Bayesian theory and assumes that the blur image satisfies the Poisson distribution, and deblurring by iteratively seeking maximum likelihood, however, due to the interference of noise factors, a reasonable number of iterations is the key to determine the quality of the restored image.

The research of blind deblurring becomes necessary because the blur kernel is mostly unknown in practice. Finding a blur function for each pixel is an ill-posed problem, and most of the existing algorithms [4] rely on heuristics, image statistics. These methods first obtain the prior information of the image, model the information and fit a function that is consistent with the prior information distribution of the sharp image, and use the fitted function as the blur kernel to estimate the sharp image. These methods run faster on a single image, but these priors can only deal with a limited type of blur image. If prior knowledge is invalid, the deblurring effect will be significantly reduced, so the generalization performance of the method is poor. Huang *et al.* [7] proposed a matrix variable optimization method to obtain latent sharp images. That is to carry out singular value decomposition of the fuzzy kernel matrix and minimize the matrix variable optimization problem with blur kernel constraints, so as to better estimate the blur kernel and clear image.

With the rapid development of deep learning, deblurring algorithms based on deep learning have been proposed. Sun *et al.* [8] used convolutional neural network (CNN) to predict the probability distribution of motion blur and then estimated the image blur kernel, Chakrabarti *et al.* [9] predicted the Fourier coefficient of the blur kernel and then deblurred in the Fourier space, but these methods are too time-consuming and the deblurring results rely too much on blur kernel estimation. Nah *et al.* [10] proposed a multi-scale neural network called Deep-Deblur. It adopts an end-to-end method from coarse to fine, and does not need to estimate the distribution of the blur kernel, greatly improves the performance and effect of deblurring. However, because the parameters of different scales of the model are independent of each other, problems such as excessive model parameters, unstable training, and easy overfitting occur. Tao *et al.* [11] proposed the scale recurrent network

(SRN), which combines multi-scale and recurrent networks to share parameters between modules at different scales, which greatly improves the deblurring performance, but trains a multi-scale model still requires a lot of memory and also very time-consuming.

Generative adversarial networks (GAN) [12] are widely used in computer vision because they can generate high-quality images. Kupyn *et al.* [13] proposed the DeblurGAN algorithm, which realizes end-to-end image deblurring based on conditional GAN; then the team proposed the Deblur-GANv2 [14] algorithm, which uses the pyramid model for image deblurring to generate high-quality images while greatly improving computational efficiency; Wu *et al.* [15] proposed a circular prior GAN and used a stacked estimation residual network for blind deblurring, and obtained high-quality deblurring. However, the restored images generated by these methods still have problems such as texture details and unclear edge structures.

In this paper, we take GAN as the basic framework and propose a blind image deblurring network based on multi-scale feature fusion. We make three contributions.

1) We use U-Net [16] as the generator framework, and design a back-projection-based feature fusion residuals block (FFRB) according to the principle of error feedback, which effectively fuses features of different levels and solve the problem of spatial information preservation in U-Net structure, and preserve the spatial information of high-resolution features. It is suitable for methods with less training data, and can extract feature maps with more detailed information.

2) The convolutional block attention module (CBAM) [17] is introduced into the generator, which make the feature map pay more attention to the specific details of the target in the image and highlight the features of the important target in the image.

3) Use the idea of adversarial learning to generate deblurred images with more natural details and textures. In terms of loss function, we use the relative average discriminator [18] to determine the probability that the real image is more realistic than the fake image. And using perceptual loss, mean square error loss and relative generative adversarial loss to seek joint minimization and effectively alleviate the problems of gradient disappearance and mode collapse in standard GAN. Make the training process more stable and efficient.

2 Model structure

In this paper, the network framework is based on GAN [12], is to define a game between two competing networks: the discriminator and the generator. The generator receives noise as an input and generates a sample. A discriminator receives a real and generated sample and trying to distinguish between them. The goal of the generator is to fool the discriminator by generating perceptually convincing samples that can't be distinguished from the real one. The goal of the discriminator is to receive real data and generated data, and make a distinction.

2.1 Generator network

In this paper, the generator network uses U-Net as the framework to build a simple, robust and computationally efficient network model for the single image deblurring problem, as shown in Fig. 1. It consists of three modules: encoder module, attention module called CBAM and decoder module. In this paper, in order to ensure that the images generated by the network achieve the best visual quality, we not only perform end-to-end residual learning for the generator network, but also performs skip connections between the encoding layer and the decoding layer with the same number of feature channels, so that the encoded information is passed to the decoder, allowing the network to retain more detailed information.

Encoder: in order to obtain sufficient feature information, the blurred image is input into the encoder for learning. To obtain a larger receptive field, the first layer of the encoder uses a 7×7 convolution kernel to convolve the 256×256 sized input map into 32 feature maps of the same size. Features are then extracted through a dense residual block. Then, we sequentially stack four down-sampling blocks to achieve gradual dimensionality reduction and increase of feature channels, all of which consist of a convolutional layer with stride 2, a FFRB module, and a dense residual block [20]. The final output is 256 feature maps of size 16×16. The FFRB module will be discussed in sections 2.1.1.

CBAM module: in order to make the learned features pay more attention to the details of the target object in the image, generate a picture with clearer details. We add a lightweight attention module CBAM to the generator network. Since CBAM is a lightweight module, it can be integrated into any CNN architecture with negligible overhead. The CBAM module will be discussed in Section 2.1.2.

Decoder: once the required hidden encoded features are learned, the decoder needs to up-sample these feature maps to size 256×256. The input of size 16×16 is sequentially passed through four up-sampling blocks to achieve gradual dimensionality upscaling and feature channel reduction, all of which consist of a stride 2 deconvolutional layer, a dense residual block, and an FFRB module. Finally, a 3×3 ordinary convolutional layer is further used to output the deblurred image.

2.1.1 FFRB module

The U-Net architecture encoder lacks spatial information during down-sampling and lacks sufficient connections between features in non-adjacent layers. The back-projection [19] technique in the image super-resolution task is an effective feature fusion method. First, the high-resolution input is down-sampled to obtain low- resolution feature map and then the error is calculated with the real

low-resolution feature map, and an efficient iterative

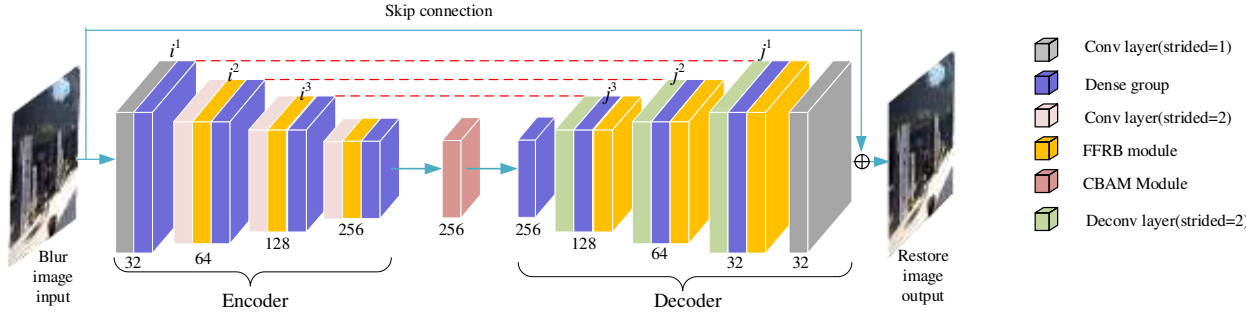


Fig.1 Generator network in this paper

process minimizes the error, based on which the high-resolution output is optimally reconstructed to achieve enhanced feature maps.

Inspired by the back-projection algorithm, we propose the FFRB module. This module uses the generator network to enhance the extracted feature map through the error feedback mechanism, effectively repairing the missing information caused by the cross-scale operation of the network and making full use of the features of non-adjacent layers. The positions of the modules in the generator are shown as the yellow modules in Fig. 1. In the encoder and decoder, the FFRB module performs feature fusion before and after the dense residual block, respectively. The FFRB module of the encoder is defined as follows.

scale up: $$H^t = (L_0^t * p_t)\uparrow_{s2} \quad (1)$$

residual: $$e^h = H^t - H^0 \quad (2)$$

scale residual down: $$L_1^t = (e^h * q_t)\downarrow_{s2} \quad (3)$$

output: $$L^t = L_1^t + L_0^t \quad (4)$$

where '*' represents the convolution operation, '$\uparrow_{s2}$' and '$\downarrow_{s2}$' represent s times of up and down-sampling operations with a step size of 2, respectively. And pt and qt represent t times of transposed convolution and convolution operation, respectively. t represents the number of network layers.

As shown in the upper part of Fig. 2, the low-resolution feature map L_0^t is used as the input of the encoder FFRB module, and the 256×256 high-resolution feature map H^t is obtained through up-sampling operation, Then, e^h, the residual between H^t and feature graph H^0 at i^0 layer of encoder network is calculated. The error feature map L_1^t is obtained by down-sampling the residual, where the dimension of L_1^t is the same as that of L_0^t. Finally, the enhanced feature map L^t is obtained by summing the two low-resolution feature maps.

The FFRB definition of the decoder is similar to that of the encoder, as shown in the lower half of Fig. 2, That is, the high-resolution feature map H_0^t is down-sampling to a size of 16 × 16 and named L^t, then e^l is the residual between L^t and the output feature map L^3 of the decoder network j^3 layer, and the residual feature is up-sampling and summed with the input feature map of this module to obtain the enhanced feature map H^t.

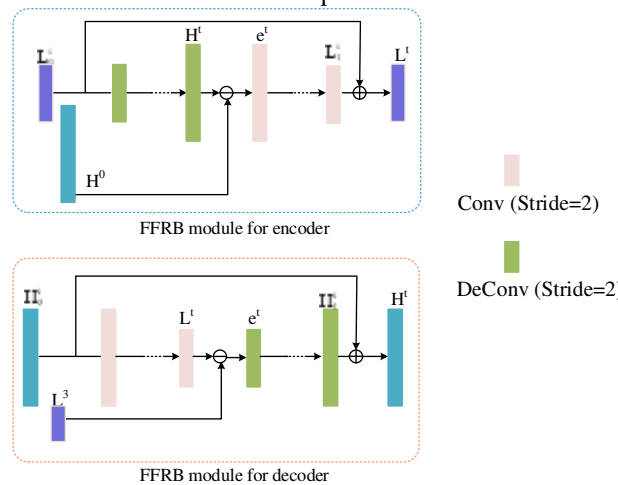


Fig.2 FFRB modular in this paper

scale down: $$L^t = (H_0^t * q_t)\downarrow_{s2} \quad (5)$$

residual: $$e^l = L^t - L^3 \quad (6)$$

scale residual up: $$H_1^t = (e^l * p_t)\uparrow_{s2} \quad (7)$$

output: $$H^t = H_1^t + H_0^t \quad (8)$$

Compared with serial fusion method, FFRB can better extract high-dimensional information from high-resolution features on the encoder module due to the applied error feedback mechanism. By fusing these feature differences back to down-sampled low-dimensional features, the lost spatial information can be better repaired. On the other hand, the module uses these feature differences to enhance the input feature map to a certain extent. In order to obtain more sufficient feature information and better deblurring results.

2.1.2 CBAM module

The CBAM [17] module includes a channel attention module and a spatial attention module. This module enables the network to pay more attention to the specific details of the target in the image while learning the image features, and highlights the characteristics of the important

target in the image, thereby generating higher-quality pictures. The network structure of the module is shown in Fig. 3.

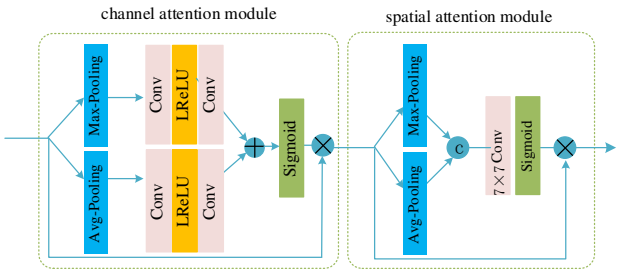


Fig.3 CBAM modular

The channel attention module aims to pay attention to what is meaningful in the input feature maps, that is, to judge the importance of the input feature channels, and then assign appropriate weights, so that the network can extract more useful and detailed feature maps. The spatial attention module complements the channel attention module and aims to focus on which are the most informative parts of the feature map.

2.2 Loss function

There are also many problems with the standard GAN, such as gradient vanishing, mode collapse [21]], etc. In the standard GAN, the role of the discriminator is to determine the probability that the input image is real, and the training process only continuously increases the probability that the fake data is real, which will lead to the scores of the real data and the generated data approaching 1. And cannot account for a prior knowledge that half of the discriminator input samples are fake.

With the emergence of the relative GAN [18], this problem is well solved. Use a relative discriminator to calculate the probability that the given real data is more real than the fake data generated by the generator. The final score of the discriminator converges to 0.5, which is in line with the prior knowledge and increase training stability. In this paper, a relative average discriminator (RaD) is used to replace the discriminator loss of the traditional GAN, which makes the model training more stable and efficient.

We modify the loss function of the generator and the discriminator. The calculation formulas are shown in equations (9) and (10) respectively:

$$L_D^{Ra} = -E_{X_r}[\log(D_{Ra}(x_r, x_f))] - E_{X_f}[\log(1 - D_{Ra}(x_f, x_r))] \tag{9}$$

$$L_G^{Ra} = \lambda E_{X_r}[\log(1 - D_{Ra}(x_r, x_f))] - E_{X_f}[\log(D_{Ra}(x_f, x_r))] + \eta L_{per} + \mu L_2 \tag{10}$$

L_D^{Ra} is the discriminator loss function, L_G^{Ra} is the generator loss function, where X_r represents the real sharp image, X_f represents the generated deblurred image, and E represents the mathematical expectation. L_{per} is the perceptual function, and L_2 is the Mean Square Error (MSE) loss. In our experiment, λ is 0.01, η is 0.006, and μ is 0.5. L_2 is the average of the sum of squares of the difference between the predicted value and the target value. We no longer use the L_2 loss alone, because the loss function is averaged over the pixel space, and as the only optimization, the image is still blurred in detail. we use perceptual loss [24] L_{per} add L_2 loss and seeks joint minimization as content loss. Makes the restored image better in detail.

2.3 Discriminator network

The traditional discriminator network scores the authenticity of the entire image, which consumes time and generates low-quality images. We use PatchGAN [23] as the discriminator network, which maps the input image to an N×N matrix, it is judged that the image is a real image or a generated image through the calculation of the mean value of the matrix. Compared with the traditional discriminator, the advantage of the PatchGAN discriminator is that its output is a matrix, which can fully consider the influence of different areas of the image, so that the model pays more attention to the details of the picture, which helps the generator to generate pictures with better details and sharper edges. The specific structure of the discriminator is shown in Fig. 4, with a total of 5 layers. Except for the first and last layers, the other three layers are added with an instantiated normalization layer and an LReLU [24] activation function. After the last layer, the Sigmoid function is used to activate the output, and n represents the number of feature maps. s represents the step size.

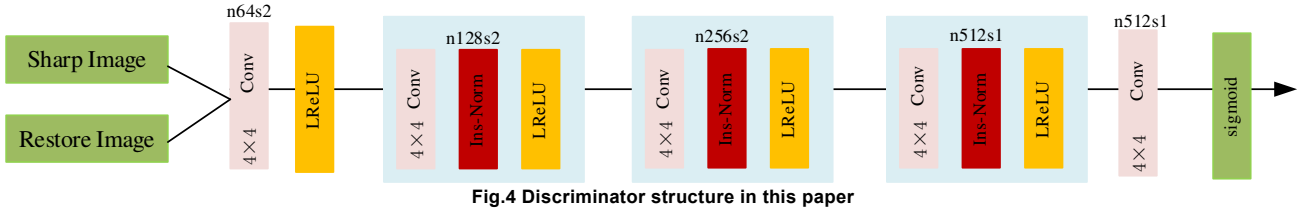


Fig.4 Discriminator structure in this paper

3 Experiments results with relevant analysis

In order to verify the efficiency of our method, this section discusses the experimental data sets, experimental settings and results analysis, and compares them with the current mainstream deblurring methods from the subjective visual level and objective evaluation indicators and get convincing results.

3.1 Experimental settings

Our experiments use the pytorch deep learning open-source framework, the CPU configuration is Intel Core i7-7700HQ, the GPU configuration is NVIDIA GeForce GTX 2080Ti, the memory size is 12GB, and the operating

system is Ubuntu18.04 to train the model on the platform. The GoPro training set is randomly cropped into a 256×256 image as the input of the model. The optimizer selects Adam optimizer. A total of 400 epochs were trained, and the batch size was set to 1. In our experiment, the learning rate is set to 0.0001 for the first 120 times, and then the learning rate decays to the original 0.75 every 20 times.

Our experimental training data uses the GoPro [10] dataset, which consists of 2103 pairs of blurred and sharp image pairs with a resolution of 720p, and a total of 1100 pairs of images are used for testing in the test set.

3.2 Comparative experiment

3.2.1 Network module effectiveness analysis

In order to verify the effectiveness of adding related modules to the generator network for image deblurring tasks. In this section, the convolutional residual block [25], dense residual block, CBAM module and FFRB module are added to the autoencoder model respectively, and the subjective visual effects and objective evaluation indicators are compared respectively. The performance comparison of each module is shown in Table 1. The deblurring effect on the GoPro test set is shown in Fig. 5.

It can be seen from Table 1 that after replacing the residual block with the dense residual block, the peak signal to noise ratio (PSNR) is increased by 0.685, and the structural similarity (SSIM) is increased by 0.019; on this basis, the effect of adding the FFRB module is significantly improved, and the PSNR and SSIM are increased by 1.825dB and 0.056 respectively; In order to further improve the deblurring effect of the model, the CBAM lightweight attention module is added on this basis, and the optimal index is reached at this time.

It can be seen from the restoration of the car and the window in Fig. 5 that the effect of adding three modules to the autoencoder at the same time is the best, so this paper chooses to integrate three modules into the generator network for image deblurring.

Table 1 Objective comparison of different methods on GoPro

	PSNR	SSIM
Res-Block	26.957	0.857
Des-Block	27.642	0.876
Des + FFRB	29.467	0.932
Des + FFRB + CBAM	**30.183**	**0.941**

3.2.2 Loss function effectiveness analysis

In order to verify the effectiveness of the loss function in this paper on the image deblurring task, this section uses different loss functions to test on the GoPro test set under the premise that the generator network is optimal. The objective evaluation indicators and subjective effects are shown in Table 2 and Fig. 6, respectively.

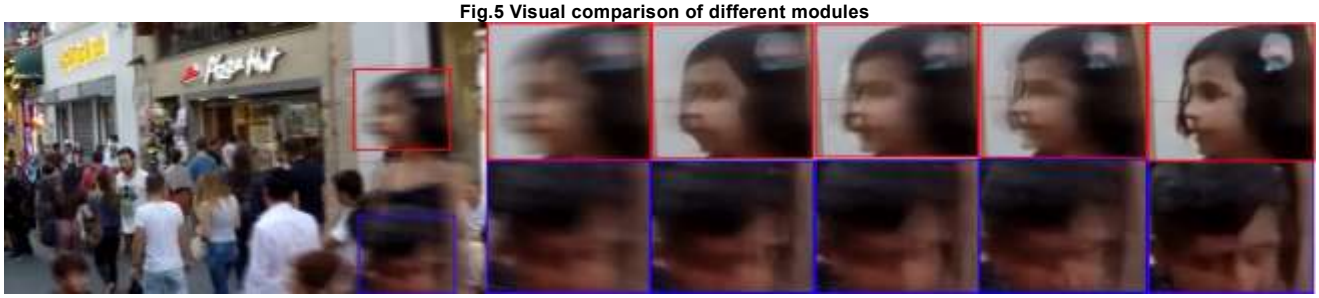
It can be seen from the Table 2 and the face restoration in Fig. 7 that the effect of the loss function using RaGAN is better than standard GAN; on this basis, adding L_2 loss further improves the performance; in this paper, the combination of RaGAN, L_{per} [26], and L_2 losses is used to achieve the optimal result, which is improved compared with other comparison groups, which proves the effectiveness of the loss function in this paper.

Table 2 Performance comparison of different loss function

	PSNR	SSIM
GAN	28.055	0.904
RaGAN	29.173	0.923
RaGAN + L_2	29.343	0.931
RaGAN + L_2 + L_{per}	**30.183**	**0.941**

Blur Image (a)blur (b)Res-Block (c)Des-Block (d)Des + FFRB (e)Ours

Fig.5 Visual comparison of different modules



Blur Image (a)blur (b)GAN (c)RaGAN (d)RaGAN+L_2 (e)Ours

Fig.6 Visual comparison of different loss function

3.2.3 GoPro dataset comparison

We conduct image deblurring experiments on the proposed method and other methods on the GoPro test set. The comparison results of the objective evaluation indicators of different method are shown in Table 3, and two images are selected to show the subjective effects as shown in Fig. 7. Among them, although the method of Xu *et al.*

[5] and Kim et al. [6] is effective, it is poor in processing the edge part of the image. Since the scene and blur degree of the GoPro dataset are not the same, the method of Sun *et al.* [8] cannot correctly estimate the blur kernel on the GoPro dataset, and the deblurring is completely invalid. The image generated by the method of Isola *et al.* [23] has larger noise and part of the image content is lost, so the evaluation index is the lowest. Although Deep-Deblur [10] has better deblurring effects than the previous methods, the details of the recovered images are still blurred. Although the Deblur-GAN [13] method is superior to other methods, the clarity still needs to be improved; Although Wu *et al.* [30] and the DPPHN [29] method have obvious advantages in evaluation indicators, but the images generated by these two methods have some irregularity, as shown in Fig. 7 distorted fonts on license plates and posters. In this paper, this situation is better improved, and the best 0.941 is achieved on SSIM, which is 0.017 higher than the method of Wu et al; however, because our loss function does not directly act on PSNR, so our method is suboptimal in PSNR. As can be seen that the deblurring model proposed in this paper has obvious advantages over other method in the image deblurring task.

For the single image deblurring task, the running time of each method is also an important indicator to measure the method. The shorter the running time, the smaller the algorithm complexity of the model. This paper selects two blurred images in the test set to test the processing time of the above algorithms and compare them. The tests are based on GPU acceleration on the same platform. Each method is run 5 times for each image, and the average of the 5 running times is taken as the final result. It can be seen that our method has the fastest processing time for a single image.

Table 3 Objective comparison of different methods on GoPro

	Xu et al	Kim et al	Cnn-Deblur	Isola et al	Deep-Deblur	Deblur-GAN	DPPHN	Wu et al	Ours
PSNR	24.813	23.641	23.328	21.619	29.085	28.726	29.826	**30.25**	30.183
SSIM	0.884	0.842	0.823	0.805	0.903	0.917	0.909	0.923	**0.941**
Times/s	13.413	1800	1200	1.483	4.302	1.553	8.578	3.261	**1.172**

Blur Image (a)Blur (b)Xu et al (c)Sun et al (d)Isola et al (e)Deep-Deblur (f)Deblur-GAN (g) Wu et al (h)Ours

Fig.7 Comparison of deblur effects of different methods on GoPro data set

3.2.4 Kohler dataset comparison

The Kohler dataset [27] is a commonly used dataset for evaluating image deblurring methods. It consists of 48 images, including 4 scenes and 12 blurred images of different degrees. First, the camera is used to shoot the moving video, and then the video is played back on the machine platform to analyze the movement trajectory of the camera and obtain the blur kernel to obtain the blurred image. The evaluation index adopts PSNR and SSIM.

The comparison results of the evaluation indicators of different methods are shown in Table 4, and the subjective effect of deblurring of two blurred images of different scenes under the blessing of different algorithms is shown in Fig. 8. Since there are as many as 12 blur kernels on the Kohler data, the methods of Kim *et al.* and Sun *et al.* cannot estimate all the fuzzy kernels well, which leads to the failure of these two methods on the Kohler data set, and the evaluation index scores are both low. The method of Isola *et al.* still has problems such as high noise and loss of image content, so the performance is poor. Compared with the previous ones, the Deep-Deblur method and Deblur-GAN have improved significantly, but the details are still unsatisfactory. In contrast, our method is visually clear and can clearly restore image details such as fonts on clocks and patterns on buildings. In terms of evaluation indicators, our method achieves the best performance, with PSNR and SSIM higher than Wu's method by 0.131 and 0.016, respectively. And the running time is also the shortest.

In summary, the restoration performance of our method

on the Kohler dataset is overall better than other methods, which also proves the effectiveness of our method.

Table 4 Objective comparison of different methods on Kohler

	Xu et al	Cnn-Deblur	Isola et al	Deep-Deblur	Deblur-GAN	DPPHN	Wu et al	Ours
PSNR	25.684	24.224	22.574	26.487	26.103	26.54	26.831	**26.962**
SSIM	0.794	0.773	0.732	0.808	0.816	0.819	0.821	**0.837**
Times/s	14.382	840	1.415	5.233	1.652	9.327	4.549	**1.366**

(a)Blur Image (b)Isola et al (c)Deep-Deblur (d)Deblur-GAN (e)Ours

Fig.8 Comparison of deblur effects of different methods on Kohler data set

3.2.5 Real dataset comparison

Most deblurred test image datasets are synthesized using high-speed cameras and still differ from real blurred images. Lai [28] proposed a real blur dataset, which contains 100 blurred images of different qualities and resolutions collected in different types of scenes. There is no corresponding clear image corresponding to these blurred images, so the corresponding objective evaluation index results cannot be produced.

In this paper, two images are selected for visual subjective comparison with the Deep-Deblur method and the Deblur-GAN method respectively. The results are shown in Fig. 9 and Fig. 10. The real images recovered by the Deep-Deblur method and the Deblur-GAN method have a high degree of blur, and the small text part in Fig. 9 is almost unrecognizable. The method in this paper is optimal in the overall and local vision of the image, and can better restore the details of the image as shown in the small text in Fig. 9 and the word "JOY" in Fig. 10.

(a)Blur Image (b)Deep-Deblur (c)Deblur-GAN (d)Ours

Fig.9 Comparison of deblur effects of different methods on real blur data set (Image 1)

(a)Blur Image (b)Deep-Deblur (c)Deblur-GAN (d)Ours

Fig10 Comparison of deblur effects of different methods on real blur data set (Image 2)

4 Conclusion

Aiming at the problem of image blurring, we propose an

image deblurring method based on GAN as the basic framework and based on back-projection feature fusion, which adopts an end-to-end method to perform blind image deblurring in the case of unknown blur kernel. The generator of our method is based on the U-Net network, which integrates the FFRB module and the attention module; uses the PatchGAN model as the discriminator network to speed up network training; uses the relative generative adversarial loss to better ensure the stability of model training, making the model more robust. Finally, through experiments, the deblurring effect of this method and other classical mainstream methods is visually compared, and through the objective comparison and analysis of the three evaluation indicators of PSNR, SSIM and processing time, it is proved that the deblurring effect of this method in different data sets is significant. It is also superior to most methods in terms of computational efficiency and has high application value.

Acknowledgements This work was supported in part by the National Natural Science Foundation of China under Grant 61772396 and 61902296, Natural Science Foundation of Shannxi Province of China under Grant 2022JM-369, and the Funded by Guangxi Key Laboratory of Trusted Software Research Project KX202061.

Ethics approval and consent to participate This article does not involve relevant ethical issues, and the relevant authors agree to participate

Availability of data Data sets generated during the current study are not publicly available due to funding restrictions, but are available from the corresponding authors upon reasonable request.

Consent for publication We have read and understood the publication policy and agree to submit this manuscript.

Competing interests This article has no relevant competing interests.

Authors' contributions Chi Li and Weiwei Kong wrote the main manuscript text and Jiawei Xue prepared for relevant ablation experiments and prepared figures 1-4. Ze Wang and Liang Chang prepared figures 5-10 and Table 1-4. All authors reviewed the manuscript.

REFERENCES

[1] J. Xiao, Z. Jin, and H. Zhang, “A general model compression method for image restoration network,” *Sign Process: Image Comm*, Vol. 93, pp. 182-192, 2021.

[2] M. S. Hosseini, K. N. Plataniotis, “Convolutional deblurring for natural imaging,” *IEEE Trans. Image Process*, Vol. 29, pp. 250-264, 2019.

[3] T. Li, Y. Yuan, and B. Zhang, “Experimental verification of three-dimensional temperature field reconstruction method based on Lucy-Richardson and nearest neighbor filtering joint deconvolution algorithm for flame light field imaging,” *Appl Therm Engineer*, Vol.162, pp. 325-334, 2019.

[4] D. Perrone, P. Favaro, “Total variation blind deconvolution: The devil is in the details,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 2909-2916, 2014.

[5] L. Xu, S. Zheng, and J. Jia, “Unnatural l0 sparse representation for natural image deblurring,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 1107-1114, 2013.

[6] T. H. Kim and K. M. Lee, “Segmentation-free dynamic scene deblurring,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 2766-2773, 2014.

[7] L. Huang, Y. Xia and T. Ye, “Effective blind image deblurring using Matrix-Variable optimization,” *IEEE Trans. Image Process*, Vol. 30, pp. 4653-4666, 2021.

[8] J. Sun, Wenfei Cao, Zongben Xu and J. Ponce, “Learning a convolutional neural network for non-uniform motion blur removal,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 769-777, 2015.

[9] A. Chakrabarti, “A neural approach to blind motion deblurring,” *Europ Conf. Comp Vis*, pp. 221-235, 2016.

[10] S. Nah, T. H. Kim and K. M. Lee, "Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring," *IEEE Conf. Comp Vis. Pattern Recog*, pp. 3883-3891, 2017.

[11] X. Tao, H. Gao, X. Shen, J. Wang and J. Jia, "Scale-Recurrent Network for Deep Image Deblurring," *IEEE Conf. Comp Vis. Pattern Recog*, pp. 8174-8182, 2018.

[12] I. Goodfellow, J. Pouget-Abadie, and M. Mirza, “Generative adversarial nets,” *Advan in Neural Infor Process Sys*, pp. 2672-2681, 2014.

[13] O. Kupyn, V. Budzan and M. Mykhailych, “Deblur-GAN: Blind motion deblurring using conditional adversarial networks,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 8183-8192, 2018.

[14] O. Kupyn, T. Martyniuk, and J. Wu, “Deblur-GANv2: Deblurring (orders-of-magnitude) faster and better,” *IEEE Inter Conf. Comp Vis*, pp. 8878-8887, 2019.

[15] J. Wu, X. Di, “Integrating neural networks into the blind deblurring framework to compete with the end-to-end learning-based methods,” *IEEE Trans. Image Process*, Vol 29, pp. 6841-6851, 2020.

[16] O. Ronneberger, P. Fischer, T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” *Intern Conf. Medic Image Comp and Comp-Assis Inter*, pp. 234-241, 2015.

[17] S. Woo, J. PARK, and J. Y. LEE, “CBAM: Convolutional block attention module,” *Europ Conf. Comp Vis*, pp. 3-19, 2018.

[18] M. A. Jolicoeur, “The relativistic discriminator: a key element missing from standard GAN,” *arXiv preprint arXiv:1807.00734*, 2018.

[19] M. Haris, G. Shakhnarovich and N. Ukita, “Deep back-projection networks for super-resolution” *IEEE Confe. Comp Vis. Pattern Recog*, pp. 1664-1673, 2018.

[20] G. Huang, Z. Liu and D. Van, “Densely connected convolutional networks,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 4700-4708, 2017.

[21] T. Salimans, I. Goodfellow and W. Zaremba, “Improved techniques for training gans,” *Advan. Neural Infor Process Sys*, Vol. 29, pp. 2234-2242, 2016.

[22] M. Arjovsky, S. Chintala and L. Bottou, “Wasserstein generative adversarial networks,” *Intern Conf. Mach Learn. PMLR*, pp. 214-223, 2017.

[23] P. Isola, J. Y. Zhu, and T. Zhou, “Image-to-image translation with conditional adversarial networks,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 1125-1134, 2017.

[24] K. He, X. Zhang and S. Ren, “Delving deep into rectifiers: Surpassing human-level performance on image-net classification,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 1026-1034, 2015.

[25] K. He, X. Zhang and S. Ren, “Deep residual learning for image recognition,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 770-778, 2016.

[26] C. Ledig, L. Theis and F. Huszár, “Photo-realistic single image super-resolution using a generative adversarial network,” *IEEE Conf. Comp Vis. Pattern Recog*, pp. 4681-4690, 2017.

[27] R. Köhler, M. Hirsch and B. Mohler, “Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database,” *Europ Confe. Comp Vis*, pp. 27-40, 2012.

[28] W. S. LAI, J. B. HUANG and Z. HU, “A comparative study for single image blind deblurring,” *IEEE Confer. Comp Vis. Pattern Recog*, pp. 1701-1709, 2016.

[29] Z. Zhao, B. Xiong, S. Gai and L. Wang, "Improved Deep Multi-Patch Hierarchical Network With Nested Module for Dynamic Scene Deblurring," *IEEE Access*, Vol. 8, pp. 62116-62126, 2020.

[30] Y. Wu, P. Qian and X. Zhang, "Two-Level Wavelet-Based Convolutional Neural Network for Image Deblurring," *IEEE Access*, Vol. 9, pp. 45853-45863, 2021.