

# Generative Adversarial Network Based on Multi-feature Fusion Strategy for Motion Image Deblurring

Zhou-Xiang Jin, Hao Qin\*

School of Tai Chi Martial Arts, Jiaozuo University, 3066 Renmin Road, Jiaozuo City, Henan Province, 454000 China  
snowycry@qq.com;qinhao123@126.com

Received 27 September 2021; Revised 15 October 2021; Accepted 27 October 2021

**Abstract.** Deblurring of motion images is a part of the field of image restoration. The deblurring of motion images is not only difficult to estimate the motion parameters, but also contains complex factors such as noise, which makes the deblurring algorithm more difficult. Image deblurring can be divided into two categories: one is the non-blind image deblurring with known fuzzy kernel, and the other is the blind image deblurring with unknown fuzzy kernel. The traditional motion image deblurring networks ignore the non-uniformity of motion blurred images and cannot effectively recover the high frequency details and remove artifacts. In this paper, we propose a new generative adversarial network based on multi-feature fusion strategy for motion image deblurring. An adaptive residual module composed of deformation convolution module and channel attention module is constructed in the generative network. Where, the deformation convolution module learns the shape variables of motion blurred image features, and can dynamically adjust the shape and size of the convolution kernel according to the deformation information of the image, thus improving the ability of the network to adapt to image deformation. The channel attention module adjusts the extracted deformation features to obtain more high-frequency features and enhance the texture details of the restored image. Experimental results on public available GOPRO dataset show that the proposed algorithm improves the peak signal-to-noise ratio (PSNR) and is able to reconstruct high quality images with rich texture details compared to other motion image deblurring methods.

**Keywords:** motion image deblurring, generative adversarial network, multi-feature fusion strategy, deformation convolution module, channel attention module

## 1 Introduction

The purpose of motion image deblurring is to reconstruct and estimate an unknown clear image for a motion blurred image. Traditional motion image deblurring methods estimate the parameters and model of the blur kernel, and use different natural priors in the optimization of the objective function, such as dark channel priors [1], data-driven priors [2] etc., to improve the accuracy of parameter estimation and deblurring quality. However, traditional methods are limited by fuzzy kernel estimation and inaccurate modeling of fuzzy sources, resulting in poor image reconstruction quality. In addition, the motion image deblurring method based on deep learning can better capture the image features and improve the deblurring effect through the powerful feature learning ability of the convolutional neural network (CNN). However, because the geometric structure of the standard convolution is fixed, it has great limitations in extracting image features from non-uniform motion blurred images. Secondly, in network design, deepening the network model is used to expand the local receptive field, thereby, it can capture high-frequency information, but too deep network models can easily lead to training difficulties.

Aiming at the shortcomings of the above-mentioned motion image deblurring methods based on deep learning, this paper proposes a motion image deblurring method based on generative adversarial network via multi-feature fusion strategy. This method constructs an adaptive residual module (ARM) composed of a deformation convolution module (DCM) [3] and a channel attention module (CAM) [4] in the generative network. The deformation convolution module adapts the convolution kernel to the size and shape of the extracted image features, prompting the network to adapt to the non-uniformity caused by the deformation of the motion blur image from the spatial level. The channel attention module adaptively rescales the features, so that the network can not only obtain a larger global receptive field, but also enhance the ability of deep network discrimination learning, it pays more attention to the learning of high-frequency features, and reconstructs the texture details of the image.

---

\* Corresponding Author

## 2 Related Research

At present, based on the rapid development of deep learning in image restoration, the application of deep learning to the research of motion image deblurring has achieved remarkable results. Sun et al., [5] used a deep convolutional neural network to estimate the probability of the motion kernel at the patch level of the image, and then used Markov random fields to fuse the patch-based estimation into the dense domain of the motion kernel. At the same time, in order to improve the performance in motion kernel estimation, image rotation technology is used to expand the candidate motion kernel set estimated by CNN. Compared with the traditional method of manually estimating the blur kernel, the powerful feature learning ability of CNN can effectively estimate the motion kernel of spatial variation, and can better eliminate the motion blur.

Gong et al. [6] expressed the non-uniform motion blur as linear motion blur in the pixel direction, and the proposed method used a full-convolutional neural network (FCNN) to estimate the dense motion flow image for blur removal. Compared with adding an image prior to estimating the blur kernel, learning the motion flow could make the model focus on the cause of blur, without the need to model the image content, reducing the complexity of learning. References [5, 6] used convolutional neural networks to estimate the blur kernel, and then used traditional non-blind deblurring algorithms to obtain clear images in motion blur. In order to avoid the ill-posedness caused by the estimation of the fuzzy kernel and the complicated calculation of the non-blind deblurring algorithm, reference [7] used an end-to-end multi-scale CNN strategy to decompose the complexity of the deblurring task, which could restore clear images under dynamic conditions such as occlusion, motion parallax, and camera rotation. On the other hand, with the proposal of generative adversarial network (GAN), it has shown strong generative ability when it is used in various image conversion tasks. Inspired by GAN, Kupyn et al., [8] used conditional adversarial networks combined with perceptual loss to perform blind motion deblurring. The adversarial network could quickly reconstruct a realistic generated image by learning the structure of the graph. At the same time, the adversarial network was constrained by perceptual loss to ensure the perceptual fidelity of the generated image.

On the other hand, the deformable convolution module further adjusts the position information of the spatial sampling to weaken the regular grid point sampling in the standard convolution structure, which makes it difficult for the network to adapt to the limitation of the geometric deformation of the object. The emergence of this kind of adaptive dynamic sampling provides a new guiding idea for solving the non-uniformity problem of motion blurred images. In addition, in view of the current problems that deepening the network model in space makes it difficult to train the model, the channel attention module adopts a brand-new feature recalibration strategy. Based on the application of this idea, ideal results have been obtained in image conversion tasks such as image super-resolution [9] and image denoising [10]. The research of the channel attention module lays the foundation for restoring the high-frequency details of the blurred image.

Based on previous work, this paper proposes a generative adversarial network based on multi-feature fusion strategy for motion image deblurring. Through the constructed adaptive residual module, the advantages of the deformable convolution module and the channel attention module are fully integrated. It can realize efficient blur image restoration.

## 3 Proposed Motion Image Deblurring

As shown in Fig. 1, based on the idea of adversarial games, the overall network structure consists of a generative network and a discriminant network. The generative network  $G$  is used to better reconstruct high-quality deblurred images. It consists of three parts: shallow feature extraction module ( $M_e$ ), adaptive residual module (ARM), and feature reconstruction module ( $M_r$ ). The input is motion blurred image  $B$ , the output is deblurred image  $\mathcal{Y}$ . The discriminant network  $D$  uses PatchGAN [11] to discriminate the learned deblurred image  $\mathcal{Y}$  and the clear image  $\hat{\mathcal{Y}}$ . Through the adversarial game between the generative network and the discrimination network, the deblurring process can be better constrained, and the quality of the deblurred image can be improved.

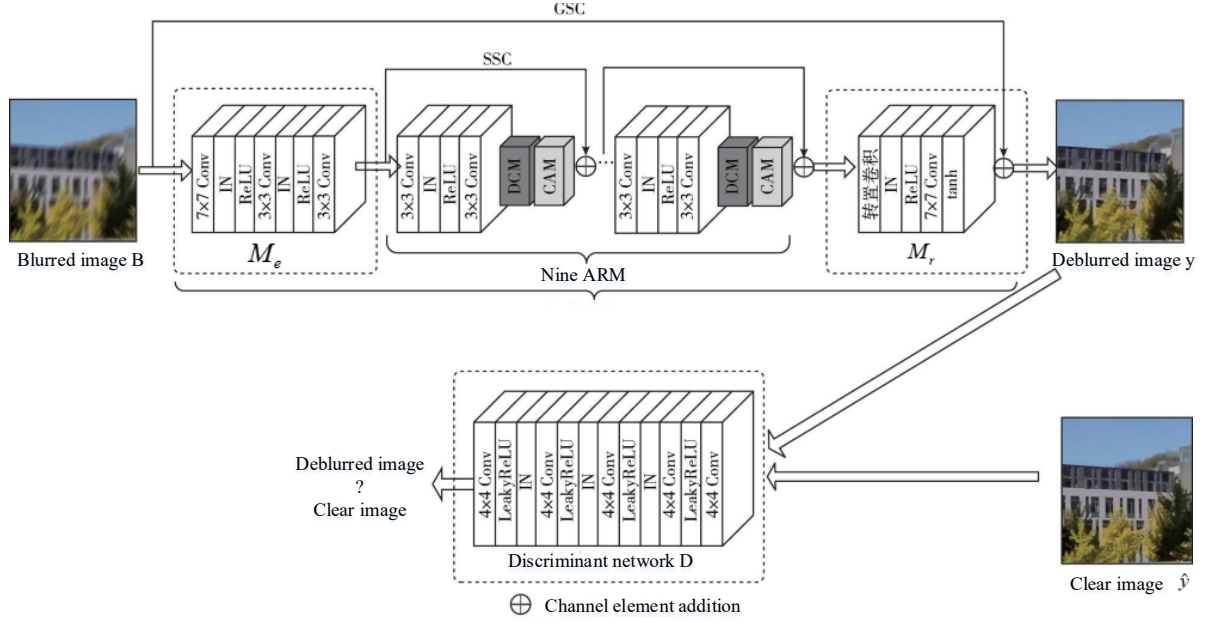


Fig. 1. Proposed network structure

### 3.1 Generative Network

As shown in Fig.1, the generative network  $G$  is divided into three parts: shallow feature extraction module ( $M_e$ ), adaptive residual module (ARM) and feature reconstruction module ( $M_r$ ).  $M_e$  consists of a  $7 \times 7$  convolutional layer and two  $3 \times 3$  convolutional layers. After each convolutional layer, instance normalization (IN) [12] and ReLU function are added. The instance normalization layer is responsible for normalizing the size of the channel features. Because for image deblurring by CNN, the generated result mainly depends on a clear image instance, and the independence between each image instance can be maintained, which can accelerate the convergence of model. The function  $M_e$  is to extract the shallow features of the input blurred image  $B$ , which can obtain the shallow feature  $P_0$ .

$$P_0 = M_e(B) \quad (1)$$

Next, it adds multiple ARMs after the shallow feature extraction module to perform deep feature learning on the shallow feature  $P_0$ .

$$P_r = P_l(P_{l-1}(\dots P_1(P_0)\dots)) \quad (2)$$

$P_l$  represents that it uses  $l$  ARMs for deep-level feature learning. This paper only uses nine ARMs to weigh the amount of model calculation and deblurring effect. ARM is composed of residual basic block, cascaded deformation convolution module and channel attention module. The deformation convolution module helps the network to better adapt to the geometric deformation of the graph when extracting features. The channel attention module enables the network to discriminately learn the features of each channel and improve the network performance. Secondly, the short skip connection (SSC) used in each ARM can promote the flow of low-frequency information in the image and help alleviate the problems of gradient disappearance and gradient explosion. Finally, the deep features  $P_r$  obtained after multiple ARMs adaptive capture features containing more high-frequency detail information. The feature reconstruction module  $M_r$  is composed of transposed convolution and  $7 \times 7$  convolution. After transposed convolution, instance normalization and ReLU are added, and tanh is used as the activation function. The feature reconstruction module can reconstruct the deep-level features  $P_r$  learned by the ARM, and combine it with the global skip connection (GSC) that can maintain the global structure to obtain the deblurred image  $y$ . The final deblurred image  $y$  generated by the generative network  $G$  can be expressed as:

$$y = G(B) = B + M_r(P_r) = B + M_r(P_i(P_{i-1}(\dots P_1(P_0)\dots))) \quad (3)$$

### 3.2 Discriminant Network

Adversarial training has been proven to effectively synthesize real images. In the adversarial network, the goal of generator G is to generate as real images as possible to deceive the discriminator. The purpose of discriminator D is to combine the real images with those generated by the generator. The images are distinguished. Through the adversarial training between the generator and the discriminator, the generation ability of the generative network is continuously improved. Finally, a deblurred image that is highly similar to the clear image can be restored. The discriminant network of this model uses PatchGAN. PatchGAN discriminates whether the image belongs to the generated deblurred image or the clear image on the image block with size  $N \times N$  in the image. In order to reduce the operating parameters of PatchGAN, we set  $N=64$ , which can quickly distinguish true and false. It can accurately restore important image information. The discriminant network shown in Fig. 1 contains five convolutional layers, and the size of the convolution kernel is set as  $4 \times 4$ . Except for the first and last convolutions, each convolutional layer is followed by instance Unified and nonlinear activation function LeakyReLU.

### 3.3 Adaptive Residual Module Design

The adaptive residual module constructed in the generative network is shown in Fig. 2, which is composed of the residual basic block cascaded deformation convolution module and the channel attention module. The adaptive residual module aims to learn the deformation information of motion blur image features to adapt to its non-uniformity and learn important high-frequency features to achieve blur image artifact removal and texture detail restoration.

The details are as follows:

(1) the input feature map is first extracted through two  $3 \times 3$  convolution layers and the residual basic block (composed of instance normalization and ReLU layer) to obtain the residual feature  $P_c$ .

(2) then  $P_c$  is sent to the deformation convolution module to obtain the deformation Features  $f_c$ .

(3) The function of the deformable convolution module is to dynamically adjust the convolution kernel according to the residual features that need to be identified currently, so that the sampling grid can be deformed freely to adapt to the non-uniformity of the motion blurred image.

(4) The subsequent channel attention module helps the model to assign different weights to each channel of the deformation features extracted by the deformation convolution module, so that the network learns each channel feature differently.

(5) Finally, it extracts more critical high-frequency features  $\hat{f}_c$ . The input feature and the feature  $\hat{f}_c$  obtained by the final screening are added through the short-hop connection, and they are input to the next basic unit in the network together.

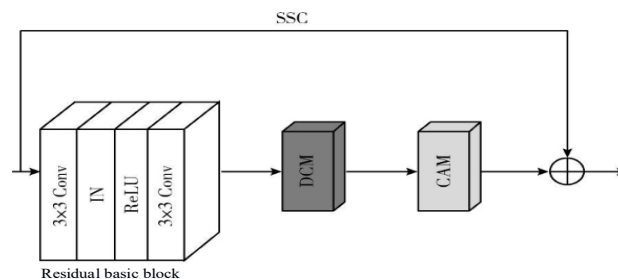


Fig. 2. Adaptive residual module

#### A. Deformation Convolution module

The input of the deformable convolution module (DCM) is the residual feature. The white dots of the input fea-

ture map shown in Fig. 3 indicate that the standard convolution structure uses a regular rectangular grid to sample the input feature map. When the traditional standard convolution structure is used for feature extraction of the input features, each position  $P_b$  of the acquired feature map  $f_s$  is expressed as follows:

$$f_s(P_b) = \sum_{P_n \in R} \omega(P_n) \bullet P_c(P_b + P_n). \quad (4)$$

$R$  represents the standard convolution structure, and  $P_n$  is the enumeration position of  $R$ . However, the traditional standard convolution structure does not consider the deformation information of the image, so in the deformation convolution module, it first uses the additional conventional convolution  $P_a$  to learn the offset  $\Delta P_n$  of the input feature, and this offset represents the deformation information of the image. And then it adds the offset  $\Delta P_n$  to the regular sampling position shown by the original white dot to obtain the new sampling position shown by the black dot in Fig. 3. The  $(\omega_1, \omega_2, \dots, \omega_9)$  represents a  $3 \times 3$  convolution kernel, which dynamically changes according to the new sampling position to adapt to the non-uniform fuzzy content. The final deformation feature  $f_c$  is shown in formula (5):

$$f_c(P_b) = \sum_{P_n \in R} \omega(P_n) \bullet P_c(P_b + P_n + \Delta P_n). \quad (5)$$

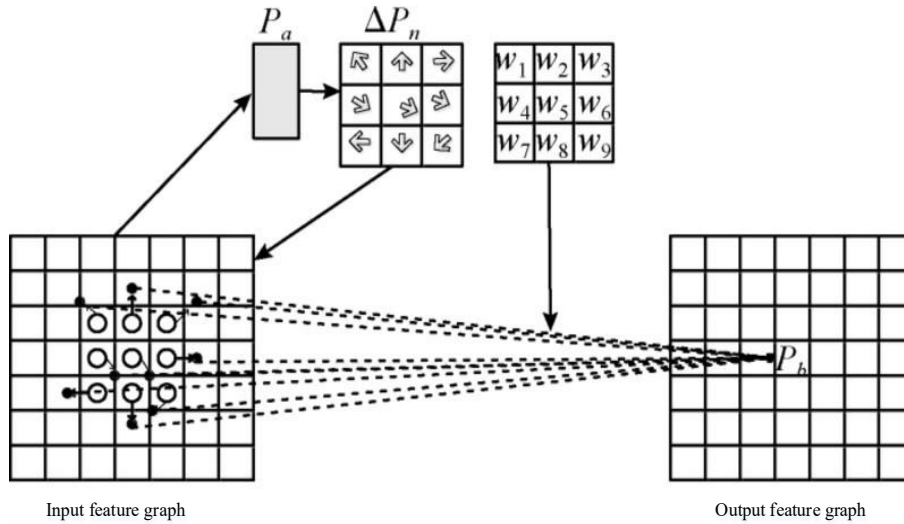


Fig. 3. DCM

## B. Channel Attention Module (CAM)

The channel attention module (CAM) structure is shown in Fig. 4. The input of the channel attention module comes from the deformation feature  $f_c$  extracted by the deformation convolution module in the previous stage. First, the global average pooling  $y_p$  is used to reduce the deformation feature  $f_c$  from  $H \times W \times C$  to  $1 \times 1 \times C$ . At this time,  $f_c$  becomes a real number  $y_c$ . This real number has a global receptive field, which represents the global distribution of the response on the feature channel, and the layer close to the input can also obtain the global receptive field, as shown in formula (6):

$$y_c = y_p(f_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W f_c(i, j). \quad (6)$$

In order to model the correlation between the channels of the image features, channel attention is introduced

into the gating mechanism  $F_{\text{ex}}$ , which is reflected by the sigmoid function and the ReLU function.

$$r_c = F_{\text{ex}}(y_c) - \beta(H_U \delta(H_D(y_c))). \quad (7)$$

In Eq.(7),  $\beta$  and  $\delta$  represent sigmoid function and ReLU function, respectively.  $H_D$  and  $H_U$  represent the dimensionality reduction and dimensionality increase weight set of the convolutional layer, respectively.  $y_c$  first performs channel reduction through  $H_D$  with a reduction ratio  $r=16$  to obtain a channel descriptor describing the features of each channel, the size is  $1 \times 1 \times \frac{C}{r}$ . After being activated by the ReLU function, the channel amplification layer with weight  $H_U$  is restored to the size of  $1 \times 1 \times C$ . Its range is scaled to between 0 and 1 through the sigmoid function. Finally, the weight set of channel features  $r_c$  is obtained, which represents the features of each channel after feature selection importance. It multiplies  $r_c$  and the input deformation feature  $f_c$  with channel elements to obtain the more important feature  $\hat{f}_c$  which is filtered by the channel attention module, its size is still  $H \times W \times C$ .

$$\hat{f}_c = r_c \bullet f_c. \quad (8)$$

In general, the deformation feature  $f_c$  input in the previous stage is adaptively rescaled and selected after being extracted by the channel attention module, which can make the network pay more attention to the channel containing high-frequency features and increase the suppression of redundant features. Thereby, it improves the network performance.

### 3.4 Multi-loss Fusion Optimization Method

This paper is based on the adversarial network structure. The objective function of adversarial training can be expressed as solving the following minimum-maximum problem:

$$\min_G \max_D \int_{y \sim Z_g} [D(y)] - \int_{\hat{y} \sim Z_r} [D(\hat{y})]. \quad (9)$$

As shown in equation (9), Wasserstein distance [13] is used as the criterion for discrimination,  $Z_g$  is the data distribution of the generated picture, and  $Z_r$  is the data distribution of the real data. The purpose of the generator G is to minimize the objective function as much as possible, and the discriminator D to maximize the objective function as much as possible. Its anti-loss reference draws on  $L_{GAN}$  in the reference [8], which can be shown as follows:

$$L_{GAN} = \int_{y \sim Z_g} [D(y)] - \int_{\hat{y} \sim Z_r} [D(\hat{y})]. \quad (10)$$

In order to further alleviate the gradient disappearance or gradient explosion during the training of the adversarial network, the gradient penalty [14] is used to constrain  $L_{GAN}$ :

$$L_{GAN} = \int_{y \sim Z_g} [D(y)] - \int_{\hat{y} \sim Z_r} [D(\hat{y})] + \lambda \int_{\bar{y} \sim Z_g} [(\|\nabla(\bar{y})\|_2 - 1)^2]. \quad (11)$$

$Z_{\bar{y}}$  means sampling in the generated image and the real image, and  $\nabla$  means the derivative operation. Based on the above analysis, in order to improve the stability of adversarial network training and image quality reconstruction, this paper combines the perceptual loss [15], pixel space loss [16] and the above adversarial loss to obtain a multi-loss objective function:

$$L = L_{pixel} + L_{GAN} + \lambda_1 L_p. \quad (12)$$

Pixel space loss  $L_{pixel}$  is used to maintain the similarity of the image content, and correctly guide the low-frequency information of the image captured by the network against the network:

$$L_{pixel} = \frac{1}{WH} \sum_{x=1}^W \sum_{y=1}^H \left( y - \hat{y} \right)^2. \quad (13)$$

At the same time, in order to maintain the perceptual fidelity of the image, the perceptual loss  $L_p$  is used to measure the high-level difference between the blurred image and the clear image, so that the content of the generated image and the clear image is close to the global structure, as shown in formula (14):

$$L_p = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} \left( \phi_{i,j}(y)_{x,y} - \phi_{i,j}(\hat{y})_{x,y} \right)^2. \quad (14)$$

$\phi_{i,j}$  represents the feature map obtained by the j-th convolution before the i-th maximum pooling layer in the pre-trained VGG-19 network. In this model, the feature map extracted by the third convolutional layer of the same pre-trained VGG-19 network as in the reference [8] is used to calculate the perceptual loss.

## 4 Experiments and Analysis

### 4.1 Experimental Data and Parameters

Traditional motion blurred image data sets mostly use fixed kernels to perform blur processing to obtain blurred images. It is difficult to simulate natural motion blurred images. In contrast, the GOPRO [17] dataset in this paper uses a GoPro Hero 4 camera [18] to capture a video sequence of 240 frames per second, and generates a blurred image by averaging consecutive short exposure frames. The motion blurred image formed in this way is more realistic. The GOPRO data set includes blur-clear image pairs collected from various motion scenes, and it simulates the non-uniform blur type formed by camera shake and object movement in natural scenes. It contains 2103 pairs of blur-clear image pairs for training and 1111 pairs of blur-clear image pairs for testing.

In the network model training, the experimental platform of this paper is configured with GTX1060Ti graphics card and Ubuntu 14.04 operating system. It is implemented by using the Pytorch deep learning framework. In this experiment,  $\lambda=10$  in Eq.(11), and the balance coefficient  $\lambda_1$  is set to 0.0001. In the training process, the initial learning rate of the generator and the discriminator is set to 0.0001, and the batch size is 1. Using Adam algorithm for stochastic gradient optimization, the total epoch=300. The first 150 epochs are learned at a learning rate of 0.0001, and the learning rate of the last 150 epochs gradually decays to zero.

### 4.2 Experiment Analysis

In order to demonstrate the effectiveness of the method in motion image deblurring, a large number of experiments have been carried out under the same conditions. The model is tested on the GOPRO test set. peak signal to noise ratio (PSNR) and structural similarity (SSIM) are used as evaluation indicators. The experiment gives the average PSNR and average SSIM values under the full data set. It is also compared with different deblurring

algorithms.

### A. Comparison of Deblurring Effects with Different Adaptive Residual Module Number

In order to explore the influence of different numbers of adaptive residual modules on the deblurring effect, experiments are carried out under different numbers of adaptive residual modules. In Table 1, ARM-3, ARM-6, and ARM-9 represent the 3, 6, and 9 adaptive residual modules respectively. It can be seen from the table that the PSNR of ARM-6 is 0.27 higher than that of ARM-3. Compared with ARM-6, the PSNR of ARM-9 is increased by 0.15, and the SSIM is increased by 0.01. In summary, with the increase of the ARM module number, the network can better adapt to the non-uniformity of motion blurred images. The more high-frequency information is extracted, the deblurring ability is gradually enhanced. However, considering the amount of model calculation, this algorithm aims to quickly verify the effectiveness of the constructed adaptive residual module in motion image deblurring. All the following experiments use nine adaptive residual modules.

**Table 1.** Performance analysis of different number of adaptive residual modules

Number of ARM	PSNR	SSIM
ARM-3	27.61	0.90
ARM-6	27.88	0.90
ARM-9	28.03	0.91

### B. Comparison with State-of-the-art Motion Image Deblurring Methods

The method in this paper is compared with the current four mainstream deblurring algorithms, such as Sun et al. [5], Xu et al. [18], Gang et al. [6], and Kupyn et al. [8]. As shown in Table 2, the deblurring algorithms such as Sun et al. [5], and Gang et al. [6] have poor performance. Compared with the previous three methods, the motion image deblurring method proposed in this paper has a significant improvement in both PSNR and SSIM values. Kupyn et al. [8] used an end-to-end adversarial network for motion image deblurring, which was currently a more efficient method for applying convolutional neural networks in the field of motion image deblurring. It can be seen from Table 2 that, although the new method in this paper is slightly lower than DeblurGAN in terms of SSIM, the PSNR is increased by 0.83. Based on the above analysis, it is verified that the adaptive residual module constructed by the deformation convolution module and the channel attention module are effective in deblurring motion images.

**Table 2.** Comparison with other algorithms

Data set	Algorithm	PSNR	SSIM
GPPRO	Sun et al.	24.68	0.85
	Xu et al.	25.18	0.89
	Gong et al.	27.19	0.90
	Kupyn et al.	27.20	0.95
	This paper	28.03	0.91

This paper compares the visualization of image details with the current advanced DeblurGAN anti-blurring algorithm, as shown in Fig. 5. From the flowerpot shown in Fig. 5(a), the part of the flowerpot generated in this paper has significantly fewer artifacts, the brightness is also close to the original image, and the image outline is clearer. In Fig. 5(b), the branch part is obviously more detailed than the texture of DeblurGAN. In Fig. 5(c), the facial features of the human face are more distinct, and they are more recognizable than the eyes, nose, and mouth of DeblurGAN. Based on the above analysis and experimental results, it can be seen that due to the instability of the adversarial network during training, the deblurred image still has serious artifacts and blurry texture details. The adaptive residual module constructed in the generation network in this paper can make the adversarial network better adapt to the non-uniformity of the motion blur image during training, it generates a deblurred image with richer texture, and effectively suppresses the blur of the image.



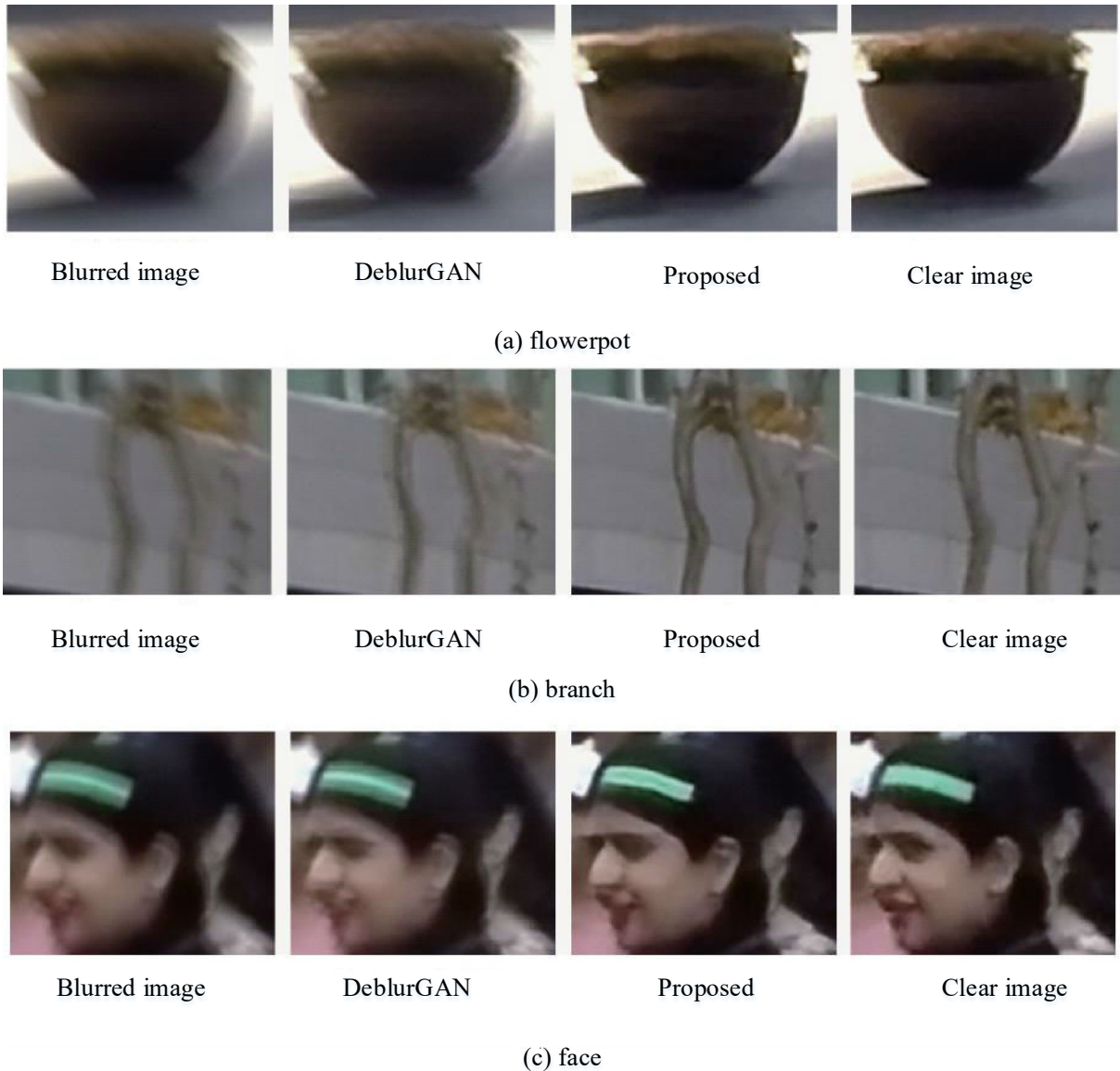


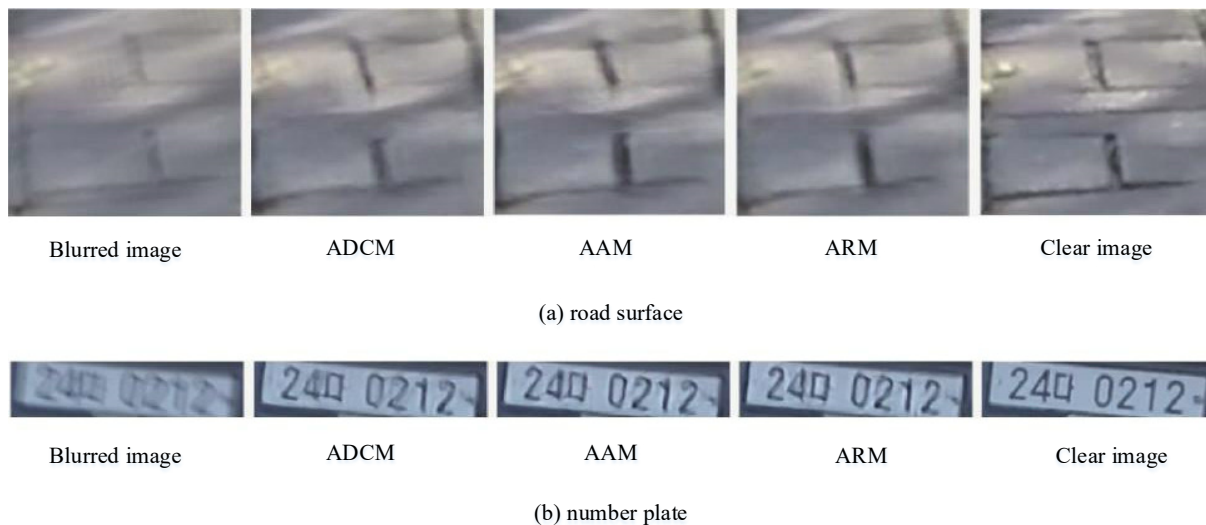
Fig. 5. Image detail comparison

### C. Performance analysis of deformable convolution module and channel attention module

In order to verify the effectiveness of the deformable convolution module and the channel attention module in image deblurring under the same experimental conditions, the adaptive deformation convolution module (ADCM) is composed of the residual basic block and the deformed convolution module is used. And the adaptive attention module (AAM) composed of residual basic block and channel attention module compares the experimental results, which are shown in Table 3 and Fig. 6.

**Table 3.** Performance comparison between deformable convolution module and channel attention module

Method	PSNR	SSIM
ADCM	28.00	0.91
AAM	28.01	0.90
ARM	28.03	0.91



**Fig. 6.** Deblurring effect comparison with different modules

As shown in Table 3, the PSNR of the adaptive deformation convolution module (ADCM) that only adds the deformable convolution module reaches 28.00 and the SSIM reaches 0.91. Compared with blurred images, the image quality is greatly improved. The adaptive attention module (AAM) has a PSNR of 28.01 and a SSIM of 0.90. The deblurring effect of the adaptive attention module is 0.01 lower than that of the adaptive deformable convolution module, but the PSNR is higher than 0.01. As shown in Fig. 6(a), for the cracks of the road surface, the AAM contains more details and the texture is clearer than the ADCM module. As shown in Fig. 6(b), the deblurring effect of AAM is clearer and more detailed than ADCM. However, the overall graphic structure in ADCM is closer to a clear image than AAM such as numbers 2 and 1. AAM has a significantly higher distortion than ADCM. This is because the deformation convolution module aims at the learning of image structure when promoting the network to adapt to the heterogeneity caused by image deformation, while the channel attention module focuses on the learning of channels containing high-frequency information, and focuses more attention on texture details. The experimental results show that both the deformable convolution module and the channel attention module can achieve better deblurring effect. When the deformable convolution module and the channel attention module are combined at the same time, the PSNR of the deblurring effect of ARM reaches 28.03, the SSIM value is 0.91, and the visualization effect shown in Fig. 6 is also better than ADCM and AAM, which can reconstruct details that are more in line with clear images. This further verifies that the adaptive residual module can fully integrate the advantages of the deformation convolution module and the channel attention module. The two complements work together, which can effectively improve the deblurring of motion images effect.

## 5 Conclusions

This paper proposes a motion image deblurring method based on adaptive residuals with the adversarial network. This method constructs an adaptive residual module composed of a deformation convolution module and a channel attention module in the generative network. Where, the deformation convolution module can make the network adapt to the non-uniformity of the blurred image and better extract the image features. The channel attention module uses the feature recalibration strategy to adjust the channels of the extracted deformation features to enhance the learning of high-frequency information. Experimental results show that the proposed algorithm can achieve better deblurring effect compared with the current advanced motion image deblurring algorithms. However, due to the instability of the training in the adversarial network, the restoration effect of the algorithm in this paper is not very good for some motion-blurred images with richer graphics structures. In the future, we will further explore a more stable confrontation network training mechanism to solve this problem.

## References

- [1] Y. Liu, W. Qin, K. Liu, F. Zhang, Z. Xiao, A Dual Convolution Network Using Dark Channel Prior for Image Smoke Classification, *IEEE Access* 7 (2019) 60697-60706.

- [2] H. Zhao, H. Yang, H. Su, S. Zheng, Natural Image Deblurring Based on Ringing Artifacts Removal via Knowledge-Driven Gradient Distribution Priors, *IEEE Access* 8 (2020) 129975-129991.
- [3] J.-F. Dai, H.-Z. Qi, Y.-W. Xiong, Y. Li, G.-D. Zhang, H. Hu, Y.-C. Wei, Deformable Convolutional Networks, in: *Proc. 2017 IEEE International Conference on Computer Vision (ICCV)*, 2017.
- [4] J. Hu, L. Shen, S. Albanie, G. Sun, E. Wu, Squeeze-and-Excitation Networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42(8)(2020) 2011-2023.
- [5] J. Sun, W.-F. Cao, Z.-B. Xu, J. Ponce, Learning a convolutional neural network for non-uniform motion blur removal, in: *Proc. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [6] D. Gong, J. Yang, L.-Q. Liu, Y.-N. Zhang, I. Reid, C.-H. Shen, A.-V.-D. Hengel, Q.-F. Shi, From Motion Blur to Motion Flow: A Deep Learning Solution for Removing Heterogeneous Motion Blur, in: *Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [7] M. Noroozi, P. Chandramouli, P. Favaro, Motion Deblurring in the Wild, In: *Pattern Recognition. GCPR 2017. Lecture Notes in Computer Science*, vol 10496. Springer, Cham.
- [8] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, J. Matas, Deblurgan: Blind motion deblurring using conditional adversarial networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [9] Y.-L. Zhang, K.-P. Li, K. Li, L.-C. Wang, B.-N. Zhong, Y. Fu, Image super resolution using very deep residual channel attention networks, in: *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.
- [10] S. Anwar, N. Barnes, Real Image Denoising With Feature Attention, in *Proc: 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.
- [11] C. Li, M. Wand, Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks, In: *Computer Vision- ECCV 2016. Lecture Notes in Computer Science*, vol 9907. Springer, Cham.
- [12] H.-J. Dai, C.-Y. Wu, T.-R. Tzong-Han, Collective Instance-Level Gene Normalization on the IGN Corpus, *Plos One* 8(11) (2013) e79517.
- [13] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein gan, *arXiv: 1701. 07875*, 2017.
- [14] M.-X. Hu, M. He, W. Su, A. Chehri, A TextCNN and WGAN-gp based deep learning frame for unpaired text style transfer in multimedia services, *Multimedia Systems* 21(2020) 1-10.
- [15] J. Johnson, A. Alahi, F.-L. Feir, Perceptual losses for realtime style transfer and super resolution, in: *Proc. European Conference on Computer Vision*, 2016.
- [16] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, A.-A. Efros, Context encoders: Feature learning by inpainting, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.
- [17] S. Nah, K.-T. Hyun, L.-K. Mu, Deep multir scale convolutional neural network for dynamic scene deblurring, In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
- [18] J.-F. Sun, Y.-T. Zhu, K. Wang, Motion deblurring based on DeblurG AN and lowrank decomposition, *Journal of South China University of Technology (Natural Science Edition)* 48(1)(2020) 32-41 (in Chinese).