Generative Adversarial Unsupervised Image Restoration in Hybrid Degradation Scenes

Fan Tang [a], Xinyu Zhu [a], Jinrong Hu [a,b], Juhong Tie [c], Jiliu Zhou [a,b], Ying Fu [a,b,*]

[a] School of Computer Science and Technology, Chengdu University of Information Technology, Chengdu, Sichuan, China
[b] Images and Spatial Information 2011 Collaborative Innovation Center of Sichuan Province, Chengdu, Sichuan, China
[c] School of Software Engineering, Chengdu University of Information Technology, Chengdu, Sichuan, China
[*] Corresponding Author: Ying Fu. Email: fuying@cuit.edu.cn

Abstract: In this paper, we propose an unsupervised blind restoration model for images in hybrid degradation scenes. The proposed model encodes the content information and degradation information of images and then uses the attention module to disentangle the two kinds of information. It can improve the ability of disentangled presentation learning for a generative adversarial network (GAN) to restore the images in hybrid degradation scenes, enhance the detailed features of restored image and remove the artifact combining the adversarial loss, cycle-consistency loss, and perception loss. The experimental results on the DIV2K dataset and medical images show that the proposed method outperforms existing unsupervised image restoration algorithms in terms of peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and subjective visual evaluation.

Keywords: unsupervised image restoration; hybrid degradation; generative adversarial network; attention mechanism; disentangled representation

1. Introduction

Image restoration is a typical under-determined problem in the field of image processing. Traditional methods use the degraded image as the known data and establish a corresponding mathematical model according to the image degradation mechanism, through which the original clear image can be restored [1,2]. Image prior information is the effective information of the image itself. Different image restoration methods proposed by researchers based on different prior information can estimate different original images. The prior information often used for image restoration includes local smoothness [3], non-local similarity [4], sparsity [5,6]. However, traditional methods heavily rely on prior knowledge of the image and the precise depiction of the degradation process, and the restoration effect is limited .

With the display of the effects of deep learning, more and more researchers have begun to use deep learning technology to solve related problems in the field of computer vision, and have achieved remarkable results [7-9]. Because of its powerful image feature learning ability, researchers in the field of image restoration also widely use this technology for research and method implementation in the field of image restoration [10,11].The currently proposed image restoration algorithms based on deep learning are mainly used for the restoration of degraded images caused by a single degradation mechanism, such as motion blur, rain, fog, noise, and super-resolution [12].Image blurring seriously affects the quality of the image itself, and will affect both image recognition and object recognition in the image. In order to obtain clear images, Kupyn et al. proposed a series of end-to-end deblurring methods based on conditional generation adversarial networks [13,14]. Nimisha proposed an unsupervised image deblurring method based on GAN, adding repetition loss and multi-scale gradient loss to the model [15]. Wang et al. presents a novel framework to deal with the non-uniform blur[16]. Most of the models that have been proposed for rain removal based on deep learning use fully supervised

methods [17-19]. However, the real rain map and the rainless map with the same background are difficult to obtain, so some researchers have proposed semi-supervised rain removal models [20] and unsupervised rain removal models [21]. Image defogging can eliminate the influence of haze environment on image quality and increase the visibility of images. Early image defogging methods based on deep learning were mainly based on atmospheric degradation models [22,23]. At present, researchers are more inclined to directly use the input foggy image output to obtain the defogging image [24,25,26]. Image denoising aims to restore a potentially relatively clean image from an image containing noise. DnCNNs[27], FFDNet[28] and CBDNet[29] proposed by Zhang et al. are very representative algorithms.

Due to the influence of multiple degradation factors, compared to a single degradation type, the image degradation of multiple degradation types is more serious, and the image restoration algorithm for a single degradation type mentioned above is not well applicable. In order to better restore images of mixed degraded types, the RL-Restore proposed by Yu et al. constructs a toolbox that contains small-scale convolutional networks for different restoration tasks, and selects appropriate according to different images to be restored The tool gradually restores damaged images [30]. The OWAN proposed by Suganuma et al. performs multiple basic operations in parallel in its core module, and selects the appropriate operation to restore the image according to the specific conditions of the image to be restored [31]. Bai et al. proposed an adaptive restoration algorithm based on hierarchical feature fusion. The algorithm directly fused the features of different receptive field branches to enhance the structure of the restored image [32].

The image restoration methods for mixed degraded images mentioned above require a large number of paired training samples. Paired training samples are difficult to obtain, and strict supervised training will cause the model to have overfitting problems and poor generalization performance. Therefore, this paper introduces a new unsupervised image restoration algorithm that can better solve the image restoration problem in complex scenes without the need for paired training samples compared to the current methods.

2. Disentanglement Method Based on Attention Mechanism

2.1 Framework of Proposed Model

The proposed framework consists of four parts: 1) content encoders E_{sc} and E_{bc} for clear image domain and degraded image domain affected by multiple degradation factors; 2) a degradation information encoder E_b ; 3) degraded and clear image discriminators D_b and D_s; and 4) degraded and clear image generators G_b and G_s. Given a training sample $simg \in S$ in the clear image domain and $bimg \in B$ in the degraded image domain, the content encoder E_{bc} extracts content information fc from clear samples, and E_b extracts degradation information fd from degraded samples. G_b then takes attention-based content information and degradation information to generate a degraded image $stbimg$, while G_s uses attention information to generate a clear image $btsimg$. The discriminators D_b and D_s distinguish between the real and generated examples. The architecture is illustrated in Fig. 1.

As shown in Fig. 1, the model first uses the encoder to perform a disentangled representation on the real image input. There are two kinds of encoders in the model, both mainly used for image feature extraction: content information encoder E_{bc} and degradation information encoder E_{sc}. E_{bc} contains three strided convolution layers and four residual blocks; E_{sc} contains four strided convolution layers and a fully connected layer, and they are used to extract content information from clear images and degraded images, respectively. Because the unpaired image data used in this model is easier to obtain, our method uses the strategy of sharing parameter weights of E_{bc} and E_{sc} to better extract effective content information from the degraded image for image restoration.
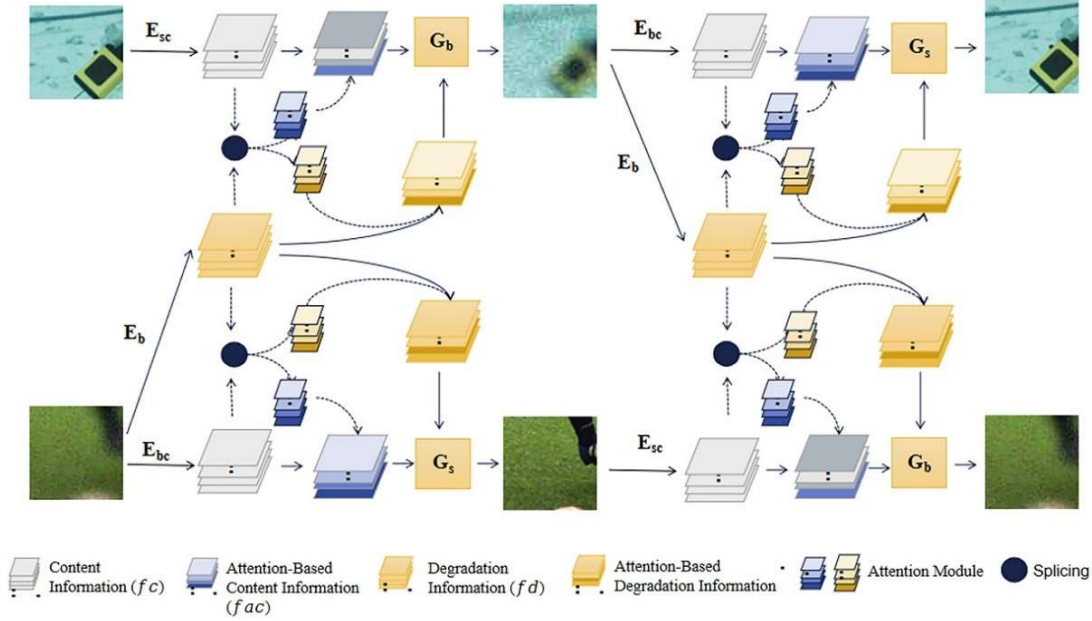
2

Fig. 1 Framework of proposed model

The degradation information encoder E_b should only encode degradation information from a degraded image $bimg$. However, in our experiments, we found that the generated degraded images $stbimg$ were also influenced by content information from $bimg$, such as color information. To help E_b suppress as much content information as much as possible, we add a KL divergence loss to regularize the distribution of degradation information $z_b = E_{bc}(bimg)$ to be close to the normal distribution $p(z) \sim N(0,1)$. The KL divergence loss is defined as:

$$KL(q(z_N) \| p(z)) = -\int q(z_N) \log \frac{p(z)}{q(z_N)} dz. \tag{1}$$

Minimizing the KL divergence is equivalent to minimizing the following loss:

$$L_{KL} = \frac{1}{2} \sum_{i=1}^{N} (\mu_i^2 + \sigma_i^2 - log(\sigma_i^2) - 1), \tag{2}$$

where μ and σ are the mean and standard deviation of z_b and N is the dimension of z_b. z_b is sampled as $z_b = \mu + z * \sigma$, and * represents element-wise multiplication.
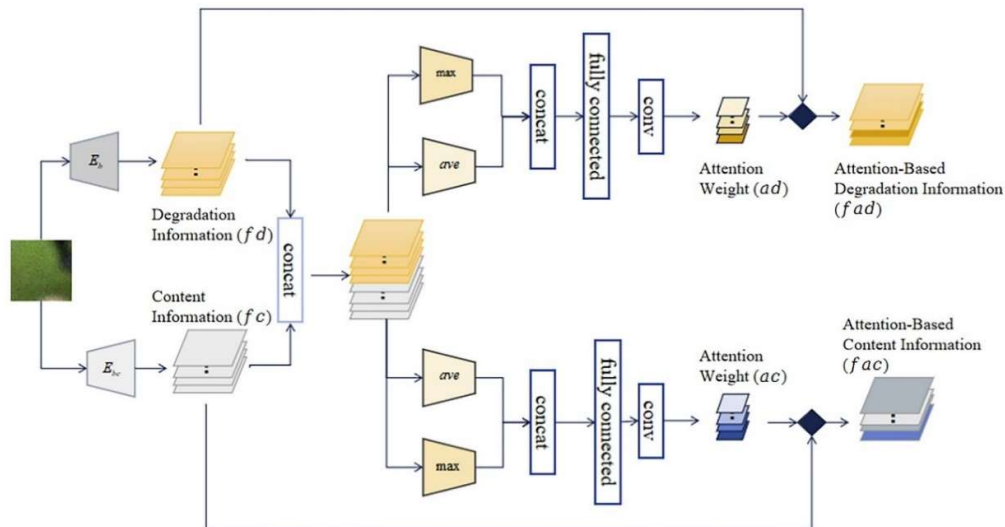

Fig. 2 Structure of attention-based disentangling operation

Second, due to the complexity of the degradation scenes of the images to be restored, there is the problem of inadequate disentanglement with only one disentangling operation. This model introduces a second disentangling operation of the attention mechanism on the content information fc and the degradation information fd encoded by the encoders to fully disentangle the image features and the degradation mechanism features, in order to improve the effectiveness of the encoded features. As shown in Fig. 2, motivated by Class Activation Mapping (CAM) [33], this model proposed in this paper first performs the splicing operation on the content information fc and the degradation information fd, and then inputs the spliced information into the upper and lower branches, respectively. We first perform pooling operations on the two branches. Pooling operations not only reduce the parameters of the model, but also keep the rotation, translation, and expansion characteristics of the features unchanged. Common pooling methods include average pooling and maximum pooling. Average pooling integrates the global spatial information of the features, while maximum pooling reduces the interference of useless information in the extracted features and extracts the features with the best response to the next module. In this paper, we use both pooling methods to calculate the attention weights, combine the advantages of both, and then feed the pooling results to the fully connected layer and the convolution layer to obtain the attention weights ac and ad of the degradation and content information. fc_i is the i-th layer of the content information fc, and ac_i is the attention weight of fc_i. The attention-based content information is then defined as:

$$fac = \{ac_i * fc_i | 1 \le i \le k\}, \tag{3}$$

where k is the number of layers of content information in the network.

Similarly, the attention-based degradation information is categorized as shown in formula (4):

$$fad = \{ad_i * fd_i | 1 \le i \le k\}. \tag{4}$$

After the disentangling operation is done, we feed the attention-based degradation information and content information that have been adequately disentangled into the generator to generate corresponding images. The generator is used to generate clear images and degraded images, and the discriminator is used to discriminate between the generated images and the real images, both of which gradually improve the effectiveness of the model in restoring images during continuous adversarial training. The generator contains four residual blocks and three transposed convolution layers; the discriminator consists of convolution layers and activation functions. To make the generated images look more realistic, we apply the adversarial loss on clear image domain S and degraded image domain B. For the clear image domain, we define the adversarial loss as formula (5):

$$L_{Dsimg} = E_{simg \sim p(simg)}[log \, D_S(simg)] + E_{bimg \sim p(bimg)}[log(1 - D_S(btsimg))]. \tag{5}$$

$btsimg$ is categorized as shown in formula (6):

$$btsimg = G_S(E_{bc}(bimg), z_b). \tag{6}$$

Similarly, we define the adversarial loss in the degraded image domain as L_{Dbimg}:

$$L_{Dbimg} = E_{bimg \sim p(bimg)}[log \, D_B(bimg)] + E_{simg \sim p(simg)}[log(1 - D_B(btsimg))]. \tag{7}$$

stbimg is categorized as shown in formula (8):

$$stbimg = G_B(E_{sc}(simg), z_b). \tag{8}$$

The discriminator mentioned above can discriminate between the images generated by the generator and the input real images, so the model can guarantee the authenticity of the input restored images. However, since the training data is unpaired, the images

4

restored by the model may appear inconsistent with the content information of the degraded images input. Motivated by CycleGAN [34], this paper solves this problem by adding cycle-consistency loss to the model. Specifically, the degraded image *stbimg* generated by the generator is re-transformed to the original clear image domain by using the attention module to perform the disentangling operation, as well as image generation of the generator, in order to obtain the generated clear image *ssimg* . Similarly, the generated degraded image *bbimg* can be obtained, and the specific process is defined as shown in formulas (9) and (10):

$$ssimg = G_S\big(E_{bc}(stbimg), E_b(stbimg)\big), \tag{9}$$

$$bbimg = G_B\big(E_{sc}(btsimg), E_b(btsimg)\big). \tag{10}$$

We define the cycle-consistency loss in both domains as formula (11):

$$L_{cyc} = E_{cimg \sim p(cimg)}[\|cimg - ccimg\|_1] + E_{bimg \sim p(bimg)}[\|bimg - bbimg\|_1]. \tag{11}$$

In addition, we use the content information and degradation information to transform the input clear images from the clear image domain to the degraded image domain. However, in the experiments, we find that the generated images often contain many artifacts. To solve this problem, we add perceptual loss to the model to further constrain the model, which is defined as:

$$L_p = \big\|f_{l-layer}(stbimg) - f_{l-lay}\ (bimg)\big\|_2^2, \tag{12}$$

where $f_{l-layer}(x)$ are the features of the *l*-th layer of the pre-trained convolutional neural network (CNN).

The full objective function is a weighted sum of all the losses from (5) to (12):

$$Loss = \lambda_{Dsbimg}\big(L_{Dsimg} + L_{Dbimg}\big) + \lambda_{KL}L_{KL} + \lambda_{cyc}L_{cyc} + \lambda_pL_p. \tag{13}$$

2.2 Parameter Settings of the Model

Table 1 Hyperparameter settings of proposed network

Parameters	Value Specified
Learning Rate γ	0.0002
Optimization Method	Adam
Batch Size	24
Iteration Times *N*	100
λ_{Dsbimg}	1
λ_{KL}	0.01
λ_{cyc}	10
λ_p	0.1

The hyperparameters of the model were set as shown in Table 1. The model was trained for a total of 100 iterations, and the model weight parameters were saved at the end of each training session. The learning rate was initially set to 0.0002, and we used the Adam algorithm to optimize our model. Considering the memory used by the model runtime, the model was trained by a small batch gradient descent algorithm with a batch size of 24. For each training of the model, the discriminator was first trained by simultaneously fixing the generator weight parameters and updating the discriminator weight parameters, and then the generator was trained by fixing the discriminator

parameters until the final training count was reached. For hyperparameters, we experimentally set: $\lambda_{Dsbimg} = 1$, $\lambda_{KL} = 0.01$, $\lambda_{cyc} = 10$ and $\lambda_p = 0.1$.

2.3 Algorithm

The training process of our model is shown in Table 2.

Table 2 Training process of proposed model

Input: the total number of iteration N, hyperparameters ($\lambda_{Dsbimg}, \lambda_{KL}, \lambda_{cyc}, \lambda_p$), optimization method *Adam*, learning rate $\gamma = 0.0002$

1.	Initialize the weights: w_{G_s} of the clear image generator G_s, w_{G_b} of the degraded image generator G_b, w_{D_s} of the clear image discriminator D_s and w_{D_b} of the degraded image discriminator D_b
2.	cycle 1 start for $n1 = 1:N$
3.	cycle 2 start for $i = 1:3$
4.	Sample a batch size of real clear images $simg_i$ and a batch size of real degraded images $bimg_i$
5.	The first disentangling operation (the content encoders E_{sc} and E_{bc} for clear and degraded image domains.) Given samples in two different image domains, the content encoder E_{sc} extracts content information $fc1_i$ from $simg_i$, then the content encoder E_{bc} and the degradation information encoder E_b extract content information $fc2_i$ and degradation information fd_i from $bimg_i$, respectively
6.	The second disentangling operation. Calculate the attention value of content information and degradation information. Obtain the attention-based content information $fac1_i$, $fac2_i$, and the attention-based degradation information fad_i according to formula (3) and formula (4), respectively
7.	Put $fac1_i$, $fac2_i$ and fad_i into the generator to obtain clear images $btsimg_i$ generated from G_s and degraded images $stbimg_i$ generated by G_b
8.	Execute the previous steps 5–7 to generate the reconstructed images $ssimg_i$ and $bbimg_i$
9.	Calculate the loss of discriminator
10.	Update parameters w_{D_s} and w_{D_b} by the gradient descent method
11.	cycle 2 end
12.	Execute steps 4–8 to calculate the loss of the generator
13.	Update weights of the encoder, the attention module and the generator by the gradient descent method
14.	cycle 1 end

Output: the updated weights w_{bc} and w_b of the degradation information encoder, w_G of the clear image generator, the updated attention weights w_{ac} and w_{ab}

3 Experimental Analysis

To demonstrate the effectiveness as well as the robustness of the proposed method, the restoration experiments were verified for images with a mixture of multiple degradation types as well as those with a single degradation type. The mixed degradation type images were from the DIV2K dataset, and the factors leading to image degradation

6

were a mix of adding Gaussian blur, Gaussian noise, and JPEG compression. The images to be tested were classified as mild, moderate, or severe, according to the degree of degradation. The single degradation type images were from the MRI dataset [35], and the factors that caused the image degradation were Rician noise of different intensities.

3.1 Experimental Data

3.1.1 DIV2K Dataset

The DIV2K dataset has 800 high-definition large-size images, which in this paper are divided into two parts, with the first 750 images as training data and the remaining 50 as test data. To facilitate training, these large images are cropped into small images for experiments. The 249,344 small images cropped from the first 750 are used as the training dataset, and the 3584 small images cropped from the last 50 are used as the test dataset.

To simulate actual complex degradation scenes, we applied multiple types of distortion to both the training and test data. Specifically, we added Gaussian blur, Gaussian noise, and JPEG compression to the images. The standard deviations of Gaussian blur and Gaussian noise were randomly generated from the ranges of [0, 5] and [0, 50], respectively, and the quality of JPEG compression was randomly selected from the range of [10, 100]. Only moderate images were used for training, and images of the three degradation degrees were used for testing.

3.1.2 MRI Dataset

The MRI dataset used in this paper was selected from the Brain Web database, which is derived from the McConnell Brain Imaging Centre of the Montreal Neurological Institute at McGill University, For the experiments we use the synthetic MRI data (T1w and PDw), which can be downloaded from http://brainweb.bic.mni.mcgill.ca/brainweb. For the experimental testing, 5%, 10%, 15%, 20%, 25%, and 30% Rician noise was added to the T1w images and PDw images. As shown in Fig. 3, the MRI images were cropped into small size images of 63*63.
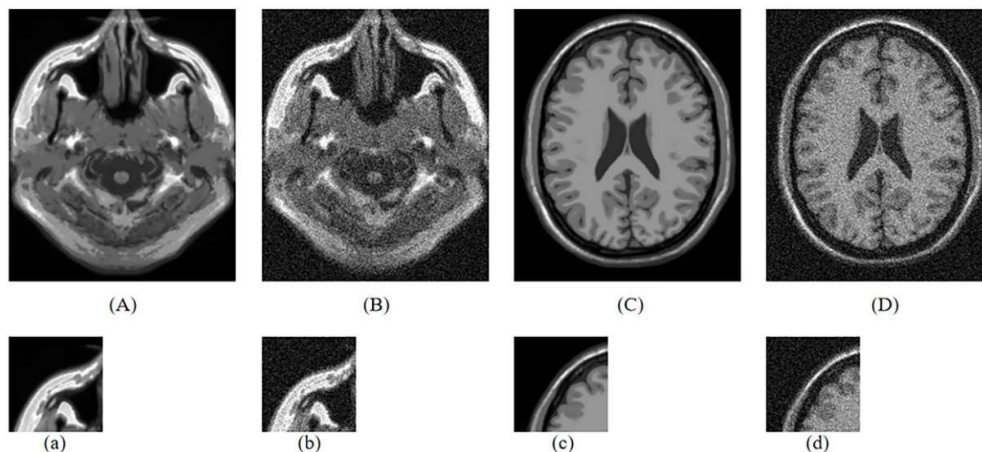

Fig. 3 Synthetic data used in our experiment. (A) and (C) are both clean MRI images. (B) and (D) are noise MRI images. (a), (b), (c), and (d) are taken from (A), (B), (C), and (D), respectively.

3.2 Experimental Environment

The experimental environment is shown in Table 3. Because the model parameters were large we required specific equipment for the experiments. We used a Linux server with an Nvidia RTX 2070 Super graphics card deployed and configured with a compliant acceleration platform and acceleration library.

Table 3 Experimental environment

Equipment	Configuration	Remark
Operating System	Ubuntu	16.04
Framework	PyTorch	1.0.1.post2
GPU	Nvidia RTX 2070 Super	8GB
Operating Platform	CUDA	11.1
GPU Accelerator Library	cuDNN	7.1.3
Language	Python3	3.6.6

3.3 Subjective Assessment of Effects of Image Restoration

3.3.1 Experiments for Images in Hybrid Degradation Scenes

We used the DIV2K dataset as a validation dataset for image restoration in complex degradation scenes. This part of the paper compares our method with the classical unsupervised image restoration algorithm CycleGAN and the newly proposed unsupervised image restoration algorithm UIDGAN [36]. The image restoration results of the DIV2K dataset are shown in Figs. 4, 5, and 6. In Fig. 4, we can see that our method can remove the noise from the image for the degradation degree of mild; the background color of the image restored by CycleGAN is not consistent with the clear image; and UIDGAN, which mainly targets domain-specific deblurring, cannot be used to restore the image in complex degradation scenes. Not only is the background color of the image restored by UIDGAN not consistent with that of the clear image, but also it does not remove the noise well. Fig. 5 shows the restoration effect of these methods on the moderate images. It can be seen that the image restored by our method is consistent with the content information of the clear image, and the texture information in the image can be restored, while the images restored by CycleGAN and UIDGAN lack texture information. The degree of damage to the severe image is the most serious, and Fig. 6 shows the results of the restoration. We can see that our method can still achieve relatively good restoration results, and in the case of the image background being influenced by degradation factors, the restored images of other methods show strange lines, which seriously affect the visual display of the recovered images.
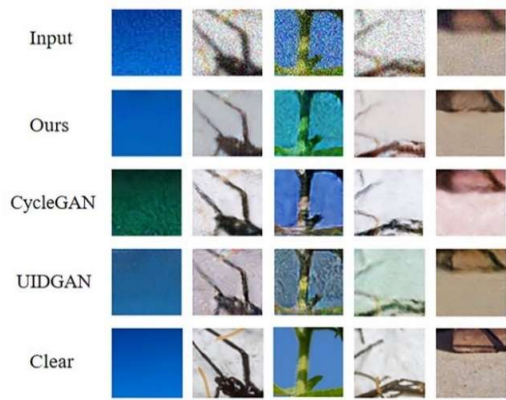

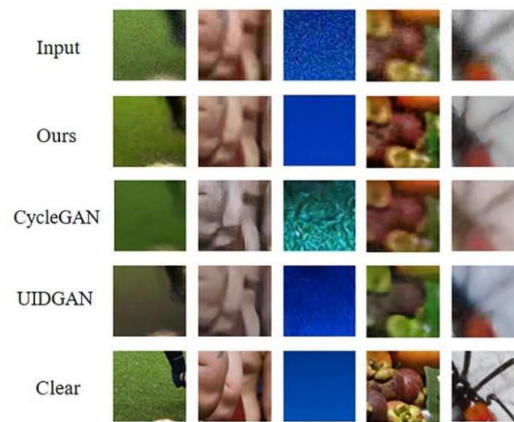**Fig. 4** Experimental results of mild images

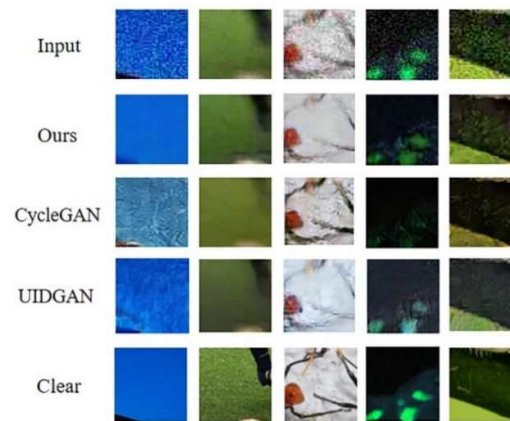**Fig. 5** Experimental results of moderate images

Fig. 6 Experimental results of severe images

3.3.2 Experiments for Images in Single Degradation Scene

To demonstrate the scalability as well as the robustness of the proposed method, we used the MRI dataset as a validation dataset for image restoration of a single degradation mechanism. This part of the paper mainly compares the proposed method with the traditional restoration method Anisotropic Diffusion Filtering (ADF) [37], the classical unsupervised image restoration algorithm CycleGAN based on deep learning, and the newly proposed unsupervised image restoration algorithm UIDGAN based on deep learning.

Figs. 7–12 show the T1w image restoration effect of each method. We can see that our method and CycleGAN have good adaptability and achieve good visual results at different Rician noise intensities. The images restored by the ADF lack details and obvious noise starts to appear in the images as the Rician noise intensity increases, which seriously impacts the visual effect. The image restored by UIDGAN showed other color artifacts at different noise intensities.

In addition to T1w images, we also conducted an experimental comparison of image restoration with PDw images, and Figs. 13–18 show the experimental results. Our method is relatively stable in the restoration effect at different noise intensities, but it is slightly deficient in the restoration of details. ADF also fails to sufficiently restore the details of the image, and the larger the noise intensity, the more noise residuals are found. The image restored by CycleGAN has a poor visual effect and very strange texture appears as the

noise intensity increases. The image restored by UIDGAN shows artifacts again at various noise intensities as PDw images.
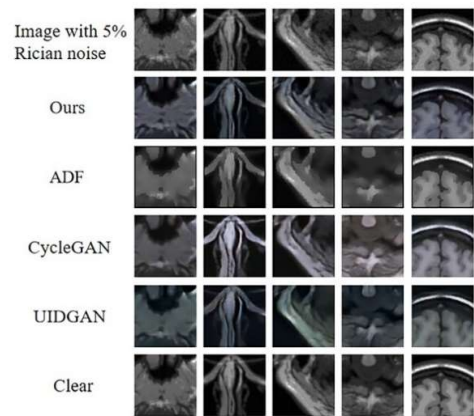

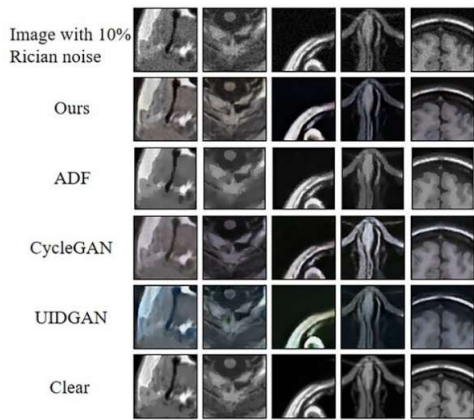**Fig. 7** Experimental results of T1w images with 5% Rician noise

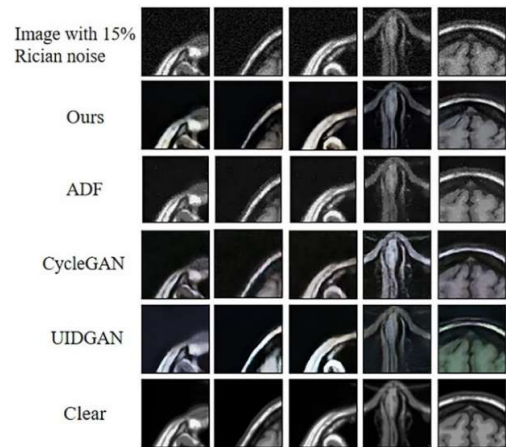**Fig. 8** Results of T1w images with 10% Rician noise

Fig.9 Results of T1w images with 15% Rician noise
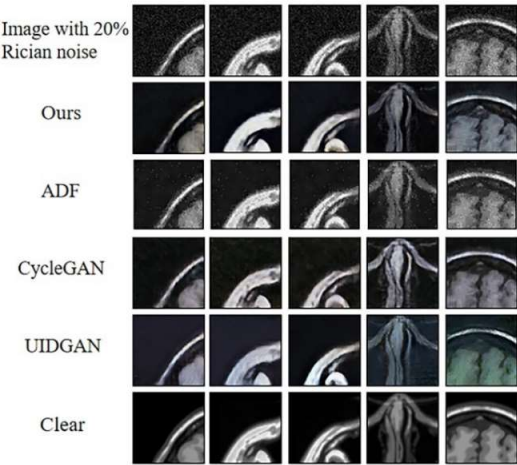
10

Fig.10 Results of T1w images with 20% Rician noise
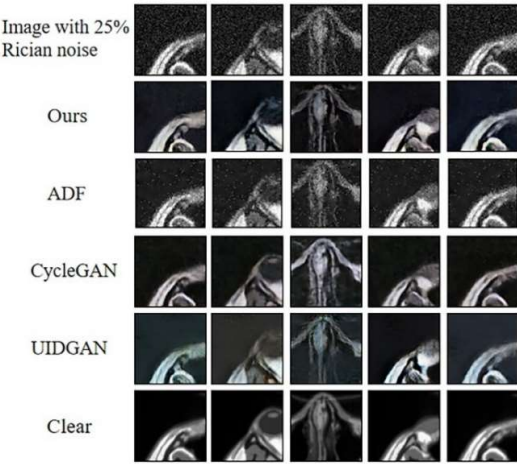

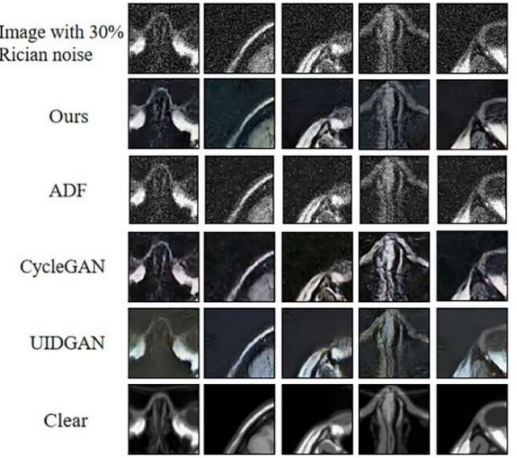**Fig.11** Results of T1w images with 25% Rician noise

Fig.12 Results of T1w images with 30% Rician noise
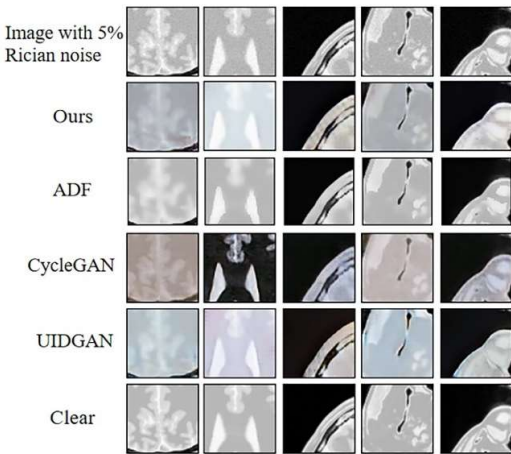
Fig.13 Results of PDw images with 5% Rician noise
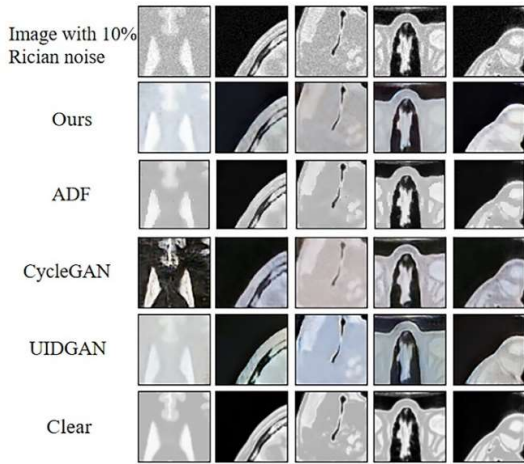

Fig.14 Results of PDw images with 10% Rician noise
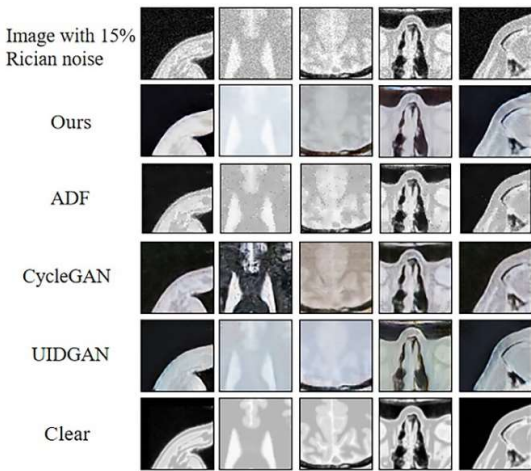

Fig.15 Results of PDw images with 15% Rician noise
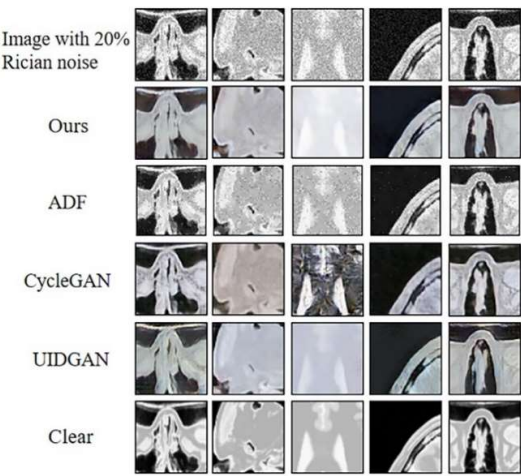
12

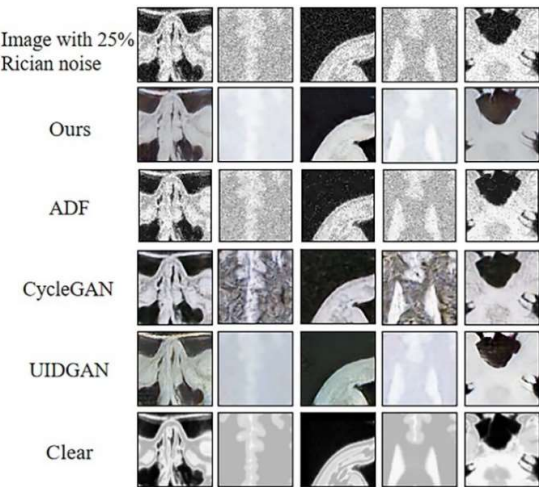**Fig.16** Results of PDw images with 20% Rician noise

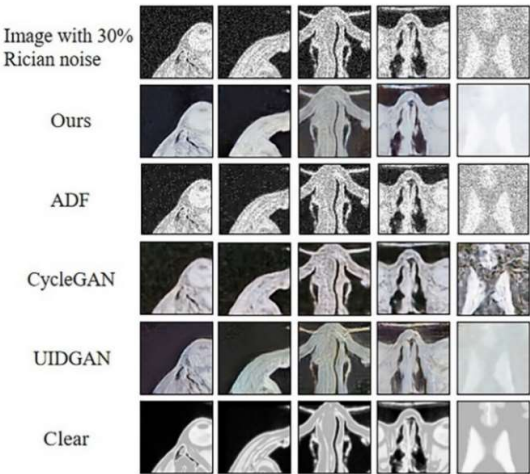**Fig.17** Results of PDw images with 25% Rician noise

Fig.18 Results of PDw images with 30% Rician noise

3.4 Objective Assessment of Effects of Image Restoration

We used structural similarity (SSIM) and peak signal-to-noise ratio (PSNR) as objective evaluation indices to evaluate the effect of image restoration.

3.4.1 Comparison of Quantitative Results of Image Restoration in Hybrid Degradation Scenes

Tables 4 and 5 show the PNSR values and SSIM values obtained by each method for images at different degradation levels. It can be seen that our method achieves the best quantitative results regardless of whether the image quality is mild, moderate, or severe, and also achieves a certain improvement compared with the recently proposed unsupervised image restoration algorithms. It can be seen that the effect of the classical unsupervised image restoration algorithm is poor, and in the comparison of SSIM values, there is a gap between our method and UIDGAN.

Table 4 PSNR values of each method for images at different degradation levels

Method	Mild	Moderate	Severe
CycleGAN	17.0965	16.1240	15.5669
UIDGAN	19.7100	18.9580	18.9565
Ours	**21.0428**	**20.3528**	**19.9386**

Table 5 SSIM values of each method for images at different degradation levels

Method	Mild	Moderate	Severe
CycleGAN	0.4876	0.3983	0.3095
UIDGAN	0.5757	0.4994	0.4437
Ours	**0.6002**	**0.5254**	**0.4618**

3.4.2 Image Restoration in a Single Degradation Scene

It can be seen from Tables 6 and 7 that the best PNSR values are achieved by our method on T1w images with the addition of different intensities of noise. For the SSIM value, our method also achieves the best value in the interval of noise intensity from 10% to 25%. The SSIM values of CycleGAN are optimal at 5% and 30% of Rician noise intensity, but the difference is not significant compared to the proposed method. The quantitative results on PDw images with the addition of different intensities of Rician noise are shown in Tables 8 and 9. The best PSNR values are obtained by our method when the noise intensity is less than 15%. When the noise intensity gradually increases, CycleGAN has the best PSNR value. ADF achieves the best SSIM value when the noise intensity is not greater than 10%, followed by CycleGAN.

Table 6 Average PSNR results on T1w images

Method	5%	10%	15%	20%	25%	30%
ADF	20.6106	20.3751	19.8617	18.9694	17.9012	16.8302
CycleGAN	20.0163	20.1530	20.2088	20.1578	20.0792	19.7676
UIDGAN	22.5134	22.4281	22.3678	21.4129	20.4036	19.3042
Ours	**24.9808**	**24.5294**	**24.1480**	**23.4785**	**22.2616**	**20.7901**

Table 7 Average SSIM results on T1w images

Method	5%	10%	15%	20%	25%	30%
ADF	0.7497	0.7125	0.6368	0.5044	0.3906	0.3104
CycleGAN	**0.7606**	0.7339	0.6901	0.6397	0.5897	**0.5404**
UIDGAN	0.7602	0.7259	0.6768	0.6191	0.5569	0.4986
Ours	0.7585	**0.7689**	**0.7208**	**0.6631**	**0.6004**	0.5377

Table 8 Average PSNR results on PDw images

Method	5%	10%	15%	20%	25%	30%
ADF	15.7612	15.7056	15.5447	15.2093	14.7566	14.2606
CycleGAN	20.1043	20.1353	20.1436	**20.1450**	**20.1828**	**20.1941**
UIDGAN	21.2514	20.9109	20.4918	20.0560	19.6652	19.1893
Ours	**21.4081**	**21.0796**	**20.6209**	20.1245	19.6098	19.1847

Table 9 Average SSIM results on PDw images

Method	5%	10%	15%	20%	25%	30%
ADF	**0.7807**	**0.7548**	0.6819	0.5378	0.4159	0.3358
CycleGAN	0.7249	0.7094	**0.6855**	**0.6581**	**0.6302**	**0.6027**
UIDGAN	0.6588	0.6457	0.6296	0.6120	0.5940	0.5721
Ours	0.6797	0.6652	0.6470	0.6250	0.6024	0.5795

4 Conclusion

In this paper, we propose a disentanglement algorithm based on the attention mechanism for solving the problem of image restoration in complex degradation scenes. Our model first performs a disentangling operation on the degraded image, using encoders to encode content information and degradation information of images, and applies a regularization constraint to the encoded degraded information, while making content information encoders for clear and degraded image domains share parameter weights. Secondly, due to the complexity of the image degradation scene and unpaired training data, in order to fully disentangle the image features and degradation mechanism features, we integrate the attention mechanism into the model for a second disentangling operation, which improves the effectiveness of the encoded features. Finally, the model combines adversarial loss, cycle-consistency loss, and perceptual loss to improve the image restoration effect under unsupervised conditions. The experimental analysis in this paper shows that our method can restore images in complex degradation scenes, and that our method not only achieves good visual effects but also achieves the most optimal quantitative results of the compared methods.

References

1. D. Krishnan, R. Fergus. Fast image deconvolution using hyper-Laplacian priors, Advances in Neural Information Processing Systems, 2009, 22: 1033-1041.
2. A. Buades, B. Coll, J. M. Morel. Non-local means denoising, Image Processing on Line, 2011, 1: 208-212.
3. Xu L., Zheng S., Jia J. Unnatural L0 sparse representation for natural image deblurring, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2013, 1107–1114. https://doi.org/10.1109/CVPR.2013.147.
4. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering, IEEE Transactions on Image Processing, 2007, 16(8): 2080-2095.
5. W. Dong, L. Zhang, G. Shi, X. Wu. Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization, IEEE Transactions on Image Processing, 2011, 20(7): 1838-1857.
6. J. Pan, Z. Hu, Z. Su, M. H. Yang. Deblurring face images with exemplars, European Conference on Computer Vision, Springer, Cham, 2014, 47-62.
7. A. Krizhevsky, I. Sutskever, G. E. Hinton. Imagenet classification with deep convolutional neural networks, Advances in Neural Information Processing Systems, 2012, 25: 1097-1105.
8. Y. LeCun, L. Bottou, Y. Bengio, P. Haffner. Gradient-based learning applied to document recognition, Proceedings of the IEEE, 1998, 86(11): 2278-2324.
9. Y. Zeng, Y. Gong, J. Zhang. Feature Learning and Patch Matching for Diverse Image Inpainting. Pattern Recognition, Volume 119, 108036, 2021.
10. Y. Tai, J. Yang, X. Liu, C. Xu. Memnet: A persistent memory network for image restoration, Proceedings of the IEEE International Conference on Computer Vision, 2017: 4549-4557.
11. J. Xie, L. Xu, E. Chen. Image denoising and inpainting with deep neural networks, Advances in Neural Information Processing Systems, 2012, 1: 341-349.
12. Y. Yang, Y. Qi. Image super-resolution via channel attention and spatial graph convolutional network, Pattern Recognition, Volume 112, 107798, 2021.
13. O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, J. Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, 8183-8192.
14. O. Kupyn, T. Martyniuk, J. Wu, Z. Wang. DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better, ICCV, 2019, 8878–8887.
15. N. T. Madam, S. Kumar, A. N. Rajagopalan. Unsupervised class-specific deblurring, Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2018, 11214 LNCS, 358–374. P. Wang, W. Sun, Q. Yan, A. Niu, Li R., Y. Zhu, J. Sun, Y. Zhang. Non-uniform motion deblurring with blurry component divided guidance, Pattern Recognition, Volume 120, 108082, 2021.
16. X. Fu, J. Huang, X. Ding, Y. Liao, J. Paisley. Clearing the skies: A deep network architecture for single-image rain removal, IEEE Transactions on Image Processing, 2017, 26(6), 2944–2956. https://doi.org/10.1109/TIP.2017.2691802.
17. W. Yang, R. T. Tan, J. Feng, J. Liu, S. Yan, Z. Guo. Deep Joint Rain Detection and Removal from a Single Image, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 1–1.
18. Li G, Xiang H, Wei Z, et al. Non-locally Enhanced Encoder-Decoder Network for Single Image De-raining[C]// 2018 ACM Multimedia Conference. ACM, 2018.
19. Qian R, Tan R T, Yang W, et al. Attentive generative adversarial network for raindrop removal from a single image[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 2482-2491.
20. Wei W, Meng D, Zhao Q, et al. Semi-supervised transfer learning for image rain removal[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 3877-3886.
21. Guo Y, Ma Z, Song Z, et al. Cycle-Derain: Enhanced CycleGAN for Single Image Deraining[C]//International Conference on Big Data and Security. Springer, Singapore, 2020: 497-509
22. Cai B, Xu X, Jia K, et al. Dehazenet: An end-to-end system for single image haze removal[J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.
23. Zhang H, Patel V M. Densely connected pyramid dehazing network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 3194-3203.
24. B. Li, X. Peng, Z. Wang, J. Xu, D. Feng. AOD-Net: All-in-One Dehazing Network, Proceedings of the IEEE International Conference on Computer Vision, 2017, 4770–4778.
25. Ren W, Ma L, Zhang J, et al. Gated fusion network for single image dehazing[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 3253-3261.
26. D. Chen, M. He, Q. Fan, J. Liao, L. Zhang, D. Hou, L. Yuan, G. Hua. Gated context aggregation network for image dehazing and deraining. Proceedings - 2019 IEEE Winter Conference on Applications of Computer Vision, WACV 2019, 2019, 1375–1383. https://doi.org/10.1109/WACV.2019.00151.
27. K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising, IEEE Transactions on Image Processing, 2017, 26(7), 3142–3155. https://doi.org/10.1109/TIP.2017.2662206.
28. K. Zhang, W. Zuo, L. Zhang. FFDNet: Toward a fast and flexible solution for CNN-Based image denoising, IEEE Transactions on Image Processing, 2018, 27(9), 4608–4622. https://doi.org/10.1109/TIP.2018.2839891.

16

29. S. Guo, Z. Yan, K. Zhang, W. Zuo, L. Zhang. Toward convolutional blind denoising of real photographs, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019, 1712–1722. https://doi.org/10.1109/CVPR.2019.00181.

30. Y. Yang, Y. Qi. Image super-resolution via channel attention and spatial graph convolutional network, Pattern Recognition, K. Yu, C. Dong, L. Lin, C. C. Loy. Crafting a toolchain for image restoration by deep reinforcement learning, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, 2443–2452. https://doi.org/10.1109/CVPR.2018.00259.

31. M. Suganuma, X. Liu, T. Okatani. Attention-based Adaptive Selection of Operations for Image Restoration in the Presence of Unknown Combined Distortions, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019, 9031–9040. https://doi.org/10.1109/CVPR.2019.00925.

32. L. Bai,H. Liu,Z. H. Shang. Mixed Degraded Image Restoration Algorithm Based on Adaptive Fusion of Hierarchical Features[J],Journal of Computer-Aided Design & Computer Graphics,33(2):8.

33. B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba. Learning deep features for discriminative localization, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 2921-2929.

34. J. Y. Zhu, T. Park, P. Isola, A.A. Efos. Unpaired image-to-image translation using cycle-consistent adversarial networks, Proceedings of the IEEE International Conference on Computer Vision, 2017, 2223-2232.

35. R.K.S. Kwan, A.C. Evans, G.B. Pike. MRI simulation-based evaluation of image-processing and classification methods, IEEE Transactions on Medical Imaging. 18(11):1085-97, Nov 1999. https://brainweb.bic.mni.mcgill.ca/brainweb/.

36. B. Lu, J. C. Chen, R. Chellappa. UID-GAN: Unsupervised image deblurring via disentangled representations, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019,10225-10234.4.

37. G. Gerig, O. Kubler, R. Kikinis, F. A. Jolesz. Nonlinear anisotropic filtering of MRI data, IEEE Transactions on Medical Imaging, 1992, 11(2): 221-232. https://doi.org/10.1109/42.141646.