A new blind image denoising method based on asymmetric generative adversarial network

Yiming Wang | Dongxia Chang | Yao Zhao

Abstract
Image denoising is a classical topic in computer vision. In recent years, with the development of deep learning, image denoising methods based on discriminative learning have received more attention. In this paper, a new blind image denoising method based on the asymmetric generative adversarial network (ID-AGAN) is proposed. In the new method, the adversarial learning is used to optimise the high-dimensional image information denoising, so as to balance the noise removal and detail retention. In order to overcome the unstability of the GAN training and improve the discriminative ability of the discriminating model, an image downsampling layer is added between the generating model and the discriminating model. Moreover, a multi-scale feature downsampling layer is utilised to extract the feature of the entire image and reducing the effect of noise on training images. Extensive experiments are conducted to verify the performance of the ID-AGAN algorithm. The results demonstrate that authors’ method has high performance and flexibility.

1 INTRODUCTION

Image denoising is a classical topic in computer vision and has been studied extensively. In real life, noise corruption is inevitable in the image sensing process and it may heavily degrade the quality of the acquired image. Therefore it is impossible to observe the image content clearly and the noisy image may effect on the high-level vision tasks. Accordingly, it is an essential step to remove the noise from the observed image in various image processing and computer vision tasks [1, 2].

In general, image denoising is a fundamental problem in the field of computer vision and image processing and its goal is to estimate the original clean image from a noisy image. The noisy image \( y \) can be expressed as \( y = x + n \), where \( x \) is the clean image corresponding to \( y \) and \( n \) is the noise. One common assumption is that \( n \) is an additive white Gaussian noise (AWGN). In the past decades, a large amount of denoising methods were proposed to recover the clean image from the noisy observation. And a gorgeous image denoising algorithm is supposed to have the following features: (1) It is flexible enough to recover the clean images with unknown noise. In other words, different noisy images can be recovered with a single model. (2) It is efficient and should not take too long to recover noisy images. (3) It is necessary to preserve the original image information while denoising, especially the high-dimensional image texture information that is easily mistaken as noise.

In the past few decades, many methods have been proposed for image denoising, including image prior based denoising methods [3–9] and discriminative learning based image denoising methods [10–14]. And image prior based denoising methods model image prior in various models, including non-local self-similarity (NSS) models [3, 15, 16], sparse models [4, 6, 17] and low rank models [7, 8, 18–20]. Image prior based denoising methods rely on image prior information, they are flexible and can be used with unknown noise. One of the classic denoising methods is BM3D [3], which is a benchmark in image denoising. For BM3D, it usually searches the similar blocks in the image and groups them firstly. After that, the similar blocks are denoised using the cooperative filtering. Finally, these blocks are fused into the original image with different weights. BM3D is faster than most image prior based methods, but its performance will deteriorate when the noise is complex. Sparse models based method KSDA [4] is based on the sparse and redundant representations over trained dictionaries. It trains a dictionary which is used to describe the image. Generally, the dictionary is supposed to be redundant so that it can...
represent different image structures. However, sparse coding with an over-complete dictionary is unstable in image restoration. NCSR method learns dictionaries from clustered image patches. It leads to a more stable and sparser representation. Consequently it can achieve better image restoration results. But it also brings greater computational complexity. In many cases, low rank based method WNNM [7] can achieve the best performance compared with other methods based on image prior. In [7], the singular values in the weighted nuclear norm minimisation (WNNM) problem are assigned different weights. And the solutions of the WNNM problem are analysed under different weighting conditions to apply it to image denoising. However, most of the existing image prior based denoising method are developed for gray scale images including WNNM. And it is hard to extend them for colour image denoising because the noise statistics in R, G, and B channels can be various for real noisy images. In [8], a multi-channel optimisation model for real colour image denoising under the weighted nuclear norm minimisation (WNNM) framework is proposed. MC-WNNM can handle noisy image with different noise statistics in the three channels. Despite their high denoising performance, most of the image prior based methods typically have two major drawbacks. First of all, these methods include complex modeling processes, which make testing time-consuming. Secondly, the image priors adopted are defined mostly based on human knowledge on image. Nevertheless, human understanding on image is incomplete, which may lead to weakness in processing complex noisy image.

As an alternative, discriminative learning based image denoising methods aim to learn the potential structure of image and discover the difference between noisy image and clean image from training. With the development of deep learning, the capacity of neural networks has been greatly exploited. Accordingly discriminative learning based denoising methods show striking denoising performance. In particular, a trainable non-linear reaction diffusion (TNRD) method was proposed which based on learning optimal nonlinear reaction diffusion models [10]. The TNRD approach is applicable for a variety of image restoration tasks by incorporating appropriate reaction force. And TNRD method is highly efficient in that it preserves the structural simplicity of diffusion models and take only a small number of diffusion steps. Subsequent work combines convolution neural network with image denoising such that the denoising performance has been greatly improved. Mao et al. [11] proposed a very deep residual encoder-decoder networks (RED-Net) for image restoration. The network consists of a chain of symmetric convolutional layers and deconvolutional layers. And some skip connections were added between corresponding convolutional and deconvolutional layers. DnCNN [12] further explores the potential of convolution neural networks. It utilises residual learning and batch normalisation to speed up the training process and boost the denoising performance. DnCNN can handle Gaussian denoising with unknown noise level as well as Gaussian noise at a certain noise level. DnCNN greatly improves the performance of denoising, but it still has the drawback that tends to produce over-smooth textures. In order to improve the performance of DnCNN, a fast and flexible denoising convolution neural network (FFDNet) [14] was proposed. The FFDNet method works on the downsampled sub-images, and a tunable noise level map is input as well as noisy image. It improves the performance of blind image denoising, but the noise level map of test image is unavailable generally. Most of the current discriminative learning based image denoising methods used deep convolution networks merely. In this way, the denoising model can be easily trained and a high PSNR can be obtained because the objective function is closely related to PSNR. However, it may lose image details that are more easily perceived by the human visual system.

The generative adversarial network (GAN) [21] that has achieved state-of-the-art results in image generation [22] can effectively solve the problem of over-smooth textures. Generative adversarial network consists of two parts: the generating model and the discriminating model. The generating model is trained to generate images that are difficult to distinguish from real images, and the discriminating model determines whether a sample is a real image or a generated image from the generating model. Two networks iteratively train by combating each other to generate realistic images. It can generate images that are closer to the real images, and can also generate images with more detailed information.

However, current image denoising methods based on the generative adversarial network didn’t combine the advantages of GAN and convolution network which may deteriorate the denoising effect. A CNN architecture for blind image denoising was proposed [23] which combines three multi-scale feature extraction layer, an $L_p$ regulariser and a three step adversarial training. It is an attempt to apply GAN to image denoising, but the denoising performance was unsatisfactory. Then, a GAN-CNN based blind denoiser (GCBD) was proposed [24] which aims to tackle the problem of lack of paired training data. In this method, a GAN is trained to generate noise samples which are used to construct a paired training dataset. Then the paired data is used to train the convolution neural network for denoising. It’s a novel way to solve the lack of paired training data, but it is weak in handle Gaussian denoising.

To overcome the drawbacks of existing generative adversarial network based denoising methods, a blind image denoising method based on the asymmetric generative adversarial network (ID-AGAN) is proposed in this paper. A convolutional auto-encoder with symmetric skip connections is used in our denoising model. In order to extract image information of different sizes and dimensions, a multi-scale feature downsampling layer is added in the beginning of the denoising model. Moreover, it also can reduce the effect of noise and make the image blind denoising model more flexible and robust. To optimises the information transfer between the denoising model and the discriminating model, an image downsampling structure is added before the discriminating model. The downsampling structure makes the training times and feature sizes of the denoising model and the discriminating model asymmetrical. It can make full use of GAN’s high-dimensional information recovery ability as well as making the whole model easier to train. The algorithm can balance the denoising performance and image detail retention by the improvement of the model structure.
The remainder of this paper is organised as follows. Section 2 provides image denoising methods based on convolution neural networks and generative adversarial networks. Then a description of our algorithm is presented in Section 3. Experimental results are provided in Section 4. The experimental results demonstrate the effectiveness of the ID-AGAN denoising method. Finally, conclusions are drawn in Section 5.

2 RELATED WORK

Before introducing the method proposed in this paper, the convolution neural networks for image denoising and GAN are briefly introduced in this section.

2.1 Convolution neural networks for image denoising

There are many methods [11, 12, 14, 24] using convolution neural networks for image denoising. In [11], the encoder–decoder with skip connections is used to improve the denoising performance and this model lead to massive feature reuse. Denoising performance has been improved while the gradient explosion and gradient dispersion caused by the network depth are avoided. In [12], residual learning [25] is used to boost the denoising performance. Residual learning of convolution neural networks was originally proposed to alleviate gradient explosion and gradient dispersion. It has achieved notable improvement in object detection and image classification. However, the residual learning used in DnCNN [12] is different from [25]. In [12], the whole network is a residual block, that is, the output of the denoising network is the noise of the noisy image instead of denoising image.

There are two advantages for taking residual learning. First of all, the Gaussian noise distribution is easy to fit. In addition, the effect of image content on training results can be mitigated by training to generate noise. Another reason for convolution neural networks based methods improving performance is the mean squared error (MSE) loss function. PSNR is one of the most commonly used indicators to evaluate the performance of denoising. And there is a close relationship between MSE and PSNR. With the development of deep learning, loss can be reduced to a diminutive level by large-scale neural networks. Meanwhile, PSNR has been upgraded.

2.2 Generative adversarial network

Generally, GAN consists of a generative model and a discriminative model. The discriminative model is trained to determine whether a sample is from data distribution or the generative model. Simultaneously, the generative model is trained to generate samples close to the learned distribution to fool the discriminative model. In training, the two parts compete with each other to generate samples close to the data distribution. GAN show the potential to learn complex distributions in many applications [26–28]. However, it is generally known that GAN is unstable in training. To solve this problem, many optimisation methods and GAN variants have been proposed. Radford et al. [29] proposed a deep convolutional generative adversarial network (DCGAN) to improve the stability of GAN training by importing convolutional network. Mao et al. [30, 31] proposed a least squares generative adversarial network (LSGAN) that adapts the least squares loss function for the discriminator. It overcomes the problem the vanishing gradient problem during the learning process. Arjovsky et al. [32, 33] proposed a Wasserstein generative adversarial network (WGAN). Wasserstein distance is introduced in WGAN. WGAN not only solves the problem of unstable training, but also provides a reliable training process indicator. If GANs could be employed to image denoising, the problem of over-smooth in the denoising method based on discriminative learning can be improved.

3 THE PROPOSED ID-AGAN METHOD

In order to improve the performance of the existing denoising methods using the generative adversarial network, a new blind image denoising method based on the asymmetric generative adversarial network (ID-AGAN) is proposed in this section. The overall model of the proposed ID-AGAN algorithm is illustrated in Figure 1. The ID-AGAN consists of two parts: denoising model and discriminating model. The denoising model is a deep convolutional auto-encoder with skip connections and multi-scale downsampling input layers. And the discriminating model is a five-layer fully convolution network with an image downsampling layer. The specific network structure and training details were given in the following.

3.1 The denoising model

Our denoising model is modified by U-net [34]. On the basis of a deep convolutional auto-encoder with skip connections, a multi-scale downsampling input layer has been added in the front. More precisely, the noisy image is input into three parallel convolutional blocks (shown in Figure 1: Downsample 1-3) and sampled to different sizes firstly. Each downsample block is composed of three convolutional layers. And the first two convolutional layers have the same size of feature maps. The last convolutional layer in different downsample blocks reduce the size of the feature maps by two times, four times and eight times, respectively. Then these three convolutional blocks are connected to the auto-encoder. More content information from different dimensions of the image can be extract by the multi-scale downsampling layer. And it can alleviate the image texture information losing obviously.

In our model, the noise information is extracted from the noisy image using a combination of the multi-scale downsampling input layer and the encoder firstly. Then, the noise is reconstructed using combination of convolutional layers and transposed convolutional layers. Here, the corresponding sampling layers and the reconstruction layers are connected, that
FIGURE 1  The overall model of the proposed ID-AGAN Algorithm. It consists of four mainly modules: (a) denoising model, which is used to extract the noise; (b) discriminating model, which can distinguish between denoising image and clean image to guide the training of denoising model. The numbers on the square are the number and size of the filters.

is, the input of a certain reconstruction layer is the output of the previous reconstruction layer and the output of the corresponding sampling layer. There are two reasons for skip connections. First, gradients are often vanishing or exploding in deeper networks, so it is difficult for the network to converge. Adding skip connections allows for direct information exchange between the shallow and deep layers of the network, which can effectively solve this problem. In addition, not only the information extracted by convolution is needed, but a lot of information of original image is needed in reconstruction.

3.2  The discriminating model

GAN [21] can achieve excellent results in image generation. However, GAN is not good enough in image denoising. Image denoising is to remove the noise of the image and preserve the original information of the image. The content difference between noisy image and original image is small. And the correlation of the image content is small. The one-to-one alternate training between the denoising model and the discriminating model makes it difficult to converge for the discriminating model, and does not improve the detail retention of the denoising model. Therefore, an image down sampling layer is introduced before the discriminating model to reshape the input image and increase the training times of discriminating model in each iteration. The denoising image is downsampled by $b \times w$ to $4 \times \frac{1}{2} b \times \frac{1}{2} w$ and then entered into the discriminating model during training. Hence the discriminating model is trained four times while the denoising model is trained once. Simultaneously in order to balance the influence of each part of the image on the training, the loss of each downsampled sub-image is calculated independently and the mean of the loss is fed back into the denoising model.

Our discriminating model is a full convolution network, which consists of five convolutional layers. The size of feature map is reduced by half from $64 \times 64$ layer by layer, and the number of feature map is doubled from 32 layer by layer. The output one-dimensional vector represents the probability that the input image is the original image rather than the denoising image.

3.3  Training

The input of our denoising network is $y = x + n$. Here $x$ is clean images, $y$ is noisy images and $n$ is noises. And we can estimate $f(y) \approx n$ by residual learning. Then the denoising images $x' = y - f(y)$ can be obtained. The loss function of the denoising model is shown in Eq.(1) and it consists of two parts: the mean square error (MSE) loss and the discriminating loss. In training, we set $\lambda = 0.01$.

$$L(\theta_G) = L_C(\theta_G) + \lambda \times L_G(\theta_G),$$ (1)

where $L_C(\theta_G)$ is the MSE loss between the noise estimate and the added noise as the loss and it can be written as:

$$L_C(\theta_G) = \|G(y) - (y - x)\|^2.$$ (2)

Here $\theta_G$ refers to the trainable parameter in the denoising model, $\{(y, x)\}$ refers to a pair of training images during training.
In addition to the MSE loss, adversarial training is imported to guide the denoising model. Adversarial training [21] aims to train a generator $G$ to generate samples from real data distribution. Some noises $z$ are entered into the generator to learn the mapping to the data space. And it generates samples $G(z)$. On the other hand, discriminator $D$ aims to distinguish between generated samples and original samples. It takes a data sample as input and outputs the probability of the sample coming from real data. Here, our denoising model is regarded as a generator. Instead of generating samples from noise $z$, our denoising model takes noisy images and generate denoising images. And our discriminating network aims to distinguish between clean images and denoising images. If the input sample is a clean image, it aims to output '1'. And denoising image corresponds to '0'. Our model is trained to find optimum parameters $\theta$ image, it aims to output '1'. And denoising image corresponds to clean images and denoising images. If the input sample is a clean image, it aims to output '1'. And denoising image corresponds to '0'. Our model is trained to find optimum parameters $\theta_C$ and $\theta_D$ satisfying

$$\theta^*_C, \theta^*_D = \min_{\theta_C} \max_{\theta_D} \log(D(x)) + \log(1 - D(y - G(y))). \quad (3)$$

Here $\theta_D$ refers to the trainable parameter in the discriminating model. For denoising model, the discriminating loss can be written as

$$L_D(\theta_D) = \log(1 - D(y - G(y))). \quad (4)$$

And for discriminating model, the discriminating loss is

$$L_G(\theta_G) = \log(1 - D(y - G(y))). \quad (5)$$

The denoising model and discriminating model are trained iteratively to optimise the parameters in both models simultaneously. The result is better than the deep convolution network optimised by the mean square error.

## 4 EXPERIMENTS

In this section, the performance of the proposed method are compared with several state-of-the-art denoising methods, including BM3D [3] based on filtering, EPLL [5] based on effective prior, NCSR [6] based on sparse coding, WNNM [7] and MC-WNNM [8] based on low rank and two discriminative learning based methods TNRD [10] and DnCNN [12]. Among them, TNRD and DnCNN are implemented by GPU. Moreover, all methods are tested and evaluated based on the parameter settings indicated in the author’s paper. The code of our ID-AGAN model can be downloaded at https://github.com/wangemm/ID-AGAN.

### 4.1 Experimental setting

In the experiments, 400 training images in Berkeley Segmentation Data Set and Benchmarks 500 (BSDS500) are used as the training set. For Gaussian denoising with unknown noise level, we set the standard deviation range of the noise distribution level as $\sigma \in [0, 55]$. The patch size is set as $256 \times 256$ and $55 \times 400$ patches are cropped to train the model. The ADAM algorithm is adopted to optimise ID-AGAN, and the mini-batch size is set as 8. We train 50 epochs for our ID-AGAN model. The learning rate starts with 1e-3 in the first 30 iterations and a smaller learning rate of 1e-4 is adopted for additional 20 epochs to fine-tune the network parameters. The denoising model and the discriminating model are trained at the same learning rate.

In order to evaluate the performance of our ID-AGAN method on gray images, two different datasets are employed in the experiments. One of them is 68 test images in the Berkeley Segmentation Data Set (BSD68), and the other is a dataset that includes 12 popular test images (SET12). It is worth noting that all these images are widely used for the evaluation of Gaussian denoising methods and they are not included in the training dataset.

In addition to gray image denoising, we also train the blind colour image denoising model referred to as CID-AGAN. We use 400 training images from colour version of the BSDS500 dataset for training the CID-AGAN model. The noise levels are also set into the range of $[0, 55]$ and $55 \times 400$ patches of size $256 \times 256$ are cropped to train the model.

In the training of colour image denoising, We train 20 epochs for our ID-AGAN model. The first 10 iterations are trained with learning rate 1e-3, and the next 10 iterations are adjusted to 1e-4 to fine-tune the network parameters. And the denoising model and the discriminating model are trained at the same learning rate. In addition to the colour version of the BSD68 testset and the colour version of the SET12 testset, we also...
TABLE 3  The PSNR(dB) results of different methods on the SET12 dataset

| Image   | C.man | Houss. | Pepp. | Star. | Mona. | Airp. | Parr. | Lena | Barb. | Boat | Man | Coup. | Average |
|---------|-------|--------|-------|-------|-------|-------|-------|------|-------|------|-----|------|---------|
| Noise level $\sigma = 15$ |
| BM3D    | 31.91 | 34.93  | 32.69 | 31.14 | 31.85 | 31.07 | 31.37 | 34.26 | 33.10 | 32.13 | 31.92 | 32.10 | 32.37    |
| EPLL    | 31.78 | 32.60  | 32.60 | 30.96 | 32.09 | 31.16 | 31.35 | 33.83 | 31.30 | 31.91 | 31.92 | 31.90 | 31.95    |
| WNNM    | 32.17 | **35.13** | 32.99 | 31.82 | 32.71 | 31.39 | 31.62 | 34.27 | **33.60** | 32.27 | 32.11 | 32.17 | 32.69    |
| NCSR    | 32.01 | 35.02  | 32.72 | 31.50 | 32.36 | 31.16 | 31.39 | 34.12 | 33.03 | 32.04 | 31.94 | 31.98 | 32.44    |
| TNRD    | 32.19 | 34.53  | 33.04 | 31.75 | 32.56 | 31.46 | **31.63** | 34.24 | 32.13 | 32.14 | 32.23 | 32.11 | 32.50    |
| DnCNN   | 32.10 | 34.93  | **33.15** | 32.02 | **32.94** | 31.56 | **31.63** | 34.56 | 32.09 | **32.35** | 32.41 | 32.41 | 32.68    |
| ID-AGAN | **32.21** | 34.99  | 33.03 | **32.28** | 32.75 | **31.62** | 31.61 | **34.63** | 32.23 | 32.34 | **32.42** | **32.42** | **32.71**   |
| Noise level $\sigma = 25$ |
| BM3D    | 29.45 | 32.85  | 30.16 | 28.56 | 29.25 | 28.42 | 28.93 | 32.07 | 30.71 | 29.90 | 29.61 | 29.71 | 29.97    |
| EPLL    | 29.16 | 32.04  | 30.11 | 28.41 | 29.39 | 28.60 | 28.93 | 31.55 | 28.51 | 29.69 | 29.53 | 29.48 | 29.62    |
| WNNM    | 29.64 | 33.22  | 30.42 | 29.03 | 29.84 | 28.69 | 29.15 | 32.24 | **31.24** | 30.30 | 29.76 | 29.82 | 30.26    |
| NCSR    | 29.43 | 32.96  | 30.08 | 28.70 | 29.53 | 28.42 | 28.82 | 31.92 | 30.06 | 29.69 | 29.52 | 29.46 | 29.88    |
| TNRD    | 29.72 | 32.53  | 30.57 | 29.02 | 29.85 | 28.88 | 29.18 | 32.00 | 29.41 | 29.91 | 29.87 | 29.71 | 30.05    |
| DnCNN   | **29.94** | 33.05  | **30.84** | 29.34 | 30.05 | **29.09** | **29.35** | 32.42 | 29.69 | 30.20 | 30.09 | **30.10** | 30.36    |
| ID-AGAN | 29.92 | **33.23** | 30.71 | **29.85** | **30.14** | 29.07 | 29.14 | **32.61** | 29.77 | **30.22** | **30.12** | **30.10** | **30.41**   |
| Noise level $\sigma = 50$ |
| BM3D    | 26.13 | 26.69  | 26.68 | 25.04 | 25.82 | 25.10 | 25.90 | 29.05 | 27.22 | 26.78 | 26.81 | 26.46 | 26.72    |
| EPLL    | 26.10 | 26.12  | 26.80 | 25.12 | 25.94 | 25.31 | 25.95 | 28.68 | 24.83 | 26.74 | 26.79 | 26.30 | 26.47    |
| WNNM    | 26.45 | 30.33  | 26.95 | 25.44 | 26.32 | 25.42 | 26.14 | 29.25 | **27.79** | 26.97 | 26.94 | 26.64 | 27.05    |
| NCSR    | 26.16 | 29.63  | 26.53 | 25.15 | 25.68 | 25.20 | 26.00 | 28.89 | 27.10 | 26.60 | 26.60 | 26.21 | 26.65    |
| TNRD    | 26.62 | 29.48  | 27.10 | 25.42 | 26.31 | 25.59 | 26.16 | 28.93 | 25.70 | 26.94 | 26.98 | 26.50 | 26.61    |
| DnCNN   | 27.03 | 30.02  | 27.39 | 25.72 | 26.83 | 25.89 | **26.48** | 29.38 | 26.38 | 27.23 | 27.23 | 26.91 | 27.21    |
| ID-AGAN | **27.07** | **30.57** | **27.48** | **26.43** | **26.90** | **25.95** | 26.36 | **29.62** | 26.54 | **27.35** | **27.33** | **27.04** | **27.39**   |

use PASCAL VOC 2010 Segmentation dataset to evaluate the denoising performance of the colour images.

Moreover, to evaluate the robustness of the proposed ID-AGAN method, we also compare the performance of speckle denoising. For speckle denoising with unknown noise level, we use the same dataset, training epochs and learning rate as colour image Gaussian denoising. The standard deviation range of the noise distribution are also set into the range of [0,55]. And we use the colour version of the BSD68 testset to evaluate the denoising performance of speckle denoising.

4.2 Evaluation

The PSNR and SSIM results of different methods on the BSD68 dataset are given in Tables 1 and 2. The best results are marked in bold. It can be seen that, our ID-AGAN can achieve the best PSNR and SSIM performance at all noise levels compared to other methods. Compared with the benchmark method BM3D, methods such as WNNM can increase by up to 0.22 dB in PSNR and 0.0047 in SSIM at these three noise levels. And discriminative learning based method TNRD can achieve an average improvement of 0.36 dB in PSNR and 0.0118 in SSIM while DnCNN can increase by about 0.58 dB in PSNR and 0.0154 in SSIM. On the contrary, the proposed ID-AGAN method can increase the PSNR by 0.66 dB, and it is the only one of these methods that can exceed 29 dB in average. It is worth mentioning that, our ID-AGAN outperform BM3D by about 0.71 dB when $\sigma = 50$, which achieves the estimated PSNR bound over BM3D (0.7 dB) in [35]. Meanwhile, it can increase the SSIM by about 0.0172 compared with BM3D.

The PSNR and SSIM results of different methods on SET12 are given in Tables 3 and 4. The best results are marked in bold. From Tables 3 and 4, it can be seen that ID-AGAN achieves the best results on most images. To be more specific, the average PSNR and SSIM obtained by ID-AGAN are better in all noise level compared to other methods. The best PSNR results are received in 6, 7 and 10 images, and the best SSIM results are received in 5, 7 and 7 images, respectively. Compared with the benchmark method BM3D, discriminative learning based method TNRD can achieve an average improvement of 0.1 dB in PSNR and 0.0057 in SSIM at these three noise levels. And our ID-AGAN
TABLE 4  The SSIM results of different methods on the SET12 dataset

| Image C.man Hous. Pepp. Star. Mona. Airp. Parr. Lena Barb. Boat Man Coup. Average |
|---------------------------------------------------------------|
| Noise level $\sigma = 15$                                      |
| BM3D  0.9015 0.8913 0.9068 0.9010 0.9364 0.9000 0.8949 0.8956 0.9233 0.8540 0.8672 0.8765 0.8957 |
| EPLL  0.9083 0.8772 0.9054 0.9039 0.9362 0.9028 0.8987 0.8997 0.9043 0.8524 0.8719 0.8721 0.8936 |
| WNNM  0.9036 0.8909 0.9098 0.9081 0.9425 0.9029 0.8968 0.8973 0.9277 0.8549 0.8701 0.8765 0.8983 |
| NCSR  0.9037 0.8950 0.9044 0.9050 0.9380 0.9029 0.8968 0.8975 0.9277 0.8549 0.8701 0.8765 0.8983 |
| TNRD  0.9050 0.8926 0.9139 0.9157 0.9424 0.9080 0.9062 0.9091 0.9186 0.8601 0.8797 0.8770 0.9018 |
| DnCNN 0.9047 0.8870 0.9171 0.9175 0.9477 0.9086 0.9064 0.9186 0.8645 0.8833 0.8855 0.9034 0.9034 |
| ID-AGAN 0.9059 0.8868 0.9148 0.9198 0.9441 0.9093 0.9062 0.9011 0.9195 0.8642 0.8833 0.8855 0.9036 |
| Noise level $\sigma = 25$                                      |
| BM3D  0.8544 0.8590 0.8676 0.8509 0.8997 0.8554 0.8499 0.8607 0.8874 0.8014 0.8047 0.8204 0.8510 |
| EPLL  0.8555 0.8469 0.8652 0.8491 0.8974 0.8594 0.8497 0.8500 0.8448 0.7969 0.8071 0.8102 0.8444 |
| WNNM  0.8600 0.8611 0.8733 0.8579 0.9089 0.8621 0.8524 0.8655 0.8978 0.8019 0.8078 0.8201 0.8557 |
| NCSR  0.8570 0.8591 0.8670 0.8531 0.9030 0.8576 0.8488 0.8604 0.8867 0.7945 0.8033 0.8108 0.8501 |
| TNRD  0.8571 0.8563 0.8772 0.8712 0.9091 0.8633 0.8618 0.8610 0.8734 0.8118 0.8170 0.8226 0.8568 |
| DnCNN 0.8582 0.8601 0.8788 0.8727 0.9163 0.8669 0.8623 0.8670 0.8772 0.8124 0.8202 0.8283 0.8600 |
| ID-AGAN 0.8583 0.8620 0.8779 0.8791 0.9159 0.8671 0.8617 0.8681 0.8769 0.8129 0.8213 0.8291 0.8609 |
| Noise level $\sigma = 50$                                      |
| BM3D  0.7828 0.8122 0.7936 0.7433 0.8200 0.7722 0.7809 0.7994 0.7946 0.7053 0.7056 0.7068 0.7681 |
| EPLL  0.7617 0.7845 0.7832 0.7392 0.8124 0.7670 0.7655 0.7718 0.7017 0.6950 0.6976 0.6901 0.7475 |
| WNNM  0.7848 0.8231 0.8008 0.7596 0.8350 0.7850 0.7847 0.8059 0.8100 0.7083 0.7091 0.7135 0.7767 |
| NCSR  0.7835 0.8161 0.7987 0.7452 0.8260 0.7746 0.7805 0.8035 0.7892 0.6971 0.6996 0.6941 0.7673 |
| TNRD  0.7891 0.8069 0.8049 0.7652 0.8353 0.7891 0.7807 0.8009 0.7727 0.7074 0.7099 0.7188 0.7734 |
| DnCNN 0.7960 0.8018 0.8052 0.7734 0.8402 0.7915 0.7814 0.8014 0.7737 0.7132 0.7127 0.7212 0.7760 |
| ID-AGAN 0.7960 0.8102 0.8071 0.7784 0.8399 0.7919 0.7801 0.8021 0.7744 0.7137 0.7141 0.7230 0.7776 |

can outperform BM3D by about 0.43 dB in PSNR and 0.0091 in SSIM in average. Moreover, it can outperform DnCNN by about 0.09 dB in PSNR and 0.0009 in SSIM in average.

Figures 2 and 3 compare the visual effects of different methods. Generally, as shown in Figure 2, the proposed ID-AGAN method produces better PSNR and SSIM compared to other methods. The removal of noise by BM3D and TNRD is not complete, and there are many noise points. DnCNN are cleaner in noise removal, they work well in smooth areas, but the details of the image are partially reduced. As can be seen from the enlarged part of the results, the proposed ID-AGAN method does better in terms of noise removal and detail retention. Figure 3 shows the denoising results of different methods on an image with $\sigma = 50$. Due to the high noise, almost all methods cannot completely remove noise. Contrarily, ID-AGAN can fully remove noise and have notable PSNR and SSIM increases.

We also compare the effects on the colour image with the benchmark method CBM3D, the low rank based method MC-WNNM and the discriminative learning based method CDnCNN. The PSNR and SSIM results on the colour image dataset CBSD68 are given in Table 5. From Table 5, we can see that CID-AGAN is improved at any noise level compared to other methods. Compared to the benchmark method CBM3D, the PSNR can be increased by 0.69 dB while SSIM can be increased by 0.0184. And as the noise increases, the improvement of CID-AGAN compared with CBM3D becomes more notable. Moreover, compared to the deep learning based approach CDnCNN, the PSNR and SSIM on the CBSD68 test set can be greatly improved. In order to evaluate the effectiveness of the proposed CID-AGAN, the PASCAL VOC 2010 Segmentation dataset with 1928 colour images are used in the following. The PSNR and SSIM results on the PASCAL VOC 2010 Segmentation dataset are given in Table 6. Here, we also record the standard deviation in the table. It can be seen clearly that the CID-AGAN almost achieves the best performance than the competing methods. Although the CDnCNN can achieve smaller standard deviation compared to CID-AGAN, our method outperforms it by about 0.34 dB in PSNR and 0.0064 in SSIM.

The visual effects of the methods CBM3D, CDnCNN, MC-WNNM and CID-AGAN for colour image denoising are shown in Figures 4 and 5. CBM3D and CDnCNN tend to
produce over-smooth textures. Meanwhile, it can be seen from the background of the results by CBM3D and MC-WNNM that noise has not been removed completely. As you can see in Figure 5, the CBM3D retains image texture, but it does not completely remove the noise. While the CDnCNN removes noise thoroughly, it also removes many image textures as noise, such as the creases of butterfly’s wings in the image.

Figure 6 shows the average PSNR improvement over BM3D/CBM3D and DnCNN/CDnCNN in respect of different noise levels by ID-AGAN/CID-AGAN method. The results are evaluated on the gray/colour BSD68 dataset. It can be seen that our ID-AGAN/CID-AGAN models consistently outperform BM3D/CBM3D by a large margin on a wide range of noise levels. And the PSNR improvement increases with the noise level increasing. Meanwhile, our ID-AGAN/CID-AGAN models consistently surpass BM3D/CBM3D by about 0.1 dB in gray image and 0.3 dB in colour image.

In the following, the performance of CBM3D, MC-WNNM, CDnCNN and CID-AGAN are compared for the colour images with speckle noise. Here, the PSNR and SSIM for the CBSD68 datasets are given in Table 7. As one can see, CID-AGAN still outperforms CBM3D, MC-WNNM and CDnCNN for speckle noise. For all noise level, it surpasses benchmark method CBM3D by a large margin. And it outperforms CDnCNN by about 1.51 dB in PSNR and 0.0324 in SSIM. The visual effects of the methods CBM3D, CDnCNN, MC-WNNM and CID-AGAN for speckle denoising are shown in Figure 7. It can be seen that CID-AGAN can produce sharp edges and fine details whereas MC-WNNM and CBM3D tend to generate blurred edges. The CDnCNN tends to generate striation in the picture.

According to the above evaluation and observation, the proposed ID-AGAN method can obtain similar or better evaluation than the current denoising methods. Moreover, it can
FIGURE 3 The visual effects of different methods on an image with $\sigma = 50$. The numbers below are PSNR (dB) and SSIM.

TABLE 5 The results of PSNR(dB) / SSIM of different methods on the color image dataset CBSD68

| Methods / $\sigma$ | 15       | 25       | 50       | Average          |
|-------------------|----------|----------|----------|------------------|
| CBM3D             | 33.49/0.9217 | 30.68/0.8674 | 27.36/0.7624 | 30.51/0.8505     |
| MC-WNNM           | 29.61/0.7812 | 26.73/0.7264 | 23.16/0.5554 | 26.50/0.6877     |
| CDnCNN            | 33.68/0.9298 | 31.22/0.8765 | 27.89/0.7768 | 30.88/0.8610     |
| CID-AGAN          | 34.03/0.9320 | 31.40/0.8851 | 28.17/0.7895 | 31.20/0.8689     |

TABLE 6 The PSNR(dB) and SSIM results of different methods on the PASCAL VOC 2010 Segmentation dataset

| Methods / $\sigma$ | 15       | 25       | 50       | Average          |
|-------------------|----------|----------|----------|------------------|
| CBM3D             | 33.25±2.5267/0.9085±0.0366 | 30.64±2.7498/0.8577±0.0527 | 27.36±2.9498/0.7594±0.0895 | 30.42±2.7421/0.8419±0.0596 |
| CDnCNN            | 33.44±2.1299/0.9100±0.0318 | 31.18±2.4089/0.8723±0.0486 | 28.26±2.7686/0.7720±0.0749 | 30.96±2.4355/0.8514±0.0518 |
| CID-AGAN          | 33.78±2.4920/0.9195±0.0358 | 31.55±2.6934/0.8759±0.0479 | 28.55±2.7440/0.7781±0.0711 | 31.29±2.6431/0.8578±0.0516 |

obtain the best details restoration of the original image. It is more prominent in colour images and larger noisy images.

4.3 Experiment on railway image denoising

In this section, we apply CBM3D, CDnCNN, MC-WNNM and CID-AGAN for railway images denoising. In the experiment, some railway images from PLACE365 dataset are used to evaluate the effectiveness of different methods. 16 images which shown in Figure 8 are selected as test images to evaluate the performance of CBM3D, CDnCNN, MC-WNNM and CID-AGAN. And all discriminative learning based methods denoise railway images without re-training. The PSNR and SSIM results are shown in Table 8. As can be seen from Table 8, our
FIGURE 4 The visual effects of different methods on a colour image with $\sigma = 40$. The numbers below are PSNR (dB) and SSIM.

(a) Clean Image  
(b) Noisy Image 16.08dB / 0.1463  
(c) MC-WNNM 25.68dB / 0.5600  
(d) CBM3D 29.32dB / 0.7461  
(e) CDnCNN 30.12dB / 0.7797  
(f) CID-AGAN 30.52dB / 0.7811

FIGURE 5 The visual effects of different methods on a colour image with $\sigma = 30$. The numbers below are PSNR (dB) and SSIM.

(a) Clean Image  
(b) Noisy Image 18.61dB / 0.4767  
(c) MC-WNNM 24.92dB / 0.7554  
(d) CBM3D 29.17dB / 0.9178  
(e) CDnCNN 29.95dB / 0.9235  
(f) CID-AGAN 30.36dB / 0.9251
### Table 7: The PSNR(dB) and SSIM results of different methods for speckle denoising with noise level 10, 20, 30, 40, 50 on CBSD68 dataset

| Methods         | 𝜎 10  | 𝜎 20  | 𝜎 30  | 𝜎 40  | 𝜎 50  | Average         |
|-----------------|-------|-------|-------|-------|-------|-----------------|
| CBM3D           | 36.89 | 31.34 | 28.41 | 26.47 | 24.82 | 29.59          |
| MC-WNNM         | 24.23 | 24.14 | 24.00 | 23.82 | 23.61 | 23.96          |
| CDnCNN          | 40.52 | 36.33 | 33.95 | 31.57 | 30.44 | 34.56          |
| CID-AGAN        | 41.66 | 37.53 | 35.19 | 33.59 | 32.38 | 34.07          |

The visual effects of the methods CBM3D, CDnCNN, MC-WNNM, and CID-AGAN for railway images denoising are shown in Figure 9. As shown in Figure 9, CBM3D and CDnCNN tend to produce over-smooth textures. Meanwhile, it can be seen from the result obtained by MC-WNNM that noise has not been removed completely. In contrast, our method CID-AGAN can recover more details as well as removing noise more cleaner.

### Conclusion

In this paper, we focus on improving the image blind denoising performance of deep learning methods and a new image denoising method based on asymmetric generative adversarial network (ID-AGAN) is proposed. The ID-AGAN consists of two parts: denoising model and discriminating model. Here, an asymmetric generative adversarial network has been used to optimise the high-dimensional image information denoising, hence it can balance the noise removal and detail retention. Moreover, the denoising model is a deep convolutional auto-encoder with skip connections and multi-scale downsampling input layer. The multi-scale feature downsampling layer used in the denoising model can extract image information of different dimensions.
and reduce the effect of noise on training images. The superiority of the ID-AGAN algorithm over BM3D, EPLL, NCSR, WNNM, MC-WNNM, TNRD and DnCNN has demonstrated by the experiments. All the experiment results described in this paper have shown that our ID-AGAN method is effective because it can better balance the noise removal effect and image details retention. Moreover, the ID-AGAN algorithm has been applied to the railway image denoising, it also illustrated its effective and superiority. In our future research, we will extend our model by importing attention mechanism as guidance to
further optimise the denoising model to preserve more texture information of the image.

ACKNOWLEDGEMENTS
The authors would like to thank the editor and the anonymous referees for their helpful comments and suggestions to improve the quality of the paper. This research is supported by The Fundamental Research Funds for the Central Universities of China (2018)BZ001).

REFERENCES
1. Hunt, B.R., Janney, D.H.: Digital image processing at Los Alamos Scientific Laboratory. IEEE Comput. 7(5), 57–62 (1974)
2. Chatterjee, P., Milanfar, P.: Is denoising dead?. IEEE Trans. Image Process. 19(4), 895–911 (2010)
3. Dabov, K., et al.: Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process. 16(8), 2080–2095 (2007)
4. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. 15(12), 3736–3745 (2006)
5. Zoran, D., Weiss, Y.: From learning models of natural image patches to whole image restoration. Proc. IEEE Int. Conf. Computer Vision, Barcelona, Spain, 2011, pp. 479–486
6. Dong, W., et al.: Nonlocally centralized sparse representation for image restoration. IEEE Trans. Image Process. 22(4), 1620–1630 (2013)
7. Gu, S., et al.: Weighted nuclear norm minimization with application to image denoising. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 2862–2869
8. Xu, J., et al.: Multi-channel weighted nuclear norm minimization for real color image denoising. Proc. IEEE Int. Conf. Computer Vision, Venice, Italy, 2017, pp. 1105–1113
9. He, N., et al.: An improved fractional-order differentiation model for image denoising. Signal Process. 112, 125–138 (2015)
10. Chen, Y., Pock, T.: Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. IEEE Trans. Pattern Anal. Mach. Intell. 39(6), 1256–1272 (2017)
11. Mao, X., et al.: Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. Proc. Advances in Neural Information Processing Systems, Barcelona, Spain, 2016, pp. 2802–2810
12. Zhang, K., et al.: Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. Image Process. 26(7), 3142–3155 (2017)
13. Lefkimmiatis, S.: Non-local color image denoising with convolutional neural networks. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017, pp. 5882–5891
14. Zhang, K., et al.: Fidnet: Toward a fast and flexible solution for CNN-based image denoising. IEEE Trans. Image Process. 27(9), 4606–4622 (2018)
15. Knaus, C., Zwicker, M.: Progressive image denoising. IEEE Trans. Image Process. 23(7), 3141–3152 (2014)
16. Zhao, L., et al.: Local activity-driven structural-preserving filtering for noise removal and image smoothing. Signal Process. 157, 427–429 (2019)
17. Mairal, J., et al.: Non-local sparse models for image restoration. Proc. IEEE Int. Conf. Computer Vision, Kyoto, Japan, 2009, pp. 2272–2279
18. Dong, W., et al.: Nonlocal image restoration with bilateral variance estimation: A low-rank approach. IEEE Trans. Image Process. 22(2), 700–711 (2013)
19. Yair, N., Michaeli, T.: Multi-scale weighted nuclear norm image restoration. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 3165–3174
20. Fan, H., et al.: Hyperspectral image denoising with bilinear low rank matrix factorization. Signal Process. 163, 132–152 (2019)
21. Goodfellow, I.J., et al.: Generative adversarial networks. arXiv:1406.2661, 2014
22. Zhang, H., et al.: Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. Proc. IEEE Int. Conf. Computer Vision, Venice, Italy, 2017, pp. 5908–5916
23. Divakar, N., Babu, R.V.: Image denoising via CNNs: An adversarial approach. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 2017, pp. 1076–1083
24. Chen, J., et al.: Image blind denoising with generative adversarial network based noise modeling. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 3155–3164
25. He, K., et al.: Deep residual learning for image recognition. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 2015, pp. 3155–3164
26. Ledig, C., et al.: Photo-realistic single image super-resolution using a generative adversarial network. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017, pp. 105–114
27. Isola, P., et al.: Image-to-image translation with conditional adversarial networks. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017, pp. 5967–5976
28. Lin, J., et al.: Conditional image-to-image translation. Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 3155–3164
29. Radford, A., et al.: Unsupervised representation learning with deep convolutional generative adversarial networks. Proc. International Conference on Learning Representations, London, England, (2016)
30. Mao, X., et al.: Least squares generative adversarial networks. Proc. IEEE Int. Conf. Computer Vision, Venice, Italy, 2017, pp. 2813–2821
31. Mao, X., et al.: On the effectiveness of least squares generative adversarial networks. arXiv:1712.06391, 2017
32. Arjovsky, M., et al.: Wasserstein GAN. arXiv:1701.07875, 2017
33. Gulrajani, I., et al.: Improved training of Wasserstein GANs. Proc. Advances in Neural Information Processing Systems, Long Beach, CA, USA, 2017, pp. 5767–5777
34. Ronneberger, O., et al.: U-net: Convolutional networks for biomedical image segmentation. Proc. Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015, pp. 234–241
35. Levin, A., et al.: Patch complexity, finite pixel correlations and optimal denoising. Proc. European Conference on Computer Vision, Florence, Italy, 2012, pp. 73–86

How to cite this article: Wang Y, Chang D, Zhao Y. A new blind image denoising method based on asymmetric generative adversarial network. IET Image Process. 2021;15:1260–1272.
https://doi.org/10.1049/ipr2.12102