Academic Journal of Computing & Information Science
ISSN 2616-5775 Vol. 3, Issue 4:102-109, DOI: 10.25236/AJCIS.030412

# Motion deblurring method based on Improved DeblurGAN

**Xinyan Zhang[1,*], Xiaofeng Wang[2]**

[1] *College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China*
[2] *Shanghai Maritime University, Shanghai 201306, Shanghai, China*
[*] *Corresponding author e-mail: 1954649331@qq.com*

***ABSTRACT.*** Generative Adversarial Networks(GANs) is a generation model that learns data distribution through the mutual against between generator network and discriminantor network.It has the advantage of generating clear and sharp samples, and has made progress in the application of image super-resolution and image repair. DeblurGAN solves the problem of end-to-end image deblurring by using conditional Generative Adversarial Networks(cGANs).In order to obtain better deblurring effect, this paper proposes an improvement based on DeblurGAN.Firstly, the method in this paper uses DenseBlock to replace the ResBlock in DeblurGAN, and adds two skip-connections. Finally, depthwise separable convolution is used to replace the common convolution block in the network, so as to reduce the network model, reduce parameters and accelerate the convergence speed of the network. The loss function uses the perceptual loss to ensure content consistency between the generated image and the clear image.

***KEYWORDS:*** motion blur, dense-connected-convolutional-network(DenseNet), skip-connection, depthwise separable convolution, perceptual loss

## 1. Introduction

In the process of object imaging,the relative motion,the camera not brought to focus,atmospheric turbulence effect and so on,which caused the appearace of blurred image.The blurred image is not good for extracting effective information of image,and it brings a lot of inconvenience to our life and scientific research,so it is of great practical significance to the research of the deblurring technology.Motion blur refers to the blur caused by the rapid motion of the object being photographed or the relative motion caused by the shaking of the imaging equipment, which is the most common blurring problem.Image deblurring is divided into non-blind deblur and blind deblur according to whether the blur kernel is known.In the early work of deblur,non-blind deblur was mainly studied.Assuming that the blur kernel is known,then the classical R-L filtering algorithm and wiener filtering algorithm [1] are used for deconvolution to obtain clear images.However,the blur image encountered in real life has unknow blur kernel,and the blur kernel is complex and diverse,difficult to estimate,which makes the problem of blind deblur more difficult.

In the traditional blind deblur method,the problem of image deblur is usually formulated as the estimation problem of blur kernel.First, the blur kernel is estimated using the prior information of the clear image,then convolution it with a clear image and add noise,can get a blur image,then the obtained blur image and blur kernel are used to calculate the estimated clear image. In the process of deblurring, the prior information of the image is used as a regular term to constrain the quality of the generated image so that the generated image approximates the original clear image as much as possible. In 2006,Fergus [2] was analyzed the gradient distribution of images by using the variable db Bayes method in mathematical statistics,a blind deblur model is formed by combining gaussian distribution and exponential distribution as constraints.The model deblurs very well.But the image characteristics of real blur images are not subject to Gaussian distribution and the model is not applicable to real blur images. In 2008,Fergus [3] use the image gradient obeys the feature of heavy tail distribution,and the super Laplace constraint term is used to quickly recover the information of blur kernel,but the image recovery quality is not high.The traditional blind deblur method relies too much on the image prior information, which makes it difficult to generalize to other models. Moreover, the estimation and computation cost of blur kernel is too high.

In recent years, with the development of deep learning technology, the deblur method based on neural network has gradually replaced the traditional method. Chakrabarti [4] was used convolutional neural network (CNN) to estimate the motion blur kernel,after the blur kernel is obtained, the non-blind deblur algorithm of traditional methods is used to deblur,although this method improves the accuracy of blur kernel estimation, it takes too long time to estimate the blur kernel process. In 2014,Generative Adversarial Networks(GANs) [5] was proposed, because of the simplicity of the model network and the advantages of the generation effect, GANs in the field of image translation and image super-resolution reconstruction has been rapidly applied and achieved good results,GANs are also began to be used in the field of deblur.Nah [6] used the combination of multi-scale CNN and GAN to realize the deblurring of dynamic scenes, and good deblurring effect is achieved. However, the model takes a long time. Inspired by research in the field of image translation [7], in 2018, Orest Kupyn [8] combining the conditional Generative Adversarial Network (cGAN) and perceptual loss based on recovered image content, DeblurGAN is constructed,it is with simple network structure and fast model processing speed, excellent results have been achieved in the field of deblur, which also greatly promotes the deblur based on cGAN.

Due to the edge blurring of DeblurGAN image, this paper modifies its network structure to improve the quality of image generation and speed up the network.The generator network adopts denseblock structure instead of residual block structure in the original structure,denseblock structure can make full use of network parameters, mitigation of gradient disappearance.At the same time, the skip connections are added between the convolutional layer of the network layer and the corresponding

ISSN 2616-5775 Vol. 3, Issue 4:102-109, DOI: 10.25236/AJCIS.030412

transposed convolutional layer to enhance the feature propagation,by the way, use the depthwise separable convolution instead of the common convolutional layer in the network, by reducing the size of the model and reducing the number of parameters and the amount of computation in the network.  The effectiveness of the proposed algorithm is finally verified by experiment.

## 2. Related work

### *2.1 Image blind deblur*

The most common formula for generating blur is as follows:

$$I_B = K(M) * I_S + N \tag{1}$$

$I_B$ is the generated blur image, $I_S$ is the sharp image, $K(M)$ is the PSF.It can be seen from this formula that the generation of blur image can be regarded as the result of the convolution operation of clear image with point diffusion function and the addition of noise.

### *2.2 Image deblur model based on cGAN*

Since the image deblur can also be seen as an image to generate another image conversion process, cGAN as a general scheme in the field of image translation, it must also be used in the field of image deblurring.The deblur model based on cGAN can learn the mapping relationship between blur image and sharp image, and its basic framework is shown in Figure 1.The blur image B with additive noise is generated by generator ,G(B).The sharp image S corresponding to the blur image B is input together with the generated image G(B) to the discriminator D, and the discriminator D outputs a probability value, namely the probability that the generated image G(B) is judged as the sharp image S. Then, the result of discriminator D is used to constrain the generation of generator G and guide the generator G to better simulate the data distribution of clear images, until the discriminator D cannot distinguish sharp image S and generated image G(B), then the network training can be considered to the optimal state, that is, to reach Nash equilibrium.

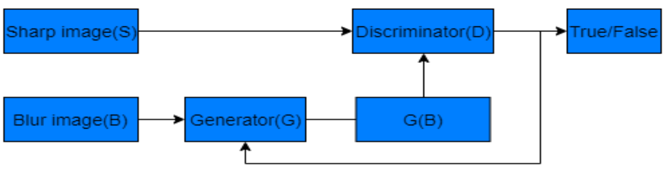


*Figure. 1 Image deblur model based on cGAN*

ISSN 2616-5775 Vol. 3, Issue 4:102-109, DOI: 10.25236/AJCIS.030412

## 3. Network structure

### *3.1 Genetator network structure*

The generation network structure of this paper is shown in Figure 2.In this paper, the common convolution is replaced by depthwise separable convolution.The network consists of two sets of 3 * 3 and 1 * 1 convolution blocks,one dense block containing 9 dense connection layers,and two sets of 3 * 1 and 1 * 3 transpose convolution blocks.Each dense connection layer consists of 3 * 3 convolution block and 1 * 1 convolution block.In addition, skip connections are added between the convolution block of the network and the corresponding transposed convolution block to obtain a better deblur effect.

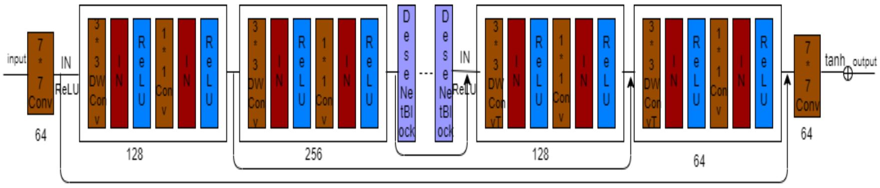


*Figure. 2 Generator network structure*

### *3.2 Discriminator network structure*

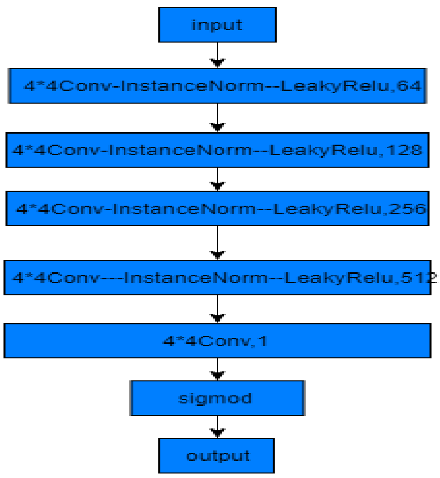


*Figure. 3 Discriminator network structure*

The discriminator network adopts the classic Markov discriminator (patchGAN), which is completely composed of convolutional layer, as shown in Figure 3. The

ISSN 2616-5775 Vol. 3, Issue 4:102-109, DOI: 10.25236/AJCIS.030412

advantage of the patchGAN is that the output is an N by N matrix,Each element in the matrix corresponds to the probability that each patch in the original figure is true. Finally, the mean value of elements in the matrix is taken as the final output of the discriminator.Experiments show that patchGAN is more suitable than ordinary GAN discriminator for areas with high image quality requirements.

### *3.3 Loss function*

The loss function of the model in this paper is composed of the perceptual loss and adversarial loss.The formula is as follows:

$$\ell = \ell_{\mathrm{adv}} + \lambda \ell_{content} \tag{2}$$

$\ell_{\mathrm{adv}}$ is the adversarial loss, focus on restoring texture detail of the image, $\ell_{content}$ is the perceptual loss, focus on restoring general content of the image.Experiments show that when $\lambda$ is 100, the result is the most best.

#### *3.3.1 Adversarial function*

GAN used JS divergence to measure the difference between the original data distribution and the generated data distribution,the smaller the JS divergence, the closer the two distributions get,But when there is no overlap between the two distributions,the JS divergence will be constant.Since there is no overlap between the two distributions, the probability is very high, it is easy to cause the gradient to disappear and difficult to train the network to converge.Arjovsky et al introduced the Wasserstein distance to effectively overcome the problem of gradient disappearance, then WGAN emerged [9]. Whether or not there is overlap between the two data distributions,Wasserstein distance can continuous transformations, therefore, the difference between the two distributions can be effectively measured.However, the W distance must satisfy the Lipscitz continuity condition, thus W distance WGAN-GP with gradient penalty mechanism is generated [10].The loss function is more robust to image deblurring,as shown in Formula (3), it is as follows.

$$W(P_r, P_g) \approx max\left\{E_{x\sim P_r}[D(x)] - E_{x\sim P_g}[D(x)] - \lambda E_{x\sim P_{penalty}}(||\Delta_x D(x)|| - 1)^2\right\} \tag{3}$$

$P_r$ is the image from the origin sharp image set, $P_g$ is the image from the generated image set, $D(x)$ is probability that the image is authenticated by the discriminator, $E$ is took the mean of the discriminator's results.

ISSN 2616-5775 Vol. 3, Issue 4:102-109, DOI: 10.25236/AJCIS.030412

### *3.3.2 Perceptual function*

In this paper, perceptual loss is used to ensure the authenticity of the generated image content.Perceptual loss is a simple L2 loss,euclidean distance is carried out between the feature map obtained by clear image convolution and the feature map obtained by generated image convolution,to make the image content and global structure closer,as shown in Formula (4), it is as follows.

$$L_X = \frac{1}{W_{i,j}H_{i,j}}\sum_{x=1}^{W_{i,j}}\sum_{y=1}^{H_{i,j}}(\varphi_{i,j}(I_S)_{x,y} - \varphi_{i,j}(G(I_B))_{x,y})^2 \tag{4}$$

$\phi_{i,j}$ is the feature map obtained by the j-th convolutional layer before the i-th maximum pooling layer in the VGG19 network [11], $W_{i,j}$ and $H_{i,j}$ are the width and height of the feature map, $I_S$ is the original sharp image, $G(I_B)$ is the Original sharp image,is the generated clear image.

## 4. Experimental results and comparison

In this experiment, GOPRO data set [12] was used for network training and verification.The data set is composed of blur-clear image pairs, including 2103 training sets images and 1111 test sets images. Since the GOPRO data sets images are all high-quality images of 1280*760, the size of the image needs to be cropped to the size of 256*256 required by the generator.

This article is under the Windows 10 system using pytorch deep learning framework for experiments, using Adam for the optimizer, the number of iterations is set to 400, the generator and the discriminator of the initial vector is set to 0.001, the convolution kernel size is 4 * 4, each round of iteration 4 for the number of images in the network, the whole experiment the effect of the best. In addition, because of the discriminator of learning faster than the slow learning speed of the generator, so the updated once a training the discriminator is 5 times the generator parameters.

The result evaluation index of image deblur can be divided into quantitative evaluation and qualitative evaluation.Quantitative evaluation is the objective evaluation of image.In this paper, two objective indicators, the peak-signal-to-noise ratio (PSNR) and the Structural Similarity index measure (SSIM) are used to evaluate the image quality before and after the deblurring.Qualitative evaluation is compared by visual identification, which is subjective to a certain extent.The PSNR and SSIM values as shown in table 1, can be seen from table 1 improved PSNR and SSIM are higher than before improvement, the result of the qualitative evaluation is through visual identification to evaluate, the results are shown in figure 4, (a) as the blurred image, (b) as the generated image, image contrast is recovering well known on the vision, the method has obvious advantages.

ISSN 2616-5775 Vol. 3, Issue 4:102-109, DOI: 10.25236/AJCIS.030412

*Table 1 Comparison of the results of training different models on our dataset*

| | DeblurGAN | Our method |
|---|---|---|
| PSNR | 26.38 | 27.19 |
| SSIM | 0.73 | 0.82 |

*Figure. 4 Experimental result. (a) is left, (b) is right*

## 5. Conclusion

In this paper, the method is improved on the basis of DeblurGAN. The ordinary convolutional layer is changed into depthwise separable convolution to make the network light. At the same time, the original residual block is replaced by the dense connection layer improved in this paper. This method comes from the dense connection network, which can make better use of the characteristic data in the network, so as to achieve better deblurring effect. Secondly also joined the jump connection, is advantageous to the characteristics of the cross layer, so as to improve the stability of the network. Model using GOPRO data sets, the experimental results show that the model in the PSNR and the SSIM has good effect , behind will continue to study the GANs, further improve to deblur.

## References

[1] Helstrom C W. Image Restoration by the Method of Least Squares [J]. J. opt. soc. amrer, 1967, 57 (3): 297-303.
[2] Fergus R, Singh B, Hertzmann A, et al. Removing Camera Shake from a Single Photograph [C] //ACM transactions on graphics (TOG). ACM, 2006, 25 (3): 787-794.
[3] Krishnan D, Fergus R. Fast Image Deconvolution using Hyper-Laplacian Priors [C] //Advances in neural information processing systems. 2009: 1033-1041.
[4] Chakrabarti. A neural approach to blind motion deblurring. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2016.
[5] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Networks. June 2014.

ISSN 2616-5775 Vol. 3, Issue 4:102-109, DOI: 10.25236/AJCIS.030412


[6] S. Nah, T. Hyun, K. Kyoung, and M. Lee. Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring. 2016.

[7] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. arxiv, 2016.

[8] KUPYN O, BUDZAN V, MYKHAILYCH M, et al DeblurGAN: blind motion deblurring using conditional adversarial networks [EB/OL].(2018-04-03) [2018-08-15]. https: //arxiv. org /pdf / 1711. 07064.pdf.

[9] Arjovsky M, Chintala S, Bottou L.Wasserstein Generative Adversarial Networks [C]// International Conference on Machine Learning, 2017: 214-223.

[10] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GANs [C]//Advances in Neural Information Processing Systems.2017: 5767-5777.


[10] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GANs [C]//Advances in Neural Information Processing Systems.2017: 5767-5777.


[11] SIMONYAN K, ZISSERMAN A. Very deep convolu- tional networks for large-scale image recognition [EB/OL]. (2015-04-10) [2018-08-17]. https://arxiv. org / pdf / 1409. 1556. pdf.

[12] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In CVPR, 2017.