

# A robust non-blind deblurring method using deep denoiser prior

Yingying Fang<sup>1</sup> Hao Zhang<sup>2</sup> Hok Shing Wong<sup>2</sup> Tiejong Zeng<sup>2\*</sup>

<sup>1</sup>Imperial College London, London, United Kingdom

<sup>2</sup>The Chinese University of Hong Kong, Shatin, Hong Kong

y.fang@imperial.ac.uk, {haozhang, hswong, zeng}@math.cuhk.edu.hk

## Abstract

*The existing non-blind deblurring methods are mostly susceptible to noise in the given blurring kernel, which is usually estimated from the observed image. This will produce undesirable ringing artifacts around the recovered edges when the given kernel is not accurate enough. Besides, the noise and outliers in the observed images may also severely degrade the performance of the deblurring methods. Considering these factors, we designed a robust non-blind deblurring method taking all these noises into account. In this paper, we propose a kernel error term to rectify the given kernel in the midst of the deconvolution process. A residual error term is also introduced to deal with the outliers caused by noise or saturation. A deep learning denoiser prior is adopted to reserve the fine textures in the recovered image. The experiments show clearly that the proposed method achieves remarkable progress in both the visual quality and the numerical results of the recovered images compared to the state-of-the-art deblurring methods.*

## 1. Introduction

Single image deblurring aims to recover a sharp image  $x$  from a single observed blurry image  $y$ . The degradation of the blurry image from the clear one is usually modeled as

$$y = k * x + n, \quad (1)$$

where  $k$  is the blurring kernel,  $*$  denotes the convolution operator and  $n$  is some additive white Gaussian noise (AWGN). In most cases, the observed information from a blurry image is only the blurry image itself. Hence, recovery of the clear image from the blurry one is decomposed into two steps: the kernel estimation (blind deblurring) and the deconvolution of the blurry image with the estimated kernel (non-blind deblurring). In this paper, we aim to explore a robust non-blind deblurring method where the kernel is estimated by the existing blind deblurring methods.

As estimating the blurring kernel is a highly ill-posed problem, variational models which consist of a fidelity term and regularization term are one of the most investigated methods for both non-blind and blind deblurring. Many models have been successfully proposed with the focus on designing the image priors, which may favor sharper images. However, these manually-designed priors are usually based on statistical features of the natural images whereas these features are not suitable for all image type. Hence, some undesirable artifacts would be easily produced for the latent images or kernels which are not exactly consistent with the proposed prior. For this reason, variational models are refined to adopt the priors driven by a large number of natural images, which can be retrieved by the dictionary learning or deep learning techniques [13, 21, 49, 52]. With the data-driven regularizations, the models usually produce better results on a wider range of images compared to those manually-designed regularizations. With the great success of the deep convolutional neural networks (CNNs) in the imaging tasks [5, 11, 20, 48], researchers have also attempted to put forward end-to-end networks for deblurring from a single image bypassing the need of estimating the blurring kernel beforehand, simplifying the traditional two-step procedures. Despite its potential efficiency in the practical applications of deblurring, deconvolution with arbitrary kernels by a single network is of great challenge, which also depends heavily on the training data [32, 38]. Therefore, various excellent networks have been proposed with the aim to increase the deconvolution ability of the proposed network as well as its generalization to different kinds of kernels [3, 16, 29, 31, 42, 46]. However, the generalization of existing networks to different kernel types is still limited and the performance of deep CNNs on blind image deblurring still falls behind conventional optimization-based approaches on handling an arbitrary blurring kernel [38]. A blurry image whose kernel is not included in the training data tends to retain the blurring effect in its deblurred result of the network. Therefore, works with deep neural networks also resort to tackle the non-blind deblurring problem with the aim to produce a sharper image with fewer

\*Corresponding authors

artifacts [10, 15, 30, 49] and reached the state-of-the-art performance compared to the traditional methods.

As we notice, both of the above-mentioned variational models and the learning-based methods are vulnerable to the problematic factors existing in original blurry images, which turns out to be visually unpleasant solutions. The major problems in this case arise from outliers, which refer to the blurry pixels whose values are inconsistent with the linear degradation model (1). A bunch of methods were designed to relieve the ringing artifacts from the outliers [4, 23, 34, 35]. However, these methods overlook the artifacts coming from the noise in the estimated kernel. Though a bunch of works were proposed to get a more accurate kernel, [9, 24, 36, 47] the kernel errors are inevitable in the derived kernels. For this reason, some works have also shed light on the non-blind deblurring methods which take the kernel error into consideration [12, 30]. The authors in [12] treated the images convolved by the error existing in the kernels and constrained it by the sparsity in the proposed model. The latest learning-based work [30] proposed a non-blind deblurring model which takes the kernel error into consideration and designed a network to mimic the optimization iterations of the proposed model. The trained model of [30] reached state-of-the-art performance, but remained sensitive to the noise contained in the blurry image.

As shown above, while there have been amounts of works to reduce the artifacts by different ways to improve the visual quality of the deblurred results, few can consider different origins of the ringing artifacts in a single deblurring scheme and thus lead to the limitations of these methods in different cases. In this paper, we aim to design a more robust non-blind deblurring method to restore the images from a single blurry image with the estimated kernel. The contributions are summarized in the following points:

- Two error terms, named the kernel error term and the residual error term, are introduced to avoid the appearance of the ringing artifacts. The kernel error term is adopted to approximate the error in the given kernel returned by the blind deblurring methods while the residual error term is introduced to approach the outliers in the blurry images. The two terms are both constrained by the sparse property based on the observation that these errors are usually presented sparsely but incur severe artifacts in this ill-posed restoration problem.
- To get finer textures in the recovered image, an enhanced texture-reserved denoiser prior is trained and applied in the alternative optimization scheme which is based on the half-quadratic splitting methods.
- The experiments show that the proposed model gets fewer ringing artifacts with different kernel estimation methods and shows more robustness to the outliers in

the blurry images. The proposed model proves to reach both numerical and visual improvement on the benchmark datasets as well as real-world images.

## 2. Preliminary

In this section, we give a brief review on the Plug-and-Play framework as well as the existing denoiser priors, which are closely related to the proposed deblurring model.

### 2.1. Denoiser prior

The denoiser prior has been widely applied in the well-known Plug-and-Play framework for image restoration problems in the past decade [7, 39, 44, 49, 51]. The core idea of this scheme is to incorporate an implicit denoiser prior term in the variational model and decouple this term into a denoising subproblem by the splitting technique, which can be solved by off-the-shelf denoisers. Specifically, the general form of restoration model with the denoiser prior is given as follows:

$$\min_u \frac{1}{2} \|Ax - y\|_2^2 + \lambda \Phi(x), \quad (2)$$

where  $A$  is the degradation linear operator and  $\Phi(\cdot)$  is the implicit denoiser prior which indicates a lower value for a more natural and sharper image. To solve the problem (2), the splitting technique will firstly transform it into an equivalent two-variable problem by introducing a new variable  $z$ :

$$\min_{x,z} \frac{1}{2} \|Ax - y\|_2^2 + \lambda \Phi(z) + \frac{\mu}{2} \|z - x\|^2, \quad (3)$$

where  $\mu$  will increase as the iteration goes up. Then the solution to the problem (2) can be approached by iterating the following two steps:

$$x^{k+1} = \arg \min_x \frac{1}{2} \|Ax - y\|_2^2 + \frac{\mu}{2} \|z^k - x\|^2, \quad (4)$$

$$z^{k+1} = \arg \min_z \lambda \Phi(z) + \frac{\mu}{2} \|z - x^{k+1}\|^2. \quad (5)$$

For Equation (5) where the implicit prior  $\Phi(\cdot)$  indicates the good property of the images such as sharpness, the solution  $z$  can be interpreted as the denoised result of the given  $x^{k+1}$ . For this reason, the implicit term  $\Phi(\cdot)$  is hence called by the denoiser prior and the solution can be derived by the existing denoising methods such as BM3D [7] and EPLL [52]. However, with these traditional denoisers, the cost of calculation and the artifacts may be enlarged with the iterations increasing. Therefore, [49] initially introduced the learning-based denoiser as the optimization method, which has higher efficiency and better denoising performance.

As we notice, while the step (5) removes the noise and artifacts, it also has a close relation to the textures reserved in  $x^{k+1}$ . Therefore, one of the contributions of paper is to explore this implicit denoiser prior to recover more textures.

## 2.2. Multi-level Wavelet-CNN denoiser

The learning-based denoising methods have gained much progress in the last decade. In [48], the author proposed a residual network which is more efficient in training and has good performance in denoising. To get a balance of the efficiency and performance of the denoiser, the author in [49] adopted the dilated filters to increase the receptive field of their proposed 7 layers network and trained it for a set of different noise levels. Later in [25], a denoising model named Multi-level Wavelet-CNN (MWCNN) was proposed and attracted much attention. This MWCNN model was motivated by multi-level wavelet packet transform (WPT) [2, 8] and integrated with the CNN block. From WPT to MWCNN, it was designed by adding the CNN block after each level of discrete wavelet transform (DWT). In this MWCNN model, a U-Net architecture [40] consisting of a contracting subnetwork and an expanding subnetwork was applied. Different from the traditional CNN, the classical pooling operation in the contracting subnetwork was replaced by an invertible discrete wavelet transform (DWT). Correspondingly, the inverse discrete transform (IWT) was used in the upsampling stage. The inversion of the down-sampling in the MWCNN proved to be effective in keeping the detailed textures in the expanding subnetwork. Considering its advantages in reserving the textures, we explored this denoising scheme as our denoiser prior in our deblurring model.

## 3. Proposed Approach

In this section, we will firstly introduce the details of our denoiser prior in Section 3.1, which will be applied in the proposed deblurring model introduced in Section 3.2. In Section 3.3, the numerical scheme as well as the framework of the proposed method will be illustrated.

### 3.1. Texture-preserving prior

In the first step, we trained our denoisers to obtain optimal solution in Equation (5). Motivated by the idea to keep more details in the recovered images, the advanced MWCNN architecture is applied as our denoiser prior which has proven to reserve more textures in the denoising task than other networks such as DnCNN [48] and FFDNet [50]. Compared with the pooling operation and dilated filtering in the other networks, the invertible DWT and IWT in MWCNN can avoid the underlying drawbacks such as gridding effects. For the wavelet transform, we use the default Haar wavelet in MWCNN. We denote the network parameters by  $\Theta$  and the output of the network by  $F_{n_\sigma}(y; \Theta)$ . Then the loss function of the MWCNN network

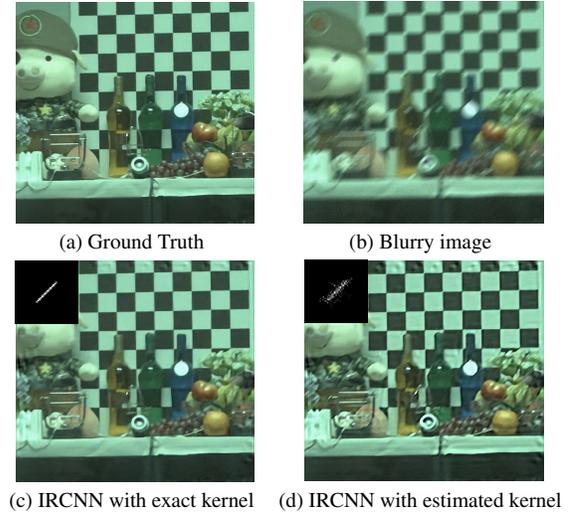


Figure 1. The deblurring results of [49] with exact kernel and estimated kernel respectively. (a) Ground truth, (b) blurry image with 2% noise, (c) deblurred result with exact kernel, (d) deblurred result with estimated kernel by [36].

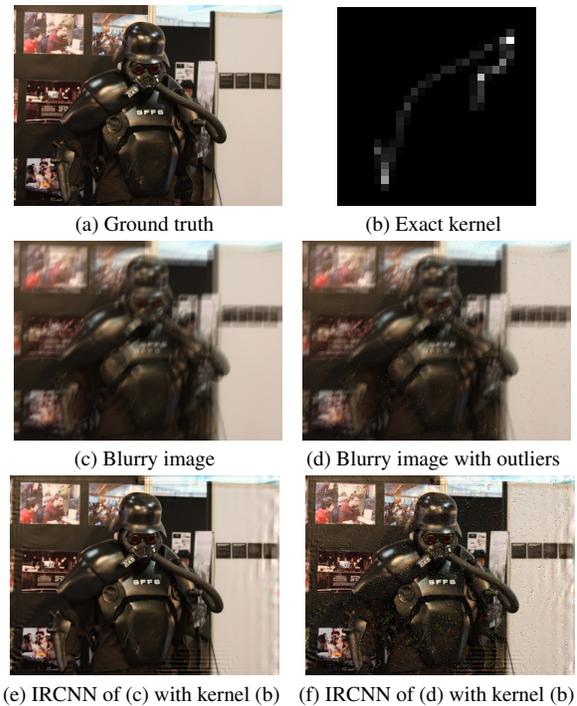


Figure 2. Sensitivity of the deblurring method to outliers in the blurry image. (a) ground truth, (b) exact kernel, (c) blurry image, (d) image (c) with outliers, (e) and (f) deblurred results of blurry image (c) and (d) using [49] with the kernel (b) respectively.

is given as the following:

$$L(\Theta) = \frac{1}{2N} \sum_{i=1}^N \|F_{n_\sigma}(y_i; \Theta) - x_i\|_F^2,$$

where  $x_i$  is the ground-truth image,  $y_i$  is the  $i$ -th corresponding noisy input image by adding Gaussian noise  $n_\sigma$  with standard variance  $\sigma$ , i.e.,  $y_i = x_i + n_\sigma$ . To recover the image from the multi-noise-level images as [49], a series of MWCNN denoisers with  $\sigma$  from high to low need to be trained in advance to solve Equation (5) in different iteration. As the proposed MWCNN-based denoisers show superiority in reserving the textures in the deblurred results, we refer to them as texture-preserving denoisers in the following presented deblurring model.

### 3.2. Non-blind deblurring model with kernel errors

To facilitate the better recovery of deblurring, we consider two sources of the ringing artifacts and avoid them by introducing two terms named by the kernel error and the residual error respectively. The proposed model, is given in the following equation:

$$\min_{x,t,s} \frac{1}{2} \|(k+t)*x - y - s\|_2^2 + \lambda \Phi(x) + \frac{\alpha}{2} \|t\|_2^2 + \beta \|s\|_1 \quad (6)$$

where  $k$  is the given estimated kernel returned by the blind deblurring method,  $t$  and  $s$  denote the kernel error and residual error terms respectively.

The kernel error term  $t$  is used to rectify the error in the given estimated kernel. As Fig. 1 shows, the kernel in (d) which is estimated from the blurry image (b) contains the noise compared to the exact kernel in (c). As we can see, the deblurred result by IRCNN [49] with the estimated kernel has ringing artifacts around the edges compared to the one recovered by the exact kernel. Therefore, the proposed term  $t$  is used to rectify this error by approaching the difference between the estimated kernel and the exact kernel. To illustrate the role of the residual error term, we synthesize two blurry images with and without outliers in Fig. 2 (c) and (d) respectively. We can see the separated outliers in (d) incur the ringing artifacts around their locations in the deblurred results due to their deviation from the linearity of the model (1). Therefore, the residual error term  $s$  in the proposed model is applied to make up the values of the outliers in  $y$  and make the compensated blurry image  $y + s$  a linear convolution as following

$$y + s = (k + t) * x. \quad (7)$$

Substituting the blurry image as  $y + s$ , can therefore avoid the ringing artifacts arisen from the non-linear values in  $y$ .

Considering the sparsity of the two difference terms and the calculation efficiency, we constrain  $t$  and  $s$  by the  $l_2$  and  $l_1$ -norms respectively, so that the fast fourier transformation and thresholding can be applied as given in Equation (16) and (17) respectively. With the two parameters  $\alpha$  and  $\beta$ , the two error terms in the proposed model can be flexibly adjusted for the specific situations of the blurry images.

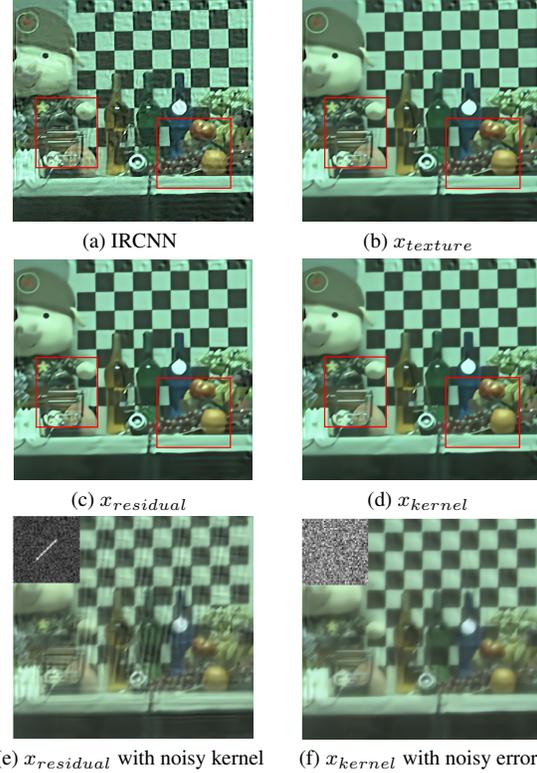


Figure 3. (a)-(d): the deblurred results of Fig.1 (b) with estimated kernel in Fig. 1 (d). (e) and (f): the deblurred results of Fig.1 (b) with synthetic noisy kernel.

### 3.3. Iterative solver with texture-preserving prior

To solve the proposed model (6), we need to split the form-free regularization from the other terms. We use the Plug-and-Play technique and adopt the half quadratic splitting (HQS) method. By introducing an extra variable  $z$ , the original model can be transformed as:

$$\min_{x,t,s} J(x, t, s) = \frac{1}{2} \|(k + t) * x - y - s\|_2^2 + \lambda \Phi(z) + \frac{\alpha}{2} \|t\|_2^2 + \beta \|s\|_1, \text{ subject to } z = x. \quad (8)$$

By introducing an additional quadratic penalty term, HQS transforms Equation (8) into the unconstrained function as:

$$L_\mu(x, t, s, z) = \frac{1}{2} \|(k + t) * x - y - s\|_2^2 + \lambda \Phi(z) + \frac{\alpha}{2} \|t\|_2^2 + \beta \|s\|_1 + \frac{\mu}{2} \|z - x\|_2^2, \quad (9)$$

where  $\mu$  is a penalty parameter which varies in a non-descending order. The solution to the model (6) is then approached by iteratively solving the four variables  $x, t, s, z$

presented in the function (9) as following:

$$\min_x \|(k+t^k) * x - y - s^k\|_2^2 + \mu^k \|z^k - x\|_2^2, \quad (10)$$

$$\min_z \frac{1}{2(\sqrt{\lambda/\mu^k})^2} \|z - x^{k+1}\| + \Phi(z), \quad (11)$$

$$\min_t \frac{1}{2} \|(k+t) * x^{k+1} - y - s^k\|_2^2 + \frac{\alpha}{2} \|t\|_2^2, \quad (12)$$

$$\min_s \frac{1}{2} \|(k+t^{k+1}) * x^{k+1} - y - s\|_2^2 + \beta \|s\|_1. \quad (13)$$

To solve the  $x$ -subproblem (10) where the data fitting term is associated with a quadratic regularized least-squares term, we update the image  $x$  efficiently by using the fast Fourier transform (FFT) in the frequency domain:

$$x^{k+1} = \mathcal{F}^{-1} \left( \frac{\overline{\mathcal{F}(k+t^k)} \mathcal{F}(y+s^k) + \mu^k \mathcal{F}(z^k)}{|\mathcal{F}(k+t^k)|^2 + \mu^k} \right), \quad (14)$$

where  $\overline{\cdot}$  is the complex conjugate operator,  $\mathcal{F}(\cdot)$  and  $\mathcal{F}^{-1}(\cdot)$  denote Fourier transform and its inverse transform.

For Equation (11) where the implicit prior  $\Phi(z)$  prefers the clear images, the solution can be interpreted as the denoised result for the image  $x^{k+1}$  with the denoising level defined by the parameter  $\sqrt{\lambda/\mu^k}$ . Following the Plug-and-Play scheme, the existing denoisers can be applied as the optimization to solve this equation. Following [49], we denote the solver for (11) in the following form:

$$z^{k+1} = \text{Denoiser}(x^{k+1}, \sqrt{\lambda/\mu^k}), \quad (15)$$

where the larger value of  $\sqrt{\lambda/\mu^k}$  implies the image  $x^{k+1}$  requires a denoiser of higher noise level to get the solution  $z_{k+1}$ . Considering the great advancement in the denoising methods, we trained our Gaussian denoisers based on the MWCNN model. Since it has proven to be effective in recovering the sharp structures from the noisy images, we referred to the implicit prior  $\Phi(\cdot)$  in the proposed model as the texture reserving prior.

Next, to solve the kernel error  $t^{k+1}$  in Equation (12), we also derive the solution by FFT. The closed-form solution is

$$t^{k+1} = \mathcal{F}^{-1} \left( \frac{\overline{\mathcal{F}(x^{k+1})} \mathcal{F}(y+s^k - k * x^{k+1})}{|\mathcal{F}(x^{k+1})|^2 + \alpha} \right). \quad (16)$$

Finally, the subproblem (13) containing the  $l_1$ -norm can be solved by the soft shrinkage operator. The solution of  $s^{k+1}$  is given component-wisely as

$$s_i^{k+1} = \text{sgn}(v_i) * \max(|v_i| - \beta, 0), \quad (17)$$

where  $v_i = [(k+t^{k+1}) * x^{k+1} - y]_i$ .



Figure 4. The deblurred results of (a) IRCNN (b) the proposed method only considering the texture reserving prior (c) the proposed model only considering only the kernel error term and (d) the proposed model considering only the residual error term.

Table 1. Average PSNR(dB)/SSIM values on 4 images from Levin’s dataset. Note that the value on each image is an average result of 8 test blurry images synthesized by 8 different kernels. The models are compared using the estimated kernels returned by [36].

Levin <i>et al.</i>	Image 1	Image 2	Image 3	Image 4	Average
[49]	29.70/0.8761	31.50/0.8913	33.14/0.9263	28.69/0.8554	30.76/0.8873
$x_{texture}$	31.87/0.9208	33.62/0.9448	34.50/0.9433	29.85/0.8790	32.46/0.9220
Proposed model (6)	<b>34.46/0.9362</b>	<b>35.13/0.9489</b>	<b>37.54/0.9651</b>	<b>32.90/0.9201</b>	<b>35.01/0.9426</b>

## 4. Experiments

In this section, we first illustrate the experimental details (Sec. 4.1.) and the effectiveness of each proposed modules (Sec. 4.2.), after which we present the comparison of the proposed model to other state-of-the-art deblurring methods in both synthetic images (Sec. 4.3.) and real-world images (Sec. 4.4.). To achieve valuable insights from the comparison, the chosen methods include both variational models and learning based methods.

Table 2. Average PSNR(dB)/SSIM of the non-blind deblurring results on Levin’s dataset [18] with noise  $\sigma = 1\%$  using different estimated blurring kernels by [6], [19], [36], [43] and [24].

Kernel	Non-Blind Methods							
	[12]	[14]	[52]	[33]	[49]	[26]	[22]	Ours
[6]	27.81/0.83	28.12/0.82	29.00/0.86	28.52/0.84	28.12/0.83	28.15/0.81	28.32/0.82	<b>29.86/0.87</b>
[19]	27.96/0.84	28.17/0.83	29.06/0.87	28.57/0.85	28.51/0.84	28.29/0.83	28.41/0.83	<b>29.54/0.87</b>
[36]	29.20/0.86	29.84/0.86	31.68/0.91	30.53/0.88	30.63/0.89	30.49/0.87	30.32/0.87	<b>31.98/0.91</b>
[43]	28.86/0.85	29.30/0.85	30.71/0.89	29.79/0.86	29.73/0.86	29.82/0.85	29.71/0.85	<b>31.05/0.89</b>
[24]	29.10/0.86	29.67/0.86	31.49/0.91	30.28/0.88	30.22/0.88	30.60/0.89	29.15/0.82	<b>31.68/0.91</b>

### 4.1. Training details and parameter setting

The denoisers for the subproblem (11) are trained with the MWCNN network. The training data consists of 200 images from Berkeley Segmentation Dataset (BSD) [28], 800 images from DIV2K [1], and 4744 images from Wa-

Table 3. Average PSNR(dB)/SSIM of the comparison on Levin’s dataset with the the estimated blurring kernels returned by [36] under different noise levels.

$\sigma$	[14]	[33]	[49]	[26]	[22]	Ours
noise-free	31.06/0.90	32.22/0.92	30.76/0.89	33.22/0.94	32.98/0.93	<b>35.01/0.94</b>
0.59%	30.21/0.89	31.29/0.91	30.27/0.88	30.33/0.88	31.40/0.88	<b>32.59/0.92</b>
1%	30.08/0.84	30.53/0.88	30.42/0.86	30.49/0.87	30.32/0.87	<b>31.98/0.91</b>
2%	28.56/0.80	27.80/0.76	30.06/0.88	29.09/0.84	27.79/0.78	<b>30.66/0.89</b>

terloo Exploration Database (WED) [27] and are further cropped into patches of size  $240 \times 240$ . We set the batch size as 24 and  $24 \times 9000$  small patches are used in each epoch of the training process. We trained the MWCNN model to learn a mapping from synthesized images with a specific level of Gaussian noise to the clean images. Following the settings in MWCNN models, the Adam solver is used during the training with parameters  $\alpha = 0.010$ ,  $\beta_1 = 0.900$ ,  $\beta_2 = 0.999$  and  $\epsilon = 10^{-8}$ . The learning rate is decreased from  $10^{-4}$  to  $10^{-5}$ . Data augmentation of rotation and flip are used in mini-batch learning. Since a series of models with different noise levels from 2 to 50 with step size 2 are needed, each model is trained by 10 epochs for early stop. The models are trained in Matlab R2018b under Ubuntu 18.04 with an NVIDIA GeForce GTX1080Ti. The overall training procedure is illustrated in Fig. 5. Albeit this training procedures for the denoisers bring about the extra time and memory consumption in the preparation stage, they are fixed in the deblurring algorithm. Compared to the end-to-end neural networks, this plug-and-play deblurring method keeps the interpretability and flexibility to different images.

In the non-blind deblurring scheme, we use the pre-trained texture reserving denoisers to optimize (15). For the selection of the denoiser in each iteration, we basically adopt the strategy of multi-noise-levels formulated in [49]. In the proposed deblurring model, there are four parameters, i.e.,  $\mu_k$ ,  $\lambda$ ,  $\alpha$  and  $\beta$ . Among them,  $\sqrt{\lambda/\mu^k}$  controls the noise level  $\sigma_\eta$  of the texture reserving denoiser and changes for every iteration. From (11), we set  $\sqrt{\lambda/\mu^k}$  decayed linearly and exponentially from noise level  $\sigma_\eta = 50$  to a value  $\sigma_\eta$  in the interval [2, 10]. During our experiments, empirical evidence showed that the proposed numerical scheme with the texture reserving denoisers reached the solution with a pleasant visual quality by 15 iterations. Therefore, we fix the number of iterations as 15, which is less than 30 iterations used by IRCNN [49]. Since the introduced error terms in the proposed method are solved efficiently with the closed-form solutions, we say the proposed method brings about the improvement without the loss of efficiency. The parameters  $\alpha$  and  $\beta$  flexibly adjust the role of the kernel error term and the residual error term. While the kernel is not well estimated, the value of  $\alpha$  should be lessened. While there are noise and outliers in the blurry images, the value of  $\beta$  should be lessened. The sensitivity to different values of  $\alpha$  and  $\beta$  are given in Fig. 6. Generally, they can be

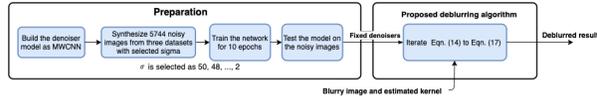


Figure 5. Flowchart of the training and testing procedure.

chosen from a small range to deal with different case, i.e.,  $\alpha \in \{1, 10, 10^2, 10^3\}$  and  $\beta \in \{0.01, 0.1\}$ .

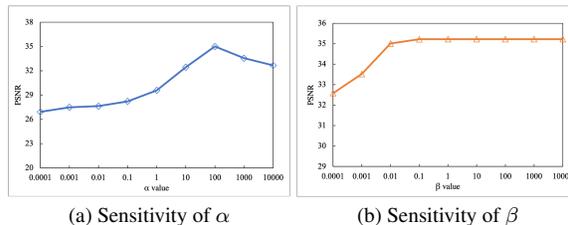


Figure 6. Average PSNR(dB) values on Levin’s dataset with different  $\alpha$  and  $\beta$ .

## 4.2. Ablation study

In this section, an ablation study is firstly presented to verify the effectiveness of proposed kernel error term and residual error term as well as the proposed texture reserving prior introduced in the proposed model by comparing the following three models:

$$\begin{aligned}
 x_{texture} &= \arg \min_x \frac{1}{2} \|k * x - y\|_2^2 + \lambda \Phi(x), \\
 x_{kernel} &= \arg \min_{x,t} \frac{1}{2} \|(k+t) * x - y\|_2^2 + \lambda \Phi(x) + \frac{\alpha}{2} \|t\|_2^2, \\
 x_{residual} &= \arg \min_{x,s} \frac{1}{2} \|k * x - y - s\|_2^2 + \lambda \Phi(x) + \beta \|s\|_1,
 \end{aligned}$$

where  $x_{texture}$  is the result of the proposed model only considering the texture-preserving prior,  $x_{kernel}$  is the result of the proposed model only with the kernel error term while  $x_{residual}$  denotes the one only with residual error term.

In Fig. 1 (b), we synthesize a blurred image with a moderate noise level ( $\sigma = 2\%$ ) and obtain a noisy kernel as illustrated in Fig. 1 (d). The results of IRCNN and  $x_{texture, kernel, residual}$  are illustrated in Fig. 3 (b), (c) and (d). As one can see, the solution  $x_{texture}$  is better than the one of IRCNN [49], which indicates the texture reserving prior is better than the denoiser proposed in IRCNN. However, the result of  $x_{texture}$  still reserves the ringing artifacts while  $x_{kernel}$  further corrects the artifacts and improve the deblurred quality by considering the noise in the given kernel. Additionally, we synthesize a noisy kernel by adding Gaussian noise to the exact kernel to justify the effectiveness of the kernel error term. The result of  $x_{texture}$  and  $x_{kernel}$  with the noisy kernel error are given in (e) and (f) respectively.

Table 4. Average PSNR(dB)/SSIM of the comparison on Sun’s dataset with estimated blurring kernels returned by [45] and Lai’s dataset with estimated blurring kernels returned by [37]

Dataset	[14]	[33]	[49]	[12]	[26]	[22]	Ours
Sun	29.00/0.82	29.76/0.84	29.56/0.84	27.85/0.76	30.01/0.85	29.20/0.80	<b>30.41/0.85</b>
Lai	18.73/0.50	19.50/0.59	19.11/0.59	19.22/0.55	19.31/0.56	19.20/0.57	<b>19.87/0.60</b>

For the proposed residual error term, we deblur the blurry image with saturated pixels shown in Fig. 2 (d) with the estimated kernel returned by [36]. The results are shown in Fig. 4. Clearly, the solution  $x_{residual}$  recovered the location of the outliers in the deblurred result while the solution  $x_{kernel}$  still suffers from the noise and ringing artifacts.

From the above examples, we can see the necessity of the two proposed error terms as well as the texture reserving prior for alleviating ringing artifacts arisen from different causes. By tuning parameters, the proposed method can be flexibly adjusted according to the situation of the degraded image. We further compare the proposed texture reserving result  $x_{texture}$  with the one of IRCNN on the Levin’s dataset. The numerical results are reported in Tab. 1. One can see, the texture reserving prior enhances the traditional IRCNN model by a wide margin and the proposed model reaches the best results with the introduction of the two error terms. Hence, the proposed model is both flexible and effective.

### 4.3. Synthetic Dataset

In this section, a thorough evaluation will first be given in the benchmark of Levin *et al.* [18], which contains 32 grayscale blurry images. Secondly, we evaluate the methods on the larger dataset of Sun *et al.* [43], which has 640 blurry images generated from 80 clear images and the same 8 blur kernels. Furthermore, we test the proposed method on a more challenging Lai *et al.*’s dataset [17], which contains 4 large ground truth kernels and 25 color images of different types. We synthesize the blurry images by applying valid convolution with a specific level of noise added to it. The edge taper boundary extension in [41] is also adopted. As for the numerical metrics, we adopt the method in [19] by aligning the deblurred images with the sharp images and cutting off the boundary pixels before calculating the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM), which are also adopted in the measurements by other studies on deblurring.

#### 4.3.1 Levin *et al.*’s dataset

To verify the robustness of the proposed method to the estimated kernels, we test the non-blind deblurring methods with given kernels returned by 5 estimation methods, [6], [19], [36], [43] and [24]. With kernels of each method, our proposed method is compared with other competitive non-blind deblurring methods including [12], [14], [52], [33],

[49], [22], [26] respectively. Note that, for the compared methods, we run their public codes and models on the test sets with the parameters that authors suggest. Tab. 2 shows the average PSNR and SSIM values of different methods in terms of 5 different kernel estimation methods on the Levin *et al.*’s dataset with 1% noise. It can be seen that the proposed model outperforms all the competing methods by 0.4-2dB in PSNR, which implies the proposed model is robust and effective with different kernel estimations. The visual comparison listed in the first row of Fig. 7 shows the proposed method reserves more details than any other methods.

To illustrate the robustness of our model to different noise cases, Tab. 3 presents the comparisons with other methods in [14], [33], [49], [26], [22] on Levin’s dataset with different levels of noise. In Tab. 3, our proposed model achieves the best performance compared to other methods in all cases. It definitively demonstrates that the proposed method is more robust to the noisy blurry images.

#### 4.3.2 Sun *et al.*’s dataset and Lai *et al.*’s dataset

For additional comparison, we evaluate our model on Sun *et al.*’s and Lai *et al.*’s datasets added with 1% noise, with the numerical results listed in Tab. 4. In Sun *et al.*’s dataset, the average PSNR of our model outperforms the competing methods by about 0.4-2.5dB. It is worth mentioning that our numerical measurements are also about 0.4 dB higher than the latest end-to-end learning method [26]. Indeed, the examples listed in Fig. 7 also prove that the visual quality of the recovered images is better than those of model-based and end-to-end non-blind network methods. For the images in Lai *et al.*’s dataset, it is challenging for all methods to get fine results, since the kernel is large and more areas get influenced from the boundary. Still, the proposed method surpasses all the methods and is competitive to the learning-based method [26]. As seen from Fig. 7, for images recovered successfully by all methods, the result of the proposed model has the most textures with few ringing artifacts while Li’s methods, though having the fewest ringing artifacts, still reserves the blur in the results.

### 4.4. Real-world image

In this section, we focus on the visual comparison of the deblurred results of the blurry images from the real world due to the lack of ground truth. Two examples are reported in Fig. 8. For the first example with complex textures, ours recovers the images with most realistic colors without smoothing crucial textures such as the fine lines and characters while other methods recover the sharp edges with ringings damaging the surrounding colors. For the characters in the second example, which are easily influenced by the kernel inaccuracy, ours recovers them with least ringings which proves it can reduce the error in the estimated kernel.

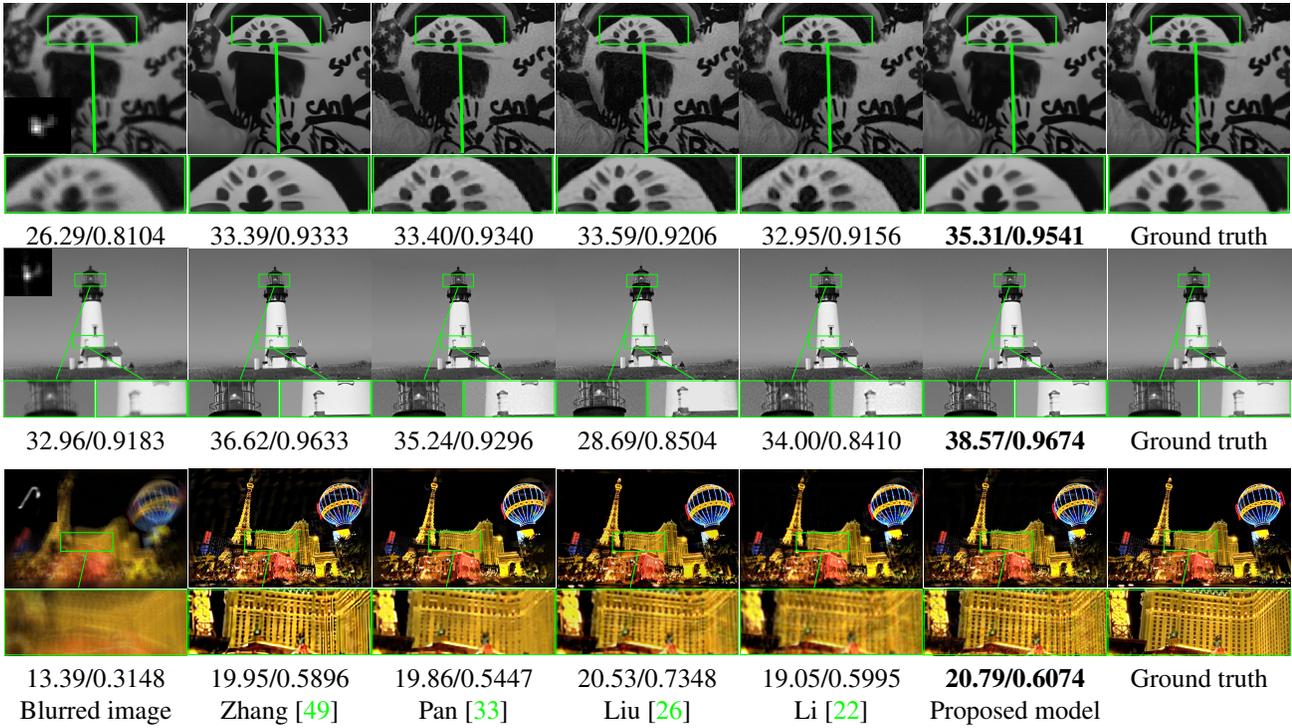


Figure 7. The 1st columns are the blurry images with 1% noise from Levin’s dataset (1st row) with the kernels by [36], Sun’s dataset (2nd row) with kernels by [45] and Lai’s dataset (3rd row) with the kernels returned by [17]. From the 2rd to the last columns, there are the results of [49], [33], [26], [22], ours and the ground truth images for comparison. The images are best compared by zooming in.

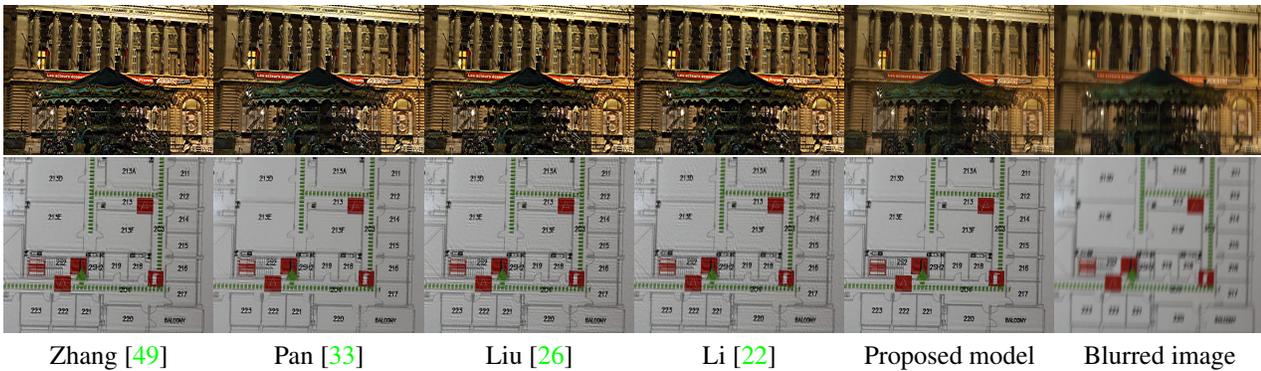


Figure 8. The comparison of the deblurred results of blurry images from the real dataset in Lai’s dataset with the kernel estimated by [24].

## 5. Conclusion

The task of non-blind image deblurring is to recover a clear image from a blurred image with a certain estimated kernel. In many situations, this kernel is inaccurate. In order to handle this issue, we propose to rectify the known kernel during deconvolution, allowing kernel error. Moreover, a residual error term was used to deal with the non-linearity caused by outlier and noise. Furthermore, a deep learning denoiser was applied to keep the fine textures in the recovered image, which can be readily improved. Extensive results show that the proposed model is robust in dealing with

a wide range of images and kernels and bring about the improvements in both quantitative metrics and visual quality.

## Acknowledgement.

This work was supported in part by the National Key R&D Program of China under Grant 2021YFE0203700, Grant NSFCRGC N\_CUHK 41519, Grant ITF MHP03820, Grant RGC 14300219, 14302920, 14301121, Grant CRF 8730063, and CUHK Direct Grant for Research under Grant 4053405, 4053460.

## References

- [1] Eirikur Agustsson and Radu Timofte. NTIRE 2017 challenge on single image super-resolution: Dataset and study. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 126–135, 2017. **5**
- [2] Ali N Akansu, Paul A Haddad, Richard A Haddad, and Paul R Haddad. *Multiresolution signal decomposition: transforms, subbands, and wavelets*. Academic press, 2001. **3**
- [3] Raied Aljadaany, Dipan K Pal, and Marios Savvides. Douglas-rachford networks: Learning both the image prior and data fidelity terms for blind image deconvolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 10235–10244, 2019. **1**
- [4] L. Chen, F. Fang, T. Wang, and G. Zhang. Blind image deblurring with local maximum gradient prior. *IEEE Conf. Comput. Vis. Pattern Recog.*, pages 1742–1750, 2019. **2**
- [5] Yunjin Chen and Thomas Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1256–1272, 2016. **1**
- [6] Sunghyun Cho and Seungyong Lee. Fast motion deblurring. In *ACM SIGGRAPH Asia 2009 papers*, pages 1–8. 2009. **5, 7**
- [7] Aram Danielyan, Vladimir Katkovnik, and Karen Egiazarian. BM3D frames and variational image deblurring. *IEEE Transactions on Image Processing*, 21(4):1715–1728, 2011. **2**
- [8] Ingrid Daubechies. *Ten lectures on wavelets*. SIAM, 1992. **3**
- [9] Jiangxin Dong, Jinshan Pan, Zhixun Su, and Ming-Hsuan Yang. Blind image deblurring with outlier handling. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2478–2486, 2017. **2**
- [10] Dong Gong, Zhen Zhang, Qinfeng Shi, Anton van den Hengel, Chunhua Shen, and Yanning Zhang. Learning deep gradient descent optimization for image deconvolution. *IEEE Transactions on Neural Networks and Learning Systems*, 2020. **2**
- [11] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5197–5206, 2015. **1**
- [12] Hui Ji and Kang Wang. Robust image deblurring with an inaccurate blur kernel. *IEEE Transactions on Image Processing*, 21(4):1624–1634, 2011. **2, 5, 7**
- [13] Mingyuan Jiu and Nelly Pustelnik. A deep primal-dual proximal network for image restoration. *IEEE Journal of Selected Topics in Signal Processing*, 15(2):190–203, 2021. **1**
- [14] Dilip Krishnan and Rob Fergus. Fast image deconvolution using hyper-laplacian priors. In *Advances in Neural Information Processing Systems*, pages 1033–1041, 2009. **5, 6, 7**
- [15] Jakob Kruse, Carsten Rother, and Uwe Schmidt. Learning to push the limits of efficient FFT-based image deconvolution. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4586–4594, 2017. **2**
- [16] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8183–8192, 2018. **1**
- [17] Wei-Sheng Lai, Jia-Bin Huang, Zhe Hu, Narendra Ahuja, and Ming-Hsuan Yang. A comparative study for single image blind deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1709, 2016. **7, 8**
- [18] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Understanding and evaluating blind deconvolution algorithms. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1964–1971, 2009. **5, 7**
- [19] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2657–2664, 2011. **5, 7**
- [20] Juncheng Li, Faming Fang, Kangfu Mei, and Guixu Zhang. Multi-scale residual network for image super-resolution. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 517–532, 2018. **1**
- [21] Lerenhan Li, Jinshan Pan, Wei-Sheng Lai, Changxin Gao, Nong Sang, and Ming-Hsuan Yang. Learning a discriminative prior for blind image deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6616–6625, 2018. **1**
- [22] Sanqian Li, Binjie Qin, Jing Xiao, Qiegen Liu, Yuhao Wang, and Dong Liang. Multi-channel and multi-model-based autoencoding prior for grayscale image restoration. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*, 29:142–156, 2019. **5, 6, 7, 8**
- [23] J. Liu, M. Yan, and T. Zeng. Surface-aware blind image deblurring. *IEEE Trans. Pattern Anal. Mach. Intell.*, 43(3):1041–1055, 2021. **2**
- [24] Jun Liu, Ming Yan, and Tiejong Zeng. Surface-aware blind image deblurring. *IEEE transactions on pattern analysis and machine intelligence*, 43(3):1041–1055, 2021. **2, 5, 7, 8**
- [25] Pengju Liu, Hongzhi Zhang, Kai Zhang, Liang Lin, and Wangmeng Zuo. Multi-level wavelet-cnn for image restoration. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2018. **3**
- [26] Risheng Liu, Long Ma, Yiyang Wang, and Lei Zhang. Learning converged propagations with deep prior ensemble for image enhancement. *IEEE Transactions on Image Processing*, 28(3):1528–1543, 2018. **5, 6, 7, 8**
- [27] Kede Ma, Zhengfang Duanmu, Qingbo Wu, Zhou Wang, Hongwei Yong, Hongliang Li, and Lei Zhang. Waterloo exploration database: New challenges for image quality assessment models. *IEEE Transactions on Image Processing*, 26(2):1004–1016, 2016. **6**
- [28] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images

- and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, volume 2, pages 416–423. IEEE, 2001. 5
- [29] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3883–3891, 2017. 1
- [30] Yuesong Nan and Hui Ji. Deep learning for handling kernel/model uncertainty in image deconvolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2388–2397, 2020. 2
- [31] Thekke Madam Nimisha, Akash Kumar Singh, and Ambasamudram N Rajagopalan. Blur-invariant deep learning for blind-deblurring. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4752–4760, 2017. 1
- [32] Jinshan Pan, Jiangxin Dong, Yang Liu, Jiawei Zhang, Jimmy Ren, Jinhui Tang, Yu-Wing Tai, and Ming-Hsuan Yang. Physics-based generative adversarial models for image restoration and beyond. *IEEE transactions on pattern analysis and machine intelligence*, 43(7):2449–2462, 2020. 1
- [33] Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang.  $l_0$ -regularized intensity and gradient prior for deblurring text images and beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(2):342–355, 2016. 5, 6, 7, 8
- [34] J. Pan, Z. Hu, Z. Su, and M. H. Yang.  $L_0$ -regularized intensity and gradient prior for deblurring text images and beyond. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(2):342–355, 2017. 2
- [35] J. Pan, D. Sun, H. Pfister, and M. H. Yang. Blind image deblurring using dark channel prior. *IEEE Conf. Comput. Vis. Pattern Recog.*, page 1628–1636, 2016. 2
- [36] Jinshan Pan, Deqing Sun, Hanspeter Pfister, and Ming-Hsuan Yang. Blind image deblurring using dark channel prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1628–1636, 2016. 2, 3, 5, 6, 7, 8
- [37] Daniele Perrone and Paolo Favaro. Total variation blind deconvolution: The devil is in the details. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2909–2916, 2014. 7
- [38] Dongwei Ren, Kai Zhang, Qilong Wang, Qinghua Hu, and Wangmeng Zuo. Neural blind deconvolution using deep priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3341–3350, 2020. 1
- [39] Arie Rond, Raja Giryes, and Michael Elad. Poisson inverse problems by the plug-and-play scheme. *Journal of Visual Communication and Image Representation*, 41:96–108, 2016. 2
- [40] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer, 2015. 3
- [41] Uwe Schmidt and Stefan Roth. Shrinkage fields for effective image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2774–2781, 2014. 7
- [42] Christian J Schuler, Michael Hirsch, Stefan Harmeling, and Bernhard Schölkopf. Learning to deblur. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(7):1439–1451, 2015. 1
- [43] Libin Sun, Sunghyun Cho, Jue Wang, and James Hays. Edge-based blur kernel estimation using patch priors. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–8. IEEE, 2013. 5, 7
- [44] Tom Tirer and Raja Giryes. Image restoration by iterative denoising and backward projections. *IEEE Transactions on Image Processing*, 28(3):1220–1234, 2018. 2
- [45] Li Xu and Jiaya Jia. Two-phase kernel estimation for robust motion deblurring. In *European Conference on Computer Vision*, pages 157–170. Springer, 2010. 7, 8
- [46] Li Xu, Jimmy SJ Ren, Ce Liu, and Jiaya Jia. Deep convolutional neural network for image deconvolution. In *Advances in Neural Information Processing Systems*, pages 1790–1798, 2014. 1
- [47] Yanyang Yan, Wenqi Ren, Yuanfang Guo, Rui Wang, and Xiaochun Cao. Image deblurring via extreme channels prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4003–4011, 2017. 2
- [48] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017. 1, 3
- [49] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep CNN denoiser prior for image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3929–3938, 2017. 1, 2, 3, 4, 5, 6, 7, 8
- [50] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Transactions on Image Processing*, 27(9):4608–4622, 2018. 3
- [51] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Deep plug-and-play super-resolution for arbitrary blur kernels. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1671–1681, 2019. 2
- [52] Daniel Zoran and Yair Weiss. From learning models of natural image patches to whole image restoration. In *2011 International Conference on Computer Vision*, pages 479–486. IEEE, 2011. 1, 2, 5, 7