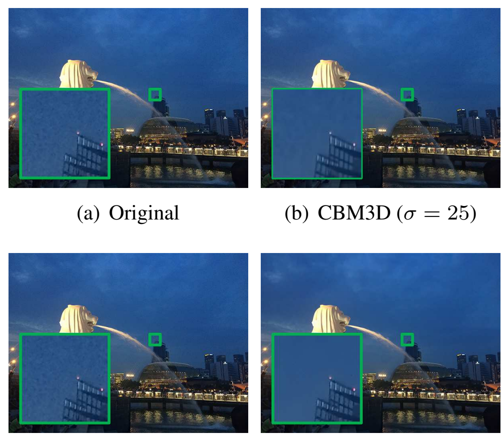
# Image Blind Denoising With Generative Adversarial Network Based Noise Modeling

Jingwen Chen, Jiawei Chen, Hongyang Chao*, Ming Yang
Sun Yat-sen University, Guangzhou, P. R. China CVTE Research, Guangzhou, P. R. China
The Key Laboratory of Machine Intelligence and Advanced Computing (Sun Yat-sen University), Ministry of Education, Guangzhou, P. R. China
{chenjw87, chenjw85}@mail2.sysu.edu.cn, isschhy@mail.sysu.edu.cn, yangming@cvte.com

## Abstract

*In this paper, we consider a typical image blind denoising problem, which is to remove unknown noise from noisy images. As we all know, discriminative learning based methods, such as DnCNN, can achieve state-of-the-art denoising results, but they are not applicable to this problem due to the lack of paired training data. To tackle the barrier, we propose a novel two-step framework. First, a Generative Adversarial Network (GAN) is trained to estimate the noise distribution over the input noisy images and to generate noise samples. Second, the noise patches sampled from the first step are utilized to construct a paired training dataset, which is used, in turn, to train a deep Convolutional Neural Network (CNN) for denoising. Extensive experiments have been done to demonstrate the superiority of our approach in image blind denoising.*



(a) Original (b) CBM3D ($\sigma = 25$)

(c) CDnCNN-B (d) Our method

Figure 1. An Example of blind denoising. (a) The original noisy image. (b) Result of CBM3D at default setting $\sigma = 25$ ($\sigma$ is the standard deviation). (c) Result of CDnCNN-B (a blind Gaussian denoising model). (d) Result of our method.

## 1. Introduction

Image denoising is a classic topic in low level vision as well as an important pre-processing step in many vision tasks. Following the degradation model $y = x + v$, image denoising targets at recovering a noise-free image $x$ from its noisy observation $y$ by reducing the noise $v$. In many cases, the noise information in the image is unavailable due to many factors such as the environment (e.g. low light) or the uncertainties of sensors. For example, photos taken by mobile phones on a specific occasion (e.g. night) are usually subject to unknown noise just as mentioned above. It is meaningful to remove these noises to improve the visual experience of users. In this paper, we focus on how to solve this blind denoising problem.

As popular solutions to denoising problems, various image prior based methods [8, 6, 21, 2], such as BM3D [4], can be extended to remove unknown noises with noise level estimated by algorithms like [19, 37, 17]. These approaches model image priors over the input image directly and achieve remarkable results. However, there are still several main drawbacks of this kind of methods. First, the image priors adopted by these methods are defined mostly based on human knowledge, so the full features of the images could be hardly captured. Therefore, the performance may be limited (see an example in Fig. 1(b)). Second, most of these approaches only use the internal information of the input image with no use of any external information. This way, though these methods can be exploited for blind denoising with some adaption, there is still room for improvement.

Besides the methods aforementioned, there are a few approaches [35, 18, 25], such as Multiscale [15], proposed to address the image blind denoising problems. These methods generally are the integration of noise model estimation and an adaptive denoising algorithm. Modeling the noise

*Corresponding author: Hongyang Chao.

plays a central role in denoising. Similar to the prior based denoising methods, most of these approaches only utilize the internal information of a single input image. Moreover, the model of noise is generally defined explicitly, which may also limit their performance.

If blind denoising is left aside, there is another type of denoising methods based on discriminative learning worth to mention. These approaches [3, 30, 31], especially those based on Convolutional Neural Network (CNN) [33, 16, 32], train a deep denoising network with paired training datasets and learn the underlying noise model implicitly, which obtains remarkable results. For the denoising problem of known noise like Gaussian noise, it is possible to form paired training data and leverage these methods to achieve state-of-the-art performance. Particularly, CNNs based approaches don't have to depend on human knowledge of image priors. They could fully exploit the great capability of the network architecture to learn from data, which breaks through the limitations of prior based methods and further improves the performance. In general, on the premise that the paired training dataset is available, this kind of approaches outperforms the previous methods. However, such a paired training dataset would be unavailable or hard to derive in reality [24]. Generally, what we can only get is noisy images with the noise information unknown. In addition, real noises are more complex so that using the existing models, which were trained for denoising known noises (e.g. Gaussian noise), to address realistic problems couldn't achieve good results (see an example in Fig. 1(c)). As such, lacking paired training datasets, these approaches might not be exploited to deal with the blind denoising problems directly.

According to the above analysis, a seductive idea comes to our mind: if it is possible to construct paired training data only from the given noisy images, the image blind denoising problem would be better solved by leveraging the advantages of discriminative learning and CNNs. Obviously, one solution to build such a training dataset is to model the noise distribution over noisy images, and then to sample noise data. As a typical approach, Gaussian Mixture Model (GMM) has been widely adopted in previous works [35, 34, 23] for noise modeling. Following their wisdom, GMM was first utilized in our experiments to model the noise over the input realistic noisy images. However, the noise samples generated from the learned model are not very similar to the observed noise. Thus, a more suitable noise modeling approach is needed under the considered scenario.

The emergence of Generative Adversarial Network (GAN) shows the possibility to us [7]. GAN is a framework for estimating generative models. This framework consists of a generative network and a discriminative network. Generally, the generative network is trained to generate samples which are hard to be distinguished from real data, while the discriminative network is trained to determine whether a sample is from real data or the generative network. GAN leverages the great capability of CNNs to learn the latent noise model implicitly, which might loosen the dependency on priors. A lot of practices [13, 28, 27] have proved that GANs could learn complex distribution. If GANs could be exploited to build a paired training dataset, the problem above would be solved. That is actually the main purpose of this paper. However, it is not trivial to realize this goal. Once we have images with unknown noise, an intuitive way to build paired training datasets is to train the generative network of GAN to learn the mapping from a clean image to an image with similar noise. Nonetheless, we found that it doesn't work this way. The generative network can be trained to learn the distribution of real noisy images and generate images with similar noise, but currently there are no mechanisms to ensure the content of the original image not to be changed. In order to solve this problem, an alternative solution is proposed: training the generative network to produce the similar kind of noises rather than noisy images.

According to the previous analysis, we propose a novel two-step framework in this paper. First, a GAN is trained to estimate the noise distribution over the input noisy images and to generate noise samples. Second, the noise patches sampled from the first step are utilized to construct a paired training dataset, which is used, in turn, to train a deep CNN for denoising the given noisy images. Details will be demonstrated in section 3. Our approach overcomes the aforementioned drawbacks of previous methods and solves the key issue of discriminative learning based denoising methods. Extensive experiments demonstrate the superiority of our approach in image blind denoising.

The major contributions of this paper are at least two folds: (1) We propose a GAN-CNN based framework to address the problem of image blind denoising, which achieves impressive results. When dealing with unknown noise, GAN is utilized to solve the key issue of building paired training datasets, and then CNN is employed for denoising. (2) To the best of our knowledge, we are the first to explore the potential of GAN in noise modeling. The ability of GAN to estimate complex distributions is exploited to learn noise distributions implicitly, overcoming the difficulty of explicitly defining the model of unknown noise.

## 2. Related Work

As the purpose of this paper is to develop an advanced denoising algorithm with GAN based noise modeling in order to improve the blind denoising performance, we briefly present three kinds of related denoising methods first: image prior based methods, noise modeling based blind denoising algorithms as well as discriminative learning based approaches. Then, the GAN is introduced subsequently.

### 2.1. Image Prior Based Denoising Methods

Most of denoising methods are based on image priors, such as BM3D [4], NSCR [6], and WNNM [8]. Most these methods do not require training data because they model the image prior over the noisy image directly, and thus can be employed to solve the denoising problem of unknown noise. One of the classic methods is BM3D, which is a benchmark in image denoising. BM3D achieves impressive results by combining the non-local self-similarity model and sparse model. More precisely, similar 2-D image fragments are grouped into 3-D data arrays which are processed by collaborative filtering subsequently. After that, jointly filtered grouped image blocks are obtained and returned to their original positions in the image. Nonetheless, there are still some problems of BM3D and most image prior based methods. First, the image priors adopted are defined mostly based on human knowledge and may limit the denoising performance. In addition, when modeling the image priors, most methods only utilize internal information of the input image while the external information from other images, such as images taken under the same condition or from a large dataset, is underused. If we could make full use of all the information and leverage the benefits of mature discriminative learning based methods as well, the denoising performance may be further boosted.

### 2.2. Noise Modeling Based Blind Denoising Methods

To our best knowledge, only a few approaches [35, 15, 18, 26] have been proposed to address image blind denoising problems. As aforementioned, these methods are the conjunction of noise modeling and an adaptive denoising algorithm generally. Multiscale [15] is an adaption of the Non-local Bayes approach [14]which assumes the noise model of each patch and its nearby patches to be zero-mean correlated Gaussian distributed. NMBD [35] proposes to model image noise with mixture of Gaussian (MoG) and develop a Low-rank MoG filter to recover the clean images. These methods only utilize the internal information of a single input image and explicitly define the noise model, which may limit the capability of noise modeling and further affect the denoising performance. If more external information, better noise modeling methods as well as advantages of discriminative learning could be exploited, the blind denoising problem would be better addressed.

### 2.3. Discriminative Learning Based Denoising Methods

To date, discriminative learning based approaches [33, 32, 16] have achieved significant success in image denoising. These methods utilize the great capability of deep network and external information from large datasets to achieve impressive results. In particular, DnCNN [32] trains a very deep CNN with residual learning and batch normalization strategies and achieves state-of-the-art results in Gaussian denoising. In addition, for blind Gaussian denoising, DnCNN trains a single network by using noisy images of different levels. Although these methods achieve high denoising quality, they cannot work in the absence of paired training data, which is often the case in reality. If the problem could be solved, this kind of methods could be exploited to better solve image blind denoising problems.

### 2.4. Generative Adversarial Network (GAN)

Recently, Generative Adversarial Network (GAN) has attracted extraordinary attention. GAN [7] has been proposed to estimate the generative model, which sidesteps some difficulties of using deep generative models, such as approximating intractable probabilistic computations. In general, GAN consists of a generative network and a discriminative network. The discriminative network is trained to determine whether a sample is from real data or the generative network. At the meantime, the generative network is trained to produce samples good enough to fool the discriminative network. During the training, the two networks compete with each other and finally the distribution which the generative network captures is as close as possible to the distribution of real data. More applications of GANs can be found in works [13, 28, 36, 5]. In these works, GANs show the potential to learn complex distributions. However, it is well known that training a GAN is tricky and unstable. DCGAN [27] has provided some useful guidelines for building and training GANs. WGAN [1, 9] further improves the training of GANs by overcoming the difficulties of maintaining the training balance between the generative network and the discriminative model and designing the network architecture. What's more, high quality samples can be generated in WGANs. If GANs could be employed to generate paired training data for blind denoising problems, aforementioned discriminative learning based methods will be easily exploited to solve the problem.

## 3. GAN-CNN Based Blind Denoiser (GCBD)

For the image blind denoising problem, the basic idea of our approach is to construct paired training data from the given noisy images, and then to train a deep denoising network for removing the noise from these given noisy images. The key issue is to build the dataset and Generative Adversarial Network (GAN) is adopted to solve this problem. As we describe in section 1, it is difficult to train the generative network generally to learn the mapping from a clean image to an image with similar noise to the given data. To ease this problem, a generative network would be trained to produce noise rather than noisy images. To simplify this problem, we assume the images dealt with having the same kind of unknown zero-mean noise [35, 20] which includes

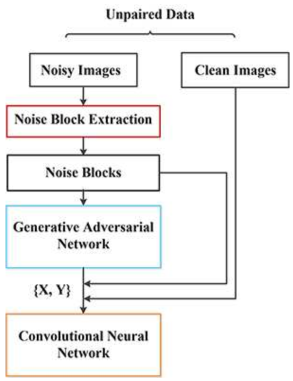



Figure 2. An overview of the proposed GCBD framework. Given unpaired data, approximate noise blocks extracted from noisy images are exploited to train a Generative Adversarial Network (GAN) for noise modeling and sampling. A large number of noise blocks are sampled from the trained GAN model. Then, both extracted and generated noise blocks are combined with clean images to obtain paired training data which is used to train a deep Convolutional Neural Network (CNN) for denoising the input noisy images.

a wide range of noises. An overview of the proposed framework is illustrated in Fig. 2. The following subsections are organized according to the steps mentioned above.

## 3.1. Noise Modeling

Before building the paired training dataset, approximate noise blocks need to be extracted from the given noisy images. Then, these blocks are used to better train the GAN for noise modeling and noise data generation.

### 3.1.1 Noise Block Extraction

This is an important step to correctly train a GAN to model the unknown noises, since the noise distribution would be better estimated from noise-dominant data. To reduce the impact of the original background, a set of approximate noise blocks (or patches), say $\mathbf{V}$, need to be extracted first from the parts with weak background in the given noisy images. This way, the noise distribution becomes the main objective for the GAN to learn, which might make the GAN model more accurate. Under the assumption that the expectation of the noise distribution is zero, an approximate noise patch can be obtained by subtracting the mean of a relatively smooth patch in noisy images. The smooth patch we discuss here refers to the region where the internal parts are very similar.

Based on the discussion above, we propose a fast smooth patch search algorithm. Let $\mathbf{p_i}$ and $\mathbf{q_j^i}$ denote a global patch with size $d \times d$ and a local patch of $\mathbf{p_i}$ with size $h \times h$ respectively. Each $\mathbf{p_i}$ is obtained by scanning the whole noisy image with stride $s_g$ and each $\mathbf{q_j^i}$ is obtained by scanning $\mathbf{p_i}$ with stride $s_l$. In the algorithm, whether $\mathbf{p_i}$ is a smooth

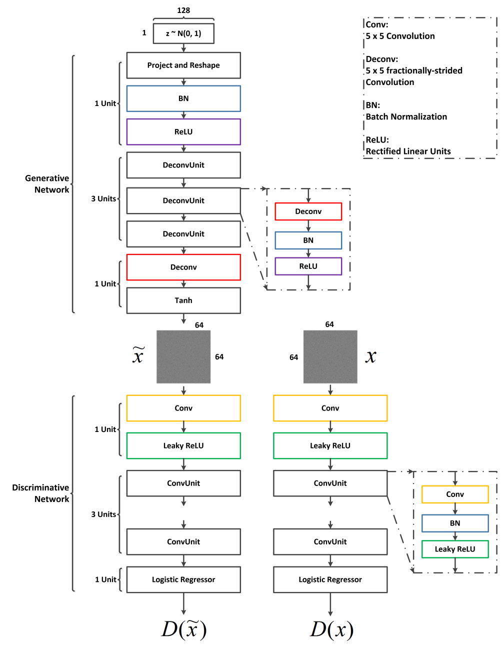



Figure 3. The network architecture of the Generative Adversarial Network. The $\tilde{x}$ is a noise block generated by the generative network, and the $x$ is a noise block extracted from noisy images. The filter number of the generative network from the second to the last unit is 256, 128, 64, and is equal to the output channel number respectively. The filter number of the discriminative network from the first to the fourth unit is 64, 128, 256, and 512 respectively.

patch is determined by the differences of mean and variance between $\mathbf{p_i}$ and $\mathbf{q_j^i}$ for each $j$. More precisely, two constraints are defined first as

$$|Mean(\mathbf{q_j^i}) - Mean(\mathbf{p_i})| \leq \mu \cdot Mean(\mathbf{p_i}), \tag{1}$$

and

$$|Var(\mathbf{q_j^i}) - Var(\mathbf{p_i})| \leq \gamma \cdot Var(\mathbf{p_i}), \tag{2}$$

where $Mean(\bullet)$ and $Var(\bullet)$ calculate the mean and the variance respectively, and $\mu, \gamma \in (0, 1)$. If for each $j$, the two constraints are satisfied, $\mathbf{p_i}$ will be regarded as a smooth patch and added to the set $\mathbf{S}$.

When $\mathbf{S} = \{\mathbf{s_1}, \mathbf{s_2}, \cdots, \mathbf{s_t}\}$ is obtained by applying the algorithm to all noisy images, the set of approximate noise blocks $\mathbf{V} = \{\mathbf{v_1}, \mathbf{v_2}, \cdots, \mathbf{v_t}\}$ can be derived by $\mathbf{v_i} = \mathbf{s_i} - Mean(\mathbf{s_i})$. Nowadays, the devices that we use generally produce high resolution images. There are a large number of eligible smooth areas in these images, such as the sky, walls and so on. Thus, sufficient smooth patches can be found in limited images, which means enough noise blocks can be extracted to train a GAN in the next step.

### 3.1.2 Noise Modeling with GAN

The patterns and quantity of noise blocks extracted in the last subsection are limited, especially when input noisy images are not enough. The results of training a deep CNN only with these blocks are not so satisfying (see discussion in Section 4.5). To boost the denoising performance, one solution is to model the noise distribution over these extracted blocks $\mathbf{V}$, and then to generate more noise data (any number of samples with more diversity) for the training of CNN. Therefore, in this subsection, the GAN will be introduced as a promising choice.

The basic idea of GAN has been briefly mentioned in section 1 and 2. As a framework to estimate generative models, GAN has the capability to learn complex distributions. What's more, GANs can be trained by back-propagation algorithm and produce noise samples by forward-propagation only without involving another component. In the proposed method, a GAN is adopted to estimate the noise distribution over a set of approximate noise blocks $\mathbf{V}$. Since WGAN [1] may improve the training of GANs and generate high quality samples as described previously in the section of related works. Therefore, in our experiments, WGAN-GP [9], which is an improved version of WGAN, is adopted to learn the noise distribution. The objective function for our task is

$$\mathcal{L}_{GAN} = \mathop{\mathbb{E}}_{\tilde{x}\sim\mathbb{P}_g}[D(\tilde{x})] - \mathop{\mathbb{E}}_{x\sim\mathbb{P}_r}[D(x)] + \lambda \mathop{\mathbb{E}}_{\hat{x}\sim\mathbb{P}_{\hat{\mathbf{x}}}}[(\parallel \nabla_{\hat{x}} D(\hat{x}) \parallel_2 - 1)^2], \tag{3}$$

where $\mathbb{P}_r$ is the distribution over $\mathbf{V}$, $\mathbb{P}_g$ is the generator distruibution, $\mathbb{P}_{\hat{x}}$ is defined as a distribution sampling uniformlly along straight lines between pairs of points sampled from $\mathbb{P}_r$ and $\mathbb{P}_g$. See more details in [9].

We adopt the similar network to DCGAN [27]. The network structure is illustrated in Fig. 3. More details can be found in [27]. The trained GAN model is used to generate noise samples for augmenting $\mathbf{V}$ and finally a larger set $\mathbf{V'} = \{\mathbf{v'_1}, \mathbf{v'_2}, \cdots, \mathbf{v'_w}\}$ is obtained.

### 3.2. Denoising with Deep CNN

Many previous works [33, 32, 16] have proposed to solve denoising problems by training a CNN with large datasets and achieved impressive results. As described in Section 1 and 2, CNNs own the great capability of the network architecture to learn the latent noise model from the paired training dataset implicitly, which loosens the dependency on human knowledge of image priors. Thus, a CNN is utilized in our framework for denoising.

In order to train the CNN, a paired training dataset need to be built first. Given the set $\mathbf{V'}$ obtained from the section 3.1.2, another set of clean images are divided into small patches of size $d \times d$ which form the set $\mathbf{X} = \{\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_e}\}$. Noise blocks in $\mathbf{V'}$ are randomly added to patches in $\mathbf{X}$ to obtain $\mathbf{Y} = \{\mathbf{y_1}, \mathbf{y_2}, \cdots, \mathbf{y_f}\}$, where $\mathbf{y_l} = \mathbf{x_j} + \mathbf{v'_k}$. The set $\mathbf{X}$ and $\mathbf{Y}$ form a paired training dataset $\{\mathbf{X}, \mathbf{Y}\}$. Actually, the dataset is built during the

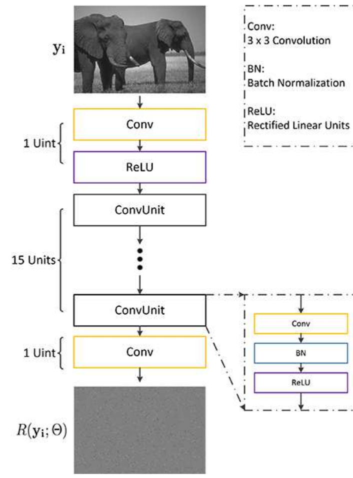



Figure 4. The network architecture of the Convolutional Neural Network. The input is a noisy image $\mathbf{y_i}$, and the output is the difference $R(\mathbf{y_i};\Theta)$ between the input and the latent clean image. The predicted clean image can be obtained by $\mathbf{y_i}$ - $R(\mathbf{y_i};\Theta)$. The filter number of the last unit is equal to the output channel number. Each of the other units contains 64 filters.

training of the denoising network. In each epoch, the combinations of $\mathbf{x_j}$ and $\mathbf{v'_k}$ are changed and a new dataset $\{\mathbf{X}, \mathbf{Y'}\}$ is obtained, which leads to further data augmentation.

Once the paired training dataset is built, a CNN can be trained for denoising finally. We adopt the similar network structure to DnCNN [32] in our experiments. The CNN is regarded as a single residual unit to predict the residual image, i.e., the difference between the input noisy image and the latent clean image. The objective function to be minimized is defined as

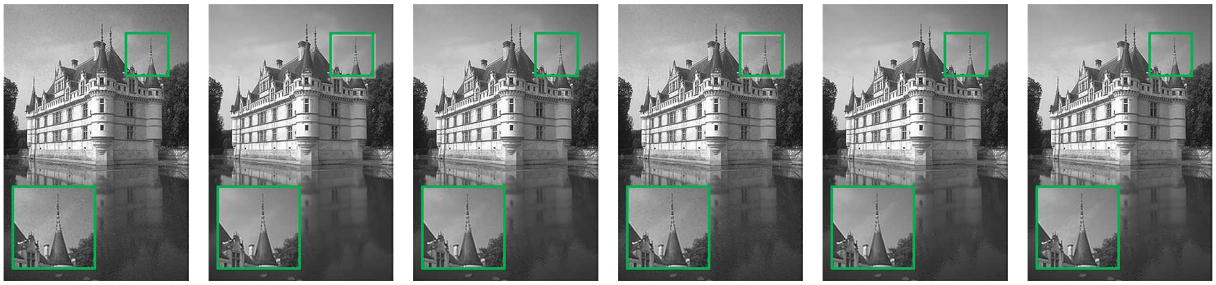
$$\mathcal{L}_{CNN}(\Theta) = \frac{1}{2N}\sum_{i=1}^{N} \parallel R(\mathbf{y_i};\Theta) - (\mathbf{y_i} - \mathbf{x_i}) \parallel_F^2, \tag{4}$$

where $\Theta$ is the parameters of the network, $N$ is the size of training data, $\mathbf{y_i}$ is a noisy image, $\mathbf{x_i}$ is the ground truth. Batch Normalization [12], ReLU [11] as well as residual learning strategy [10] are also adopted to improve the training of the deep network.

The network structure of the CNN, which consists of 17 units, is illustrated in Fig. 4. Zero padding is adopted to ensure the dimension of the input and output is consistent.

## 4. Experiments

In this section, we evaluate the proposed GCBD method on both synthetic and real-world data. Comparisons are made among several representative approaches. Four parts of experiments are carried out: (1) to verify the accuracy of noise modeling with GAN, GCBD is compared with state-of-the-art denoising methods, especially a discriminative



(a) Noisy / 34.86 (b) BM3D / 38.36 (c) DnCNN-B / 38.12 (d) Multiscale / 35.18 (e) GCBD / 40.30 (f) Ground truth

Figure 5. Comparison (PSNR/dB) on BSD68 in the evaluation of mixture noise denoising with $s = 25$. Zoom in for better view.

Table 1. The PSNR (dB) results of all the compared methods on BSD68 in synthetic noise denoising tasks.

| Gaussian Noise | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Mode | Non-Blind | | | | Blind | | | |
| Method | BM3D | EPLL | NCSR | WNNM | Multiscale | | DnCNN-B | GCBD |
| Setting | - | - | - | - | scale = 1 | scale = 2 | - | - |
| $\sigma = 15$ | 31.07 | 31.21 | 31.19 | 31.37 | 30.48 | 29.72 | 31.61 | 31.59 |
| $\sigma = 25$ | 28.57 | 28.68 | 28.62 | 28.83 | 27.58 | 26.77 | 29.16 | 29.15 |
| Mixture Noise | | | | | | | | |
| Mode | Non-Blind | | | | Blind | | | |
| Method | BM3D | EPLL | NCSR | WNNM | Multiscale | | DnCNN-B | GCBD |
| Setting | - | - | - | - | scale = 1 | scale = 2 | - | - |
| $s = 15$ | 41.08 | 41.06 | 41.06 | 41.04 | 38.99 | 37.98 | 40.75 | 42.00 |
| $s = 25$ | 37.85 | 37.76 | 37.98 | 37.63 | 35.54 | 35.12 | 37.54 | 39.87 |

learning based approach DnCNN [32], in Gaussian blind denoising tasks (in section 4.2.1); (2) to show that GCBD can deal with more complex noise besides Gaussian noise, the evaluation is conducted with mixture noise (in section 4.2.2); (3) to examine the applicability to realistic problems, another evaluation is done with a public benchmark dataset and noisy images taken by a consumer device (in section 4.3); (4) to give some discussion about the selection of noise modeling method, noise samples are shown to illustrate the reason to choose GAN rather than other traditional methods, such as GMM. Extensive experiments demonstrate the superiority of GCBD in image blind denoising problems.

## 4.1. Experimental Setting

**Experimental Data** In the experiments on synthetic data, BSD68 [29] is utilized as the test set. In the experiments on real-world data, the evaluations are conducted on a benchmark dataset Darmstadt Noise Dataset (DND) [24] and a dataset NIGHT which includes 25 high resolution noisy images (about 3000 x 2000 pixels) taken by an ordinary mobile (iPhone 5s) at night. For the proposed GCBD, a set of clean images (CLEAN1) is used to build the paired training dataset with noise data generated by GAN. In order to simulate the condition where large images are processed in reality, noises are added to another set of high resolution clean images (CLEAN2) to form the input noisy images for GCBD in the evaluation with synthetic data.

**Parameter Settings** In the noise extraction step, parameters $d$, $h$, $s_g$, $s_l$, $\mu$ and $\gamma$ are set to 64, 16, 32, 16, 0.1 and 0.25 respectively. For noise modeling with GAN, we roughly follow the parameter settings in DCGAN [27]. For the CNN, it is trained with initial learning rate 0.001 and SGD optimizer for 50 epochs.

**Compared Methods** The competing approaches include BM3D [4], EPLL [38], NCSR [6], WNNM [8], Multiscale [15], DnCNN [32] and the proposed GCBD. Particularly, in order to reveal the limitations of discriminative learning based methods when dealing with blind denoising problems, the available blind model of DnCNN for Gaussian denoising, denoted as DnCNN-B, is adopted in the evaluation. Specifically, DnCNN-B is trained with accurate Gaussian noise data from different levels (i.e. $\sigma \in [0, 55]$), which achieves state-of-the-art blind Gaussian denoising results.

## 4.2. Evaluation with Synthetic Noise

In this part, different types of zero-mean synthetic noise data are generated and added to BSD68 [29] to evaluate all the competing methods. In this evaluation, except for Multiscale, DnCNN-B and GCBD, the other methods are provided with real noise levels (i.e. the standard deviation $\sigma$).

**Gaussian Noise** It's essential to conduct experiments of blind Gaussian denoising since Gaussian noise is one of the widely-studied noises. Table. 1 shows different results of all the compared methods. Though no noise information is provided, GCBD still outperforms BM3D, EPLL, WNNM and Multiscale. Particularly, GCBD achieves comparable results with DnCNN-B. This is impressive because DnCNN-B is trained with accurate data while GCBD is trained with approximate data generated by GAN. This experiment demonstrates the accuracy of noise modeling through using GAN. More image examples can be found in the supplementary.

**Mixture Noise** Besides Gaussian noise, we further evaluate the performance of several methods in complex noise denoising tasks. The mixture noise [34] adopted in the experiments consists of 10% uniform noise $[-s, s]$, 20% Gaussian noise $N(0, 1)$ and 70% Gaussian noise $N(0, 0.01)$. Table. 1 shows the quantitative results. In this task, GCBD also performs much better than BM3D, EPLL, WNNM and Multiscale, which further shows the superiority of GCBD in blind denoising problems. Particularly, DnCNN-B cannot work well because the paired training dataset is unavailable. On the contrary, the proposed GCBD exploits GAN to estimate the noise distribution of noisy images and addresses the problem of lack of training data, which achieves remarkable denoising results. An example is shown in Fig. 5 and more examples can be found in the supplementary.

For the above cases (Gaussian and Mixture noises), the proposed method works well. Specifically, the deviation of the distribution learned by GAN over extracted noise blocks from the ground truth distribution is about 0.3% for mean and 1.5% for standard deviation on average.

### 4.3. Evaluation with Real-World Noise

**Evaluation on DND** DND [24] is a novel benchmark dataset which consists of realistic photos from 50 scenes taken by 4 consumer cameras. The competing methods are applied and evaluated in sRGB space in this part. Table. 2 shows the PSNR results. GCBD outperforms the other methods by about 1 dB at least.

Table 2. The PSNR (dB) results of all the compared methods on DND in real-world noise denoising tasks.

| Method | WNNM | EPLL | NCSR | BM3D | GCBD |
|---|---|---|---|---|---|
| PSNR (dB) | 34.44 | 33.51 | 33.81 | 34.61 | 35.58 |

**Evaluation on NIGHT** It is common that we get noisy images at night. Thus, the experiments in this part are conducted on dataset NIGHT. The goal of this evaluation is to test whether a method can reduce the noise in given noisy images and further reduce the noise in other images taken under similar conditions (i.e. same device, same scenario). The NIGHT dataset is divided into 20 images (denoted as NIGHT-A) and the other 5 images (denoted as NIGHT-B). NCSR and BM3D are evaluated at a default setting $\sigma = 25$. For the proposed GCBD, only NIGHT-A is used as input. Because images in NIGHT have no ground truth, the visual quality is the main metric. Fig. 7 shows the results


Figure 6. Noise samples (scaled). Zoom in for better view.

of different methods on NIGHT-A. As we can see, compared with the other methods, GCBD does a pretty good job at retaining details, such as the spark of the light, while removing the noise. Fig. 8 shows the results on NIGHT-B. GCBD also achieves the best result, which demonstrates that GCBD can be employed to handle noisy images taken under similar conditions.

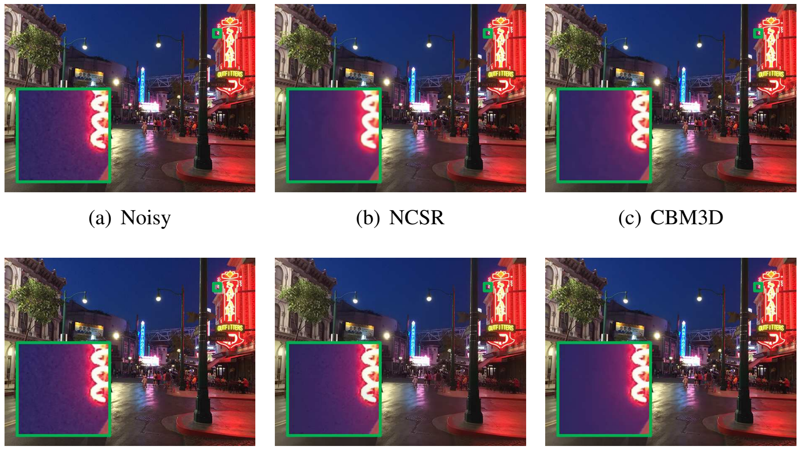
### 4.4. Selection of Noise Modeling Methods

As described in section 1, as a typical modeling method, GMM has been widely utilized in previous works like [35, 34, 23] for noise modeling. However, due to the complexity of real-world noises, a traditional pixel-based GMM may not handle realistic cases well as GAN under the considered scenario in this paper. Figure. 6 shows some noise samples to illustrate the modeling capability for real noises. The sample generated by GAN is more similar to the observed real noises than that generated by GMM, which could demonstrate that GAN may deal with more complex noises better than a simple pixel-based GMM. This is because the number of Gaussian models and the explicitly-defined model may limit the performance of GMM while GAN leverages the great capability of CNN to learn the noise model implicitly and to capture more features of noises without human knowledge of image priors. This way, GAN shows a potential in noise modeling problems. Furthermore, since the proposed framework is flexible, the noise modeling component could be replaced with better choices in the future.

### 4.5. Breakdown Performance Analysis of GCBD

The most important part of the proposed framework is noise modeling, which includes extracting noise blocks and learning the noise distribution by GAN. In this part, the effect of the first step will be investigated first followed by the discussion about the accuracy of noise modeling with GAN.

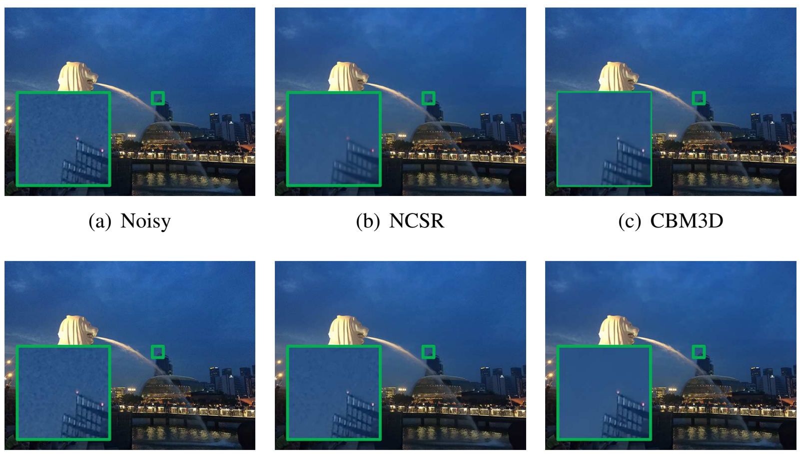
**Effect of Noise Blocks Extraction Step** If only the extracted noise blocks are used to build the dataset for CNN, the results are 0.34 dB and 0.91 dB lower than GCBD on average for Gaussian noise and Mixture noise cases respectively. This is because the extracted blocks might lack of diversity, especially when input noisy images are not enough. In contrast, GCBD leverages GAN to learn the noise distribution and then to generate samples with more diversity, which could make up for the deficiency of only using extracted blocks. Therefore, it is effective to use noise model-



(a) Noisy (b) NCSR (c) CBM3D

(d) CDnCNN-B (e) Multiscale (f) GCBD

Figure 7. Comparison on NIGHT-A in the evaluation of Real-World Noise Denoising. Zoom in for better view.



(a) Noisy (b) NCSR (c) CBM3D

(d) CDnCNN-B (e) Multiscale (f) GCBD

Figure 8. Comparison on NIGHT-B in the evaluation of Real-World Noise Denoising. Zoom in for better view.

ing for data augmentation to improve the denoising results.

**Effect of Noise Modeling with GAN** To verify the effectiveness of noise modeling with GAN, the accurate synthetic noise data were used as input to train GCBD. Taking mixture noise as an example, GCBD achieved 42.42 dB and 40.21 dB for $s = 15$ and 25 respectively, which is on par with the results of DnCNN that trained with similar accurate data (42.46 dB and 40.22 dB for $s = 15$ and 25 respectively). Moreover, when handling complex realistic noises, GCBD can learn the noise distribution well and generate good samples (see Figure. 6). All these facts indicate that using GAN to model the noises could be accurate.

## 5. Conclusion

In this paper, we attempted to improve the performance of image blind denoising by exploiting deep learning based methods in the absence of paired training data. The proposed GCBD can improve the blind denoising performance. The GAN is utilized to learn noise distribution and to build the paired training dataset to train the CNN for denoising. Extensive experiments demonstrate the superiority of our method. To the best of our knowledge, we are the first to explore the potential of GAN in noise modeling and employ it in denoising tasks. One limitation of our method is that the noise is assumed to be additive noise with zero mean. This type of noise is common in natural environment and includes a wide range of noises. If the expectation of the unknown noise is available, it would be the same procedure as our approach. Next, we will consider to overcome this limitation.

**Acknowledgments** This work is partially supported by NSF of China under Grant 61672548, U1611461, 61173081, and the Guangzhou Science and Technology Program, China, under Grant 201510010165.

## References


[1] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein gan. *arXiv:1701.07875*, 2017.

[2] A. Buades, B. Coll, and J. M. Morel. A non-local algorithm for image denoising. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 60–65, 2005.

[3] Y. Chen and T. Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1256, 2017.

[4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007.

[5] E. Denton, S. Chintala, A. Szlam, and R. Fergus. Deep generative image models using a laplacian pyramid of adversarial networks. *Advances in Neural Information Processing Systems*, pages 1486–1494, 2015.

[6] W. Dong, L. Zhang, G. Shi, and X. Li. Nonlocally centralized sparse representation for image restoration. *IEEE Transactions on Image Processing*, 22(4):1620, 2013.

[7] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial networks. *Advances in Neural Information Processing Systems*, 3:2672–2680, 2014.

[8] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2862–2869, 2014.

[9] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville. Improved training of wasserstein gans. *arXiv:1704.00028*, 2017.

[10] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2015.

[11] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *International Conference on Computer Vision*, pages 1026–1034, 2015.

[12] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, pages 448–456, 2015.

[13] P. Isola, J. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 5967–5976, 2017.

[14] M. Lebrun, A. Buades, and J. M. Morel. A nonlocal bayesian image denoising algorithm. *Siam Journal on Imaging Sciences*, 6(3):1665–1688, 2013.

[15] M. Lebrun, M. Colom, and J. M. Morel. Multiscale image blind denoising. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 24(10):3149–61, 2015.

[16] S. Lefkimmiatis. Non-local color image denoising with convolutional neural networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 5882–5891, 2017.

[17] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang. Noise estimation from a single image. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 901–908, 2006.

[18] C. Liu, R. Szeliski, K. S. Bing, C. L. Zitnick, and W. T. Freeman. Automatic estimation and removal of noise from a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):299–314, 2008.

[19] X. Liu, M. Tanaka, and M. Okutomi. Noise level estimation using weak textured patches of a single noisy image. In *IEEE International Conference on Image Processing*, pages 665–668, 2013.

[20] X. Liu, M. Tanaka, and M. Okutomi. Practical signal-dependent noise parameter estimation from a single noisy image. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 23(10):4361–71, 2014.

[21] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. Non-local sparse models for image restoration. In *IEEE International Conference on Computer Vision*, pages 2272–2279, 2010.

[22] D. R. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *IEEE International Conference on Computer Vision*, volume 2, pages 416–423, July 2001.

[23] D. Meng and F. D. L. Torre. Robust matrix factorization with unknown noise. In *IEEE International Conference on Computer Vision*, pages 1337–1344, 2013.

[24] T. Pltz and S. Roth. Benchmarking denoising algorithms with real photographs. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 2750–2759, 2017.

[25] J. Portilla. Full blind denoising through noise covariance estimation using gaussian scale mixtures in the wavelet domain. In *International Conference on Image Processing*, pages 1217–1220 Vol.2, 2004.

[26] T. Rabie. Robust estimation approach for blind denoising. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 14(11):1755–65, 2005.

[27] A. Radford, L. Metz, and S. Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *Computer Science*, 2015.

[28] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee. Generative adversarial text to image synthesis. In *International Conference on Machine Learning*, pages 1060–1069, 2016.

[29] S. Roth and M. J. Black. Fields of experts. *International Journal of Computer Vision*, 82(2):205, 2009.

[30] U. Schmidt, J. Jancsary, S. Nowozin, S. Roth, and C. Rother. Cascades of regression tree fields for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(4):677, 2016.

[31] U. Schmidt and S. Roth. Shrinkage fields for effective image restoration. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2774–2781, 2014.

[32] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Processing*, 26(7):3142–3155, 2017.

[33] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning deep CNN denoiser prior for image restoration. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 2808–2817, 2017.

[34] Q. Zhao, D. Meng, Z. Xu, W. Zuo, and L. Zhang. Robust principal component analysis with complex noise. In *The International Conference on Machine Learning*, pages 55–63, 2014.

[35] F. Zhu, G. Chen, and P. A. Heng. From noise modeling to blind image denoising. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 420–429, 2016.

[36] J. Zhu, T. Park, P. Isola, and A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv:1703.10593*, 2017.

[37] D. Zoran and Y. Weiss. Scale invariance and noise in natural images. In *IEEE International Conference on Computer Vision*, pages 2209–2216, 2009.

[38] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *International Conference on Computer Vision*, pages 479–486, 2011.