

Learning a Discriminative Prior for Blind Image Deblurring

Lerenhan Li^{1,2}

Jinshan Pan³

Wei-Sheng Lai²

Changxin Gao¹

Nong Sang^{1*}

Ming-Hsuan Yang²

¹National Key Laboratory of Science and Technology on Multispectral Information Processing,
School of Automation, Huazhong University of Science and Technology

²Electrical Engineering and Computer Science, University of California, Merced

³School of Computer Science and Engineering, Nanjing University of Science and Technology

Abstract

We present an effective blind image deblurring method based on a data-driven discriminative prior. Our work is motivated by the fact that a good image prior should favor clear images over blurred ones. In this work, we formulate the image prior as a binary classifier which can be achieved by a deep convolutional neural network (CNN). The learned prior is able to distinguish whether an input image is clear or not. Embedded into the maximum a posterior (MAP) framework, it helps blind deblurring in various scenarios, including natural, face, text, and low-illumination images. However, it is difficult to optimize the deblurring method with the learned image prior as it involves a non-linear CNN. Therefore, we develop an efficient numerical approach based on the half-quadratic splitting method and gradient decent algorithm to solve the proposed model. Furthermore, the proposed model can be easily extended to non-uniform deblurring. Both qualitative and quantitative experimental results show that our method performs favorably against state-of-the-art algorithms as well as domain-specific image deblurring approaches.

1. Introduction

Blind image deblurring is a classical problem in image processing and computer vision, which aims to recover a latent image from a blurred input. When the blur is spatially invariant, the blur process is usually modeled by

$$B = I \otimes k + n, \quad (1)$$

where \otimes denotes convolution operator, B, I, k and n denote the blurred image, latent sharp image, blur kernel, and noise, respectively. The problem (1) is ill-posed as both I and k are unknown, and there exist infinite solutions. To

*Corresponding author.

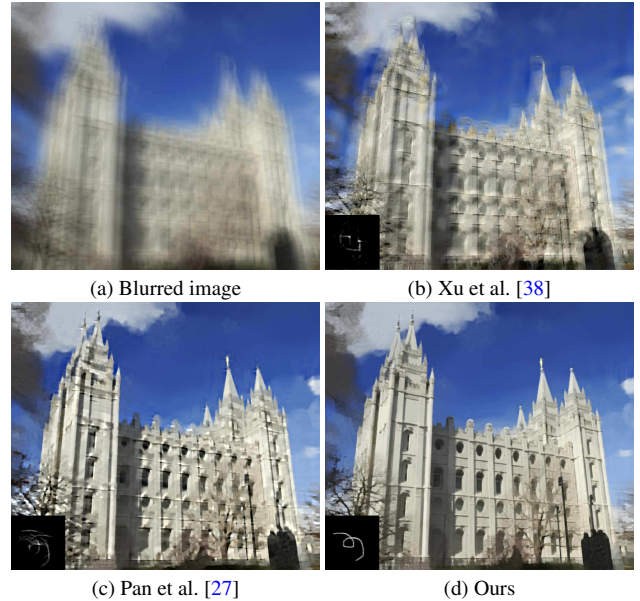


Figure 1. A deblurred example. We propose a discriminative image prior which is learned from a deep binary classification network for image deblurring. For the blurred image B in (a) and its corresponding clear image I , we can get $\frac{\|\nabla I\|_0}{\|\nabla B\|_0} = 0.85$, $\frac{\|D(I)\|_0}{\|D(B)\|_0} = 0.82$ and $\frac{f(I)}{f(B)} = 0.03$, where ∇ , $D(\cdot)$, $\|\cdot\|_0$ and $f(\cdot)$ denote the gradient operator [38], the dark channel [27], L_0 norm [27, 38] and our proposed classifier, respectively. The prior is more discriminative than the hand-crafted priors, thus leading to better deblurred results. (A larger ratio indicates that the prior responses are closer and cannot be well separated.)

tackle this problem, additional constraints and prior knowledge on both blur kernels and images are required.

The main success of the recent deblurring methods mainly comes from the development of effective image priors and edge-prediction strategies. However, the edge-prediction based methods usually involve a heuristic edge selection step, which do not perform well when strong edges are not available. To avoid the heuristic edge selection step, numerous algorithms based on natural image pri-

ors have been proposed, including normalized sparsity [16], L_0 gradients [38] and dark channel prior [27]. These algorithms perform well on generic natural images but do not generalize well to specific scenarios, such as text [26], face [25] and low-illumination images [11]. Most of the aforementioned image priors have a similar effect that they favor clear images over blurred images, and this property contributes to the success of the MAP-based methods for blind image deblurring. However, most priors are hand-crafted and mainly based on limited observations of specific image statistics. These algorithms cannot be generalized well to handle various scenarios in the wild. Thus, it is of great interest to develop a general image prior which is able to deal with different scenarios with the MAP framework.

To this end, we formulate the image prior as a binary classifier which is able to distinguish clear images from blurred ones. Specifically, we first train a deep CNN to classify blurred (labeled as 1) and clear (labeled as 0) images. To handle arbitrary image sizes in the coarse-to-fine MAP framework, we adopt a global average pooling layer [21] in the CNN. In addition, we use a multi-scale training strategy to make the classifier more robust to different input image sizes. We then take the learned CNN classifier as a regularization term w.r.t. latent images in the MAP framework. Figure 1 shows an example that the proposed image prior is more discriminative (i.e., has a lower ratio between the response of blurred and clear images) than the state-of-the-art hand-crafted prior [27].

While the intuition behind the proposed method is straightforward, in practice it is difficult to optimize the deblurring method with the learned image prior as a non-linear CNN is involved. Therefore, we develop an efficient numerical algorithm based on the half-quadratic splitting method and gradient decent approach. The proposed algorithm converges quickly in practice and can be applied to different scenarios as well as non-uniform deblurring.

The contributions of this work are as follows:

- We propose an effective discriminative image prior which can be learned by a deep CNN classifier for blind image deblurring. To ensure that the proposed prior (i.e., classifier) can handle the image of different sizes, we use the global average pooling and multi-scale training strategy to train the proposed CNN.
- We use the learned classifier as a regularization term of the latent image in the MAP framework and develop an efficient optimization algorithm to solve the deblurring model.
- We demonstrate that the proposed algorithm performs favorably against the state-of-the-art methods on both the widely-used natural image deblurring benchmarks and domain-specific deblurring tasks.
- We show the proposed method can be directly generalized to the non-uniform deblurring.

2. Related Work

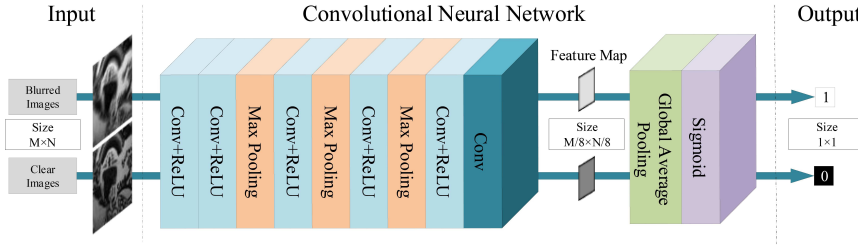
Recent years have witnessed significant advances in single image deblurring. We focus our discussion on recent optimization-based and learning-based methods.

Optimization-based methods. State-of-the-art optimization based approaches can be categorized to implicit and explicit edge enhancement methods. The implicit edge enhancement approaches focus on developing effective image priors to favor clear images over blurred ones. Representative image priors include sparse gradients [7, 19, 36], normalized sparsity [16], color-line [17], L_0 gradients [38], patch priors [32], and self-similarity [24].

Although these image priors are effective for deblurring natural images, they are not able to handle specific types of input such as text, face and low-illumination images. The statistics of these domain-specific images are quite different from natural images. Thus, Pan et al. [26] propose the L_0 -regularized prior on both image intensity and gradients for deblurring text images. Hu et al. [11] detect the light streaks in extremely low-light images for estimating blur kernels. Recently, Pan et al. [27] propose a dark channel prior for deblurring natural images, which can be applied to face, text and low-illumination images as well. However, the dark channel prior is less effective when there is no dark pixel in the image. Yan et al. [40] further propose to incorporate a bright channel prior with the dark channel prior to improve the robustness of the deblurring algorithm.

While those algorithms demonstrate state-of-the-art performance, most priors are hand-crafted and designed under limited observation. In this work, we propose to learn a data-driven discriminative prior using a deep CNN. Our prior is designed from a simple criterion without any specific assumption: the prior should favor clear images over blurred images under various of scenarios.

Learning-based methods. With the success of deep CNNs on high-level vision problems [8, 22], several approaches have adopted deep CNNs in image restoration problems, including super-resolution [6, 14, 18], denoising [23] and JPEG deblocking [5]. Hradiš et al. [10] propose an end-to-end CNN to deblur text images. Following the MAP-based deblurring methods, Schuler et al. [29] train a deep network to estimate the blur kernel and then adopt a conventional non-blind deconvolution approach to recover the latent sharp image. Sun et al. [31] and Yan and Shao [39] parameterize the blur kernels and learn to estimate them via classification and regression, respectively. Several approaches train deep CNNs as an image prior or denoiser for non-blind deconvolution [30, 42, 41], which cannot be directly applied in blind deconvolution. Recently, Chakrabarti [3] trains a deep network to predict the Fourier coefficients of a deconvolution filter. Neverthe-



(a) Network architecture

| Layers | Filter size | Stride | Padding |
|--------|----------------------------------|--------|---------|
| CR1 | $3 \times 3 \times 1 \times 64$ | 1 | 1 |
| CR2 | $3 \times 3 \times 64 \times 64$ | 1 | 1 |
| M3 | 2×2 | 2 | 0 |
| CR4 | $3 \times 3 \times 64 \times 64$ | 1 | 1 |
| M5 | 2×2 | 2 | 0 |
| CR6 | $3 \times 3 \times 64 \times 64$ | 1 | 1 |
| M7 | 2×2 | 2 | 0 |
| CR8 | $3 \times 3 \times 64 \times 64$ | 1 | 1 |
| C9 | $3 \times 3 \times 64 \times 1$ | 1 | 1 |
| G10 | $(M/8) \times (N/8)$ | 1 | 0 |
| S11 | - | - | - |

(b) Network parameters

Figure 2. Architecture and parameters of the proposed binary classification network. We adopt a global average pooling layer instead of a fully-connected layer to handle different sizes of input images. CR denotes the convolutional layer followed by a ReLU non-linear function, M denotes the max-pooling layer, C denotes the convolutional layer, G denotes the global average pooling layer and S denotes the sigmoid non-linear function.

less, the performance of deep CNNs on blind image deblurring [1, 28] still falls behind conventional optimization-based approaches on handling large blur kernels. In our work, we take advantage of both conventional MAP-based framework and the discriminative ability of deep CNNs. We embed the learned CNN prior into the coarse-to-fine MAP framework for solving the blind image deblurring problem.

3. Learning a Data-Driven Image Prior

In this section, we describe the motivation of developing the proposed image prior, network design, loss function, and implementation details of our binary classifier.

3.1. Motivation

The MAP-based blind image deblurring methods typically solve the following problem:

$$\min_{I, k} \|I \otimes k - B\|_2^2 + \gamma \|k\|_2^2 + p(I). \quad (2)$$

The key to the success of this framework lies on the latent image prior $p(I)$, which favors clear images over blurred images when minimizing (2). Therefore, the image prior $p(I)$ should have lower responses for clear images and higher responses for blurred images. This observation motivates us to learn a data-driven discriminative prior via binary classification. We train a deep CNN by predicting blurred images as positive (labeled as 1) and clear images as negative (labeled as 0) samples. Compared with state-of-the-art latent image priors [38, 27], the assumption of our prior is simple and straightforward without using any hand-crafted functions or assumptions.

3.2. Binary classification network

Our goal is to train a binary classifier via a deep CNN. The network takes an image as the input and outputs a single scalar, which represents the probability of the input image to be blurred. As we aim to embed the network as a prior

into the coarse-to-fine MAP framework, the network should be able to handle different sizes of input images. Therefore, we replace the commonly used fully-connected layers in classifiers with the global average pooling layer [21]. The global average pooling layer converts various sizes of feature maps into a single scalar before the sigmoid layer. In addition, there is no additional parameter in the global average pooling layer, which alleviates the overfitting problem. Figure 2 shows the architecture and detail parameters of our binary classification network.

3.3. Loss function

We denote the input image by x and the network parameters to be optimized by θ . The deep network learns a mapping function $f(x; \theta) = P(x \in \text{Blurred} | x)$ that predicts the probability of the input image to be blurred. We optimize the network via the binary cross entropy loss function:

$$L(\theta) = -\frac{1}{N} \sum_{i=1}^N \hat{y}_i \log(y_i) + (1 - \hat{y}_i) \log(1 - y_i), \quad (3)$$

where N is the number of training samples in a batch, $y_i = f(x_i; \theta)$ is the output of the classifier and \hat{y}_i is the label of the input image. We assign $\hat{y} = 1$ for blurred images and $\hat{y} = 0$ for clear images.

3.4. Training details

We sample 500 clear images from the dataset of Huiskes and Lew [13], including natural, manmade scene, face, low-illumination and text images. We use the method of Boracchi and Foi [2] to generate 200 random blur kernels with the size ranging from 7×7 to 51×51 . We synthesize blurred images by convolving the clear images with blur kernels and adding a Gaussian noise with $\sigma = 0.01$. We generate a total of 100,000 blurred images for training. During training, we randomly crop 200×200 patches from the training images. In order to make the classifier more robust to different

sizes of images, we adopt a multi-scale training strategy by randomly resizing the input images between $[0.25, 1]$.

We implement the network using the MatConvNet [34] toolbox. We use the Xavier method to initialize the network parameters and use the Stochastic Gradient Descent (SGD) method for optimizing the network. We use the batch size of 50, the momentum of 0.9 and the weight decay of 10^{-4} . The learning rate is set to 0.001 and decreased by a factor of 5 for every 50 epochs.

4. Blind Image Deblurring

After the training process of the proposed network converges, we use the trained model as the latent image prior $p(\cdot)$ in (2). In addition, we use the L_0 gradient prior [38, 27] as a regularization term. Therefore, we aim to solve the following optimization problem:

$$\min_{I, k} \|I \otimes k - B\|_2^2 + \gamma \|k\|_2^2 + \mu \|\nabla I\|_0 + \lambda f(I), \quad (4)$$

where γ , μ and λ are the hyper-parameters to balance the weight of each term.

We optimize (4) by solving the latent image I and the blur kernel k alternatively. Thus, we divide the problem into I sub-problem:

$$\min_I \|I \otimes k - B\|_2^2 + \mu \|\nabla I\|_0 + \lambda f(I), \quad (5)$$

and k sub-problem:

$$\min_k \|I \otimes k - B\|_2^2 + \gamma \|k\|_2^2. \quad (6)$$

4.1. Solving I

In (5), both $f(\cdot)$ and $\|\nabla I\|_0$ are highly non-convex, which make minimizing (5) computationally intractable. To tackle this issue, we adopt the half-quadratic splitting method [37] by introducing the auxiliary variables u and $g = (g_h, g_v)$ with respect to the image and its gradients in horizontal and vertical directions, respectively. The energy function (5) can be rewritten as

$$\min_{I, g, u} \|I \otimes k - B\|_2^2 + \alpha \|\nabla I - g\|_2^2 + \beta \|I - u\|_2^2 + \mu \|g\|_0 + \lambda f(u), \quad (7)$$

where α and β are penalty parameters. When α and β approach infinity, the solution of (7) is equivalent to that of (5). We can solve (7) by minimizing I , g and u alternatively and thus avoid directly minimizing the non-convex functions $f(\cdot)$ and $\|\nabla I\|_0$.

We solve the latent image I by fixing g and u and optimizing:

$$\min_I \|I \otimes k - B\|_2^2 + \alpha \|\nabla I - g\|_2^2 + \beta \|I - u\|_2^2, \quad (8)$$

Algorithm 1 Solving (12)

Input: Latent Image I
Output: the solution of u .
1: initialize $u^{(0)} \leftarrow I$
2: **while** $s < s_{\max}$ **do**
3: solve for $u^{(s+1)}$ by (12).
4: $s \leftarrow s + 1$
5: **end while**

which is a least squares optimization and has a closed-form solution:

$$I = F^{-1} \left(\frac{\overline{F(k)} F(B) + \beta F(u) + \alpha \left(\sum_{d \in \{h, v\}} \overline{F(\nabla_d)} F(g_d) \right)}{\overline{F(k)} F(k) + \beta + \alpha \left(\sum_{d \in \{h, v\}} \overline{F(\nabla_d)} F(\nabla_d) \right)} \right), \quad (9)$$

where $F(\cdot)$ and $F^{-1}(\cdot)$ denote the Fourier and inverse Fourier transforms; $\overline{F(\cdot)}$ is the complex conjugate operator; ∇_h and ∇_v are the horizontal and vertical differential operators, respectively.

Given the latent image I , we solve g and u by:

$$\min_g \alpha \|\nabla I - g\|_2^2 + \mu \|g\|_0, \quad (10)$$

$$\min_u \beta \|I - u\|_2^2 + \lambda f(u). \quad (11)$$

We solve (10) following the strategy of Pan et al. [26] and use the back-propagation approach to compute the derivative of $f(\cdot)$. We update u using the gradient descent method:

$$u^{(s+1)} = u^{(s)} - \eta \left[\beta \left(u^{(s)} - I \right) + \lambda \frac{df(u^{(s)})}{du^{(s)}} \right], \quad (12)$$

where η is the step size. We summarize the main steps for solving (12) in Algorithm 1.

4.2. Solving k

In order to obtain more accurate results, we estimate the blur kernel using image gradients [4, 26, 27]:

$$\min_k \|\nabla I \otimes k - \nabla B\|_2^2 + \gamma \|k\|_2^2, \quad (13)$$

which can also be efficiently solved by the Fast Fourier Transform (FFT). We then set the negative elements in k to 0 and normalize k so that the sum of all elements is equal to 1. We use the coarse-to-fine strategy with an image pyramid [26, 27] to optimize (4). At each pyramid level, we alternatively solve (5) and (13) with iter_{\max} iterations. The main steps are summarized in supplemental materials.

5. Extension to Non-Uniform Deblurring

The proposed discriminative image prior can be easily extended for non-uniform motion deblurring. Based on the geometric model of camera motion [33, 35], we represent

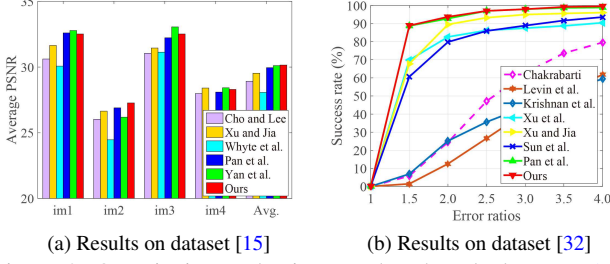


Figure 3. Quantitative evaluations on benchmark datasets [15] and [32].

the blurred images as the weighted sum of a latent clear image under geometry transformations:

$$\mathbf{B} = \sum_t k_t \mathbf{H}_t \mathbf{I} + \mathbf{n}, \quad (14)$$

where \mathbf{B} , \mathbf{I} and \mathbf{n} are the blurred image, latent image and noise in the vector forms, respectively; t denotes the index of camera pose samples; k_t is the weight of the t -th camera pose satisfying $k_t \geq 0$, $\sum_t k_t = 1$; \mathbf{H}_t denotes a matrix derived from the homography [35]. We use the bilinear interpolation when applying \mathbf{H}_t on a latent image \mathbf{I} . Therefore, we simplify (14) to:

$$\mathbf{B} = \mathbf{K} \mathbf{I} + \mathbf{n} = \mathbf{A} \mathbf{k} + \mathbf{n}, \quad (15)$$

where $\mathbf{K} = \sum_t k_t \mathbf{H}_t$, $\mathbf{A} = [\mathbf{H}_1 \mathbf{I}, \mathbf{H}_2 \mathbf{I}, \dots, \mathbf{H}_t \mathbf{I}]$ and $\mathbf{k} = [k_1, k_2, \dots, k_t]^T$. We solve the non-uniform deblurring problem by alternatively minimizing:

$$\min_{\mathbf{I}} \|\mathbf{K} \mathbf{I} - \mathbf{B}\|_2^2 + \lambda f(\mathbf{I}) + \mu \|\nabla \mathbf{I}\|_0 \quad (16)$$

and

$$\min_{\mathbf{k}} \|\mathbf{A} \mathbf{k} - \mathbf{B}\|_2^2 + \gamma \|\mathbf{k}\|_2^2. \quad (17)$$

The optimization methods of (16) and (17) are similar to those used for solving (5) and (6). The latent image \mathbf{I} and the weight \mathbf{k} are estimated by the fast forward approximation [9].

6. Experimental Results

We evaluate the proposed algorithm on natural image datasets [15, 32] as well as text [26], face [25], and low-illumination [11] images. In all the experiments, we set $\lambda = \mu = 0.004$, $\gamma = 2$, and $\eta = 0.1$. To balance the accuracy and speed, we empirically set $\text{iter}_{\max} = 5$ and $s_{\max} = 10$. Unless specially mentioned, we use the non-blind method in [26] to recover the final latent images after estimating blur kernels. All the experiments are carried out on a desktop computer with an Intel Core i7-3770 processor and 32 GB RAM. The source code and the datasets used in the paper are publicly available on the authors' websites. More experimental results are included in supplemental material.



Figure 5. Deblurred results on a real blurred image. Our result is sharper with less artifacts.

6.1. Natural images

We first evaluate the proposed algorithm on the natural image dataset of Köhler et al. [15], which contains 4 latent images and 12 blur kernels. We compare with the 5 generic image deblurring methods [4, 36, 35, 27, 40]. We follow the protocol of [15] to compute the PSNR by comparing each restored image with 199 clear images captured along the same camera motion trajectory. As shown in Figure 3 (a), our method achieves the highest PSNR on average. Figure 4 shows the deblurred results of one example. Our method generates clearer images with less ringing artifacts.

Next, we evaluate our algorithm on the dataset provided by Sun et al. [32], which consists of 80 clear images and 8 blur kernels from Levin et al. [19]. We compare with the 6 optimization-based deblurring methods [20, 16, 38, 36, 32, 27] (solid curves) and one learning-based method [3] (dotted curve). For fair comparisons, we apply the same non-blind deconvolution [43] to restore the latent images. We measure the error ratio [19] and plot the results in Figure 3 (b), which demonstrates that the proposed method performs competitively against the state-of-the-art algorithms.

We also test our method on real-world blurred images. Here we use the same non-blind deconvolution algorithm [26] for fair comparisons. As shown in Figure 5, our method generates clearer images with fewer artifacts compared with the methods [16, 38, 26]. And our result is comparable to the method [27].

6.2. Domain-specific images

We evaluate our algorithm on the text image dataset [26], which consists of 15 clear text images and 8 blur kernels from Levin et al. [19]. We show the average PSNR in Table 1. Although the text deblurring approach [26] has the highest PSNR, the proposed method performs fa-

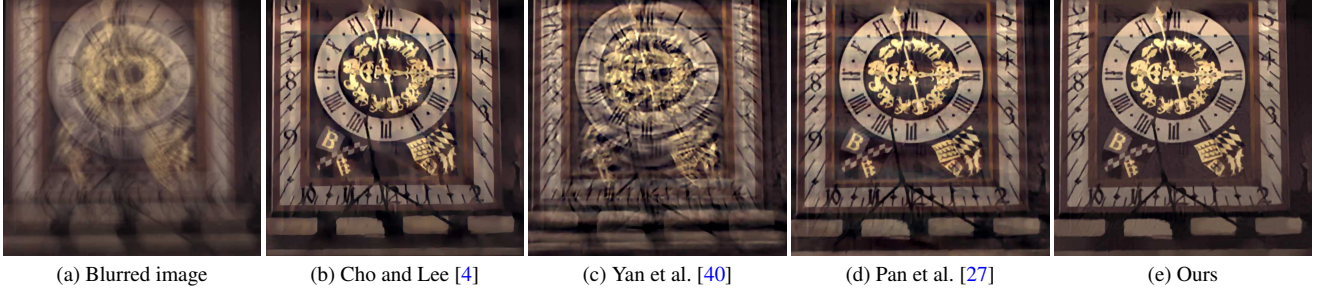


Figure 4. A challenging example from dataset [15]. The proposed algorithm restores more visually pleasing results with less ringing artifacts.

Table 1. Quantitative evaluations on text image dataset [26]. Our method performs favorably against generic image deblurring approaches and is comparable to the text deblurring method [26].

| Methods | Average PSNRs |
|-----------------------------------|---------------|
| Cho and Lee [4] | 23.80 |
| Xu and Jia [36] | 26.21 |
| Levin et al. [20] | 24.90 |
| Xu et al. [38] | 26.21 |
| Pan et al. [27] (Dark channel) | 27.94 |
| Pan et al. [26] (Text deblurring) | 28.80 |
| Ours | 28.10 |

vorably against state-of-the-art generic deblurring algorithms [4, 36, 20, 38, 27]. Figure 6 shows the deblurred results on a blurred text image. The proposed method generates much sharper results with clearer characters.

Figure 7 shows an example of the low-illumination image from the dataset of Hu et al. [11]. Due to the influence of large saturated regions, the natural image deblurring methods fail to generate clear images. In contrast, our method generates a comparable result with Hu et al. [11], which is specially designed for the low-illumination images.

Figure 8 shows the deblurred results on a face image. Our result has less ringing artifacts compared with the state-of-the-art methods [38, 40]. We note that the proposed method learns a *generic* image prior but is effective to deblur domain-specific blurred images.

6.3. Non-uniform deblurring

We demonstrate the capability of the proposed method on non-uniform deblurring in Figure 9. Compared with state-of-the-art non-uniform deblurring algorithms [35, 38, 27], our method produces comparable results with sharp edges and clear textures.

7. Analysis and Discussion

In this section, we analyze the effectiveness of the proposed image prior on distinguishing clear and blurred images, discuss the relation with L_0 -regularized priors, and

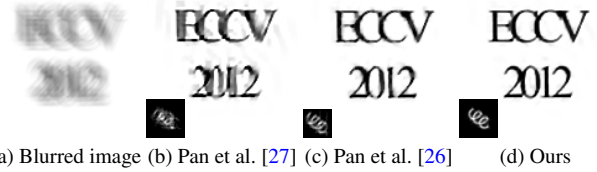


Figure 6. Deblurred results on a text image. Our method produces sharper deblurred image with more clearer characters than the state-of-the-art text deblurring algorithm [26].

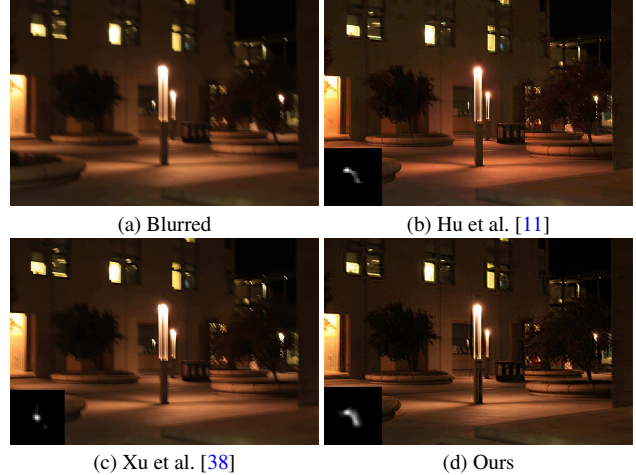


Figure 7. Deblurred results on a low-illumination image. Our method yields comparable results to Hu et al. [11], which is specially designed for deblurring low-illumination images.

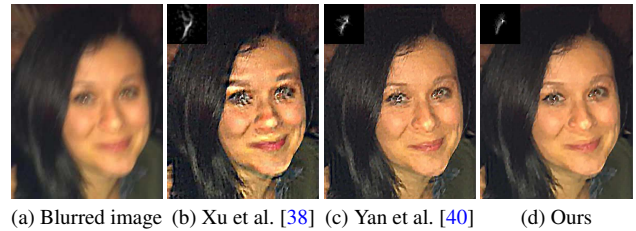


Figure 8. Deblurred results on a face image. Our method produces more visually pleasing results.

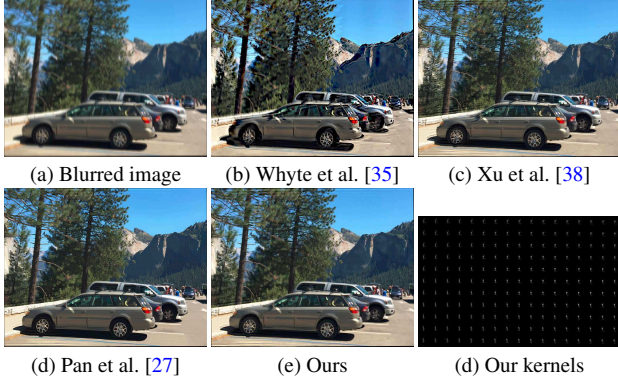


Figure 9. Deblurred results on a real non-uniform blurred image. We extend the proposed method for non-uniform deblurring and provide comparable results with state-of-the-art methods.

analyze the speed, convergence and the limitations of the proposed method.

7.1. Effectiveness of the proposed image prior

We train the binary classification network to predict blurred images as 1 and clear images as 0. We first use the image size of 200×200 for training and evaluate the classification accuracy using the images from the dataset of Köhler et al. [15], where the size of images is 800×800 . To test the performance of the classifier on different sizes of images, we downsample each image by a ratio between $[1, 1/16]$ and plot the classification accuracy in Figure 10 (green curve). When the size of test images is larger or close to the training image size, the accuracy is near 100%. However, the accuracy drops significantly when images are downsampled by more than $4\times$. As the downsampling reduces the blur effect, it becomes difficult for the classifier to distinguish blurred and clear images.

To overcome this issue, we adopt a multi-scale training strategy by randomly downsampling each batch of images between $1\times$ and $4\times$. As shown in the red curve of Figure 10(a), the performance of the classifier becomes more robust to different sizes of input images. The binary classifier with our multi-scale training strategy is more suitable to be applied in the coarse-to-fine MAP framework.

Figure 11 shows the activation of one feature map from the C9 layer (i.e., the last convolutional layer before the global average pooling) in our classification network. While the blurred image has a high response on the entire image, the activation of the clear image has a much lower response except for smooth regions, e.g., sky.

7.2. Relation with L_0 -regularized priors

Several methods [26, 38] adopt the L_0 -regularized priors in blind image deblurring due to the strong sparsity of the L_0 norm. State-of-the-art approaches [27, 40] enforce the L_0 sparsity on the extreme channels (i.e., dark and bright

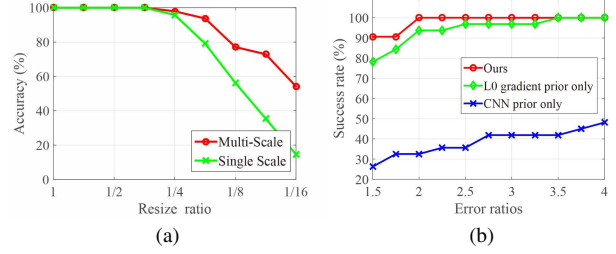


Figure 10. Effectiveness of the proposed CNN prior. (a) Classification accuracy on dataset [15] (b) Ablation studies on dataset [19]

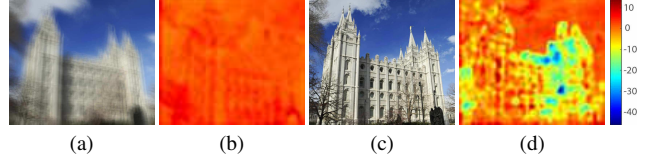


Figure 11. Activations of a feature map in our binary classification network. We show the activation from the C9 layer. The clear image has much lower responses than that of the blurred images. (a) Blurred image (b) Activation of blurred image (c) Clear image (d) Activation of clear image

channels) as the blur process affects the distribution of the extreme channels. The proposed approach also includes the L_0 gradient prior for regularization. The intermediate results in Figure 12 show that the methods based on L_0 -regularized prior on extreme channels [27, 40] fail to recover strong edges when there are not enough dark or bright pixels. Figure 12(g) shows that the proposed method without the learned discriminative image prior (i.e., use L_0 gradient prior only) cannot well reconstruct strong edges for estimating the blur kernel. In contrast, our discriminative image prior restores more sharp edges in the early stage of the optimization and improve the blur kernel estimation.

To better understand the effectiveness of each term in (4), we conduct an ablation study on the dataset of Levin et al. [19]. As shown in Figure 10(b), while the L_0 gradient prior helps to preserve more image structures, the integration with the proposed CNN prior leads to state-of-the-art performance.

7.3. Runtime and convergence property

Our algorithm is based on the efficient half-quadratic splitting and gradient decent methods. We test the state-of-the-art methods on different sizes of images and report the average runtime in Table 2. The proposed method runs competitively with state-of-the-art approaches [27, 40]. In addition, we quantitatively evaluate convergence of the proposed optimization method using images from the dataset of Levin et al. [19]. We compute the average kernel similarity [12] and the values of the objective function (4) at the finest image scale. Figure 13 shows that our algorithm converges well within 50 iterations.

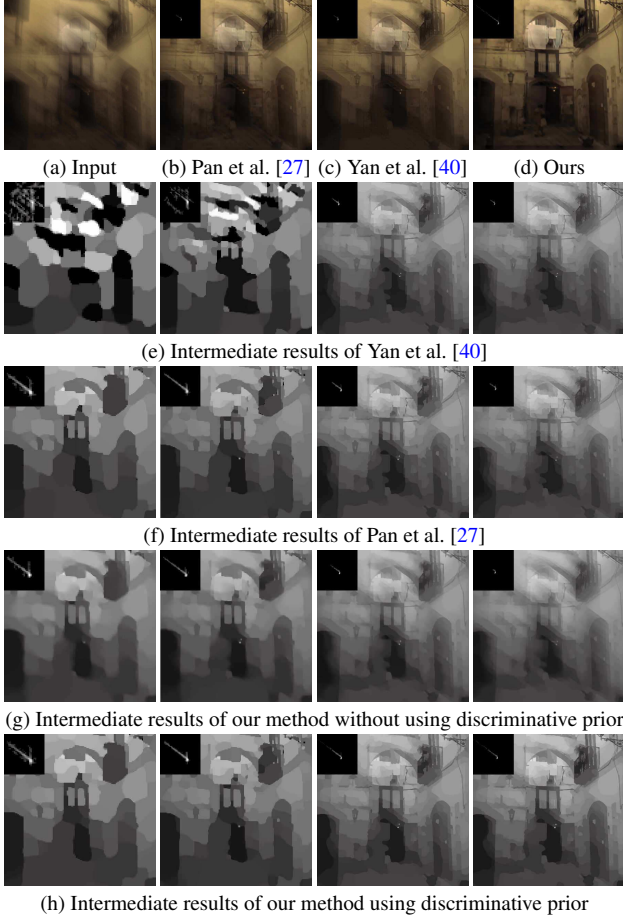


Figure 12. Deblurred and intermediate results. We compare the deblurred results with state-of-the-art methods [40, 27] in (a)-(d) and illustrate the intermediate latent images over iterations (from left to right) in (e)-(h). Our discriminative prior recovers intermediate results with more strong edges for kernel estimation.

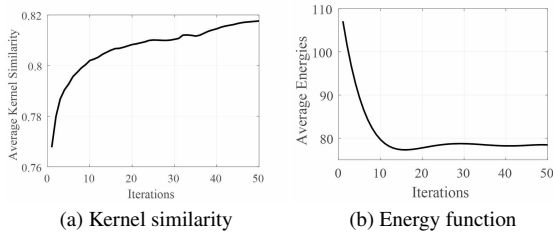


Figure 13. Convergence analysis of the proposed optimization method. We analyze the kernel similarity [12] and the objective function (4) at the finest image scale. Our method converges well within 50 iterations.

7.4. Limitations

As our classification network is trained on image intensity, the learned image prior might be less effective when input images contain significant noise and outliers. Figure 14 shows an example with salt and pepper noise in the input blurred image. In this case, our classification network

Table 2. Runtime comparisons. We report the average runtime (seconds) on three different sizes of images.

| Method | 255 × 255 | 600 × 600 | 800 × 800 |
|-------------------------------|-----------|-----------|-----------|
| Xu et al. [38] (C++) | 1.11 | 3.56 | 4.31 |
| Krishnan et al. [16] (MATLAB) | 24.23 | 111.09 | 226.58 |
| Levin et al. [20] (MATLAB) | 117.06 | 481.48 | 917.84 |
| Pan et al. [27] (MATLAB) | 134.31 | 691.71 | 964.90 |
| Yan et al. [40] (MATLAB) | 264.78 | 996.03 | 1150.48 |
| Ours (MATLAB) | 109.27 | 379.52 | 654.65 |

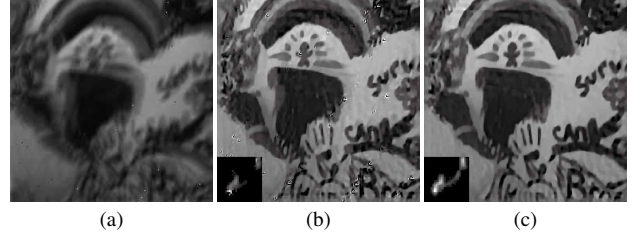


Figure 14. Limitations of the proposed method. Our learned image prior is not effective on handling images with salt and pepper noise. (a) Blurred image (b) Our deblurred results (c) Our deblurred results by first applying a median filter on blurred image.

cannot differentiate the blurred image ($f(B) \simeq 0$) due to the influence of salt and pepper noise. Therefore, the proposed prior cannot restore the image well as shown in Figure 14(b). A simple solution is to first apply a median filter on the input image before adopting our approach for deblurring. As shown in Figure 14(c), although we can reconstruct a better deblurred result, the details of the recovered images are not preserved well. Future work will consider joint deblurring and denoising in a principal way.

8. Conclusions

In this paper, we propose a data-driven discriminative prior for blind image deblurring. We learn the image prior via a binary classification network based on a simple criterion: the prior should favor clear images over blurred images on various of scenarios. We adopt a global average pooling layer and a multi-scale training strategy to make the network more robust to different sizes of images. We then embed the learned image prior into a coarse-to-fine MAP framework and develop an efficient half-quadratic splitting algorithm for blur kernel estimation. Our prior is effective on several types of images, including natural, text, face and low-illumination images, and can be easily extended to handle non-uniform deblurring. Extensive quantitative and qualitative comparisons demonstrate that the proposed method performs favorably against state-of-the-art generic and domain-specific blind deblurring algorithms.

Acknowledgements

This work is partially supported by NSFC (No. 61433007, 61401170, and 61571207), NSF CARRER (No. 1149783), and gifts from Adobe and Nvidia. Li L. is supported by a scholarship from China Scholarship Council.

References

- [1] S. A. Bigdeli, M. Zwicker, P. Favaro, and M. Jin. Deep mean-shift priors for image restoration. In *Neural Information Processing Systems*, 2017. 3
- [2] G. Boracchi and A. Foi. Modeling the performance of image restoration from motion blur. *IEEE Transactions on Image Processing*, 21(8):3502–3517, 2012. 3
- [3] A. Chakrabarti. A neural approach to blind motion deblurring. In *European Conference on Computer Vision*, 2016. 2, 5
- [4] S. Cho and S. Lee. Fast motion deblurring. *ACM Transactions on Graphics*, 28(5):145, 2009. 4, 5, 6
- [5] C. Dong, Y. Deng, C. Change Loy, and X. Tang. Compression artifacts reduction by a deep convolutional network. In *IEEE International Conference on Computer Vision*, 2015. 2
- [6] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In *European Conference on Computer Vision*, 2014. 2
- [7] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman. Removing camera shake from a single photograph. *ACM Transactions on Graphics*, 25(3):787–794, 2006. 2
- [8] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016. 2
- [9] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Schölkopf. Fast removal of non-uniform camera shake. In *IEEE International Conference on Computer Vision*, 2011. 5
- [10] M. Hradíš, J. Kotera, P. Zemčík, and F. Šroubek. Convolutional neural networks for direct text deblurring. In *British Machine Vision Conference*, 2015. 2
- [11] Z. Hu, S. Cho, J. Wang, and M.-H. Yang. Deblurring low-light images with light streaks. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2014. 2, 5, 6
- [12] Z. Hu and M.-H. Yang. Good regions to deblur. In *European Conference on Computer Vision*, 2012. 7, 8
- [13] M. J. Huiskes and M. S. Lew. The MIR flickr retrieval evaluation. In *ACM international conference on Multimedia information retrieval*, 2008. 3
- [14] J. Kim, J. K. Lee, and K. M. Lee. Accurate image super-resolution using very deep convolutional networks. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016. 2
- [15] R. Köhler, M. Hirsch, B. Mohler, B. Schölkopf, and S. Harmeling. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database. In *European Conference on Computer Vision*, 2012. 5, 6, 7
- [16] D. Krishnan, T. Tay, and R. Fergus. Blind deconvolution using a normalized sparsity measure. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2011. 2, 5, 8
- [17] W.-S. Lai, J.-J. Ding, Y.-Y. Lin, and Y.-Y. Chuang. Blur kernel estimation using normalized color-line prior. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2015. 2
- [18] W.-S. Lai, J.-B. Huang, N. Ahuja, and M.-H. Yang. Deep laplacian pyramid networks for fast and accurate super-resolution. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. 2
- [19] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Understanding and evaluating blind deconvolution algorithms. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2009. 2, 5, 7
- [20] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2011. 5, 6, 8
- [21] M. Lin, Q. Chen, and S. Yan. Network in network. *arXiv*, 2013. 2, 3
- [22] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2015. 2
- [23] X. Mao, C. Shen, and Y.-B. Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In *Neural Information Processing Systems*, 2016. 2
- [24] T. Michaeli and M. Irani. Blind deblurring using internal patch recurrence. In *European Conference on Computer Vision*, 2014. 2
- [25] J. Pan, Z. Hu, Z. Su, and M.-H. Yang. Deblurring face images with exemplars. In *European Conference on Computer Vision*, 2014. 2, 5
- [26] J. Pan, Z. Hu, Z. Su, and M. H. Yang. Deblurring text images via l0-regularized intensity and gradient prior. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2014. 2, 4, 5, 6, 7
- [27] J. Pan, D. Sun, H. Pfister, and M.-H. Yang. Blind image deblurring using dark channel prior. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2016. 1, 2, 3, 4, 5, 6, 7, 8
- [28] K. Schelten, S. Nowozin, J. Jancsary, C. Rother, and S. Roth. Interleaved regression tree field cascades for blind image deconvolution. In *IEEE Winter Conference on Applications of Computer Vision*, 2015. 3
- [29] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf. Learning to deblur. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(7):1439–1451, 2016. 2
- [30] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman. Plug-and-play priors for bright field electron tomography and sparse interpolation. *IEEE Transactions on Computational Imaging*, 2(4):408–423, 2016. 2
- [31] J. Sun, W. Cao, Z. Xu, and J. Ponce. Learning a convolutional neural network for non-uniform motion blur removal. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2015. 2
- [32] L. Sun, S. Cho, J. Wang, and J. Hays. Edge-based blur kernel estimation using patch priors. In *IEEE International Conference on Computational Photography*, 2013. 2, 5
- [33] Y.-W. Tai, P. Tan, and M. S. Brown. Richardson-lucy deblurring for scenes under a projective motion path. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(8):1603–1618, 2011. 4

- [34] A. Vedaldi and K. Lenc. Matconvnet: Convolutional neural networks for matlab. In *ACM International conference on Multimedia*, 2015. [4](#)
- [35] O. Whyte, J. Sivic, A. Zisserman, and J. Ponce. Non-uniform deblurring for shaken images. *International Journal of Computer Vision*, 98(2):168–186, 2012. [4](#), [5](#), [6](#), [7](#)
- [36] L. Xu and J. Jia. Two-phase kernel estimation for robust motion deblurring. In *European Conference on Computer Vision*, 2010. [2](#), [5](#), [6](#)
- [37] L. Xu, C. Lu, Y. Xu, and J. Jia. Image smoothing via l_0 gradient minimization. *ACM Transactions on Graphics*, 30(6):174, 2011. [4](#)
- [38] L. Xu, S. Zheng, and J. Jia. Unnatural l_0 sparse representation for natural image deblurring. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2013. [1](#), [2](#), [3](#), [4](#), [5](#), [6](#), [7](#), [8](#)
- [39] R. Yan and L. Shao. Blind image blur estimation via deep learning. *IEEE Transactions on Image Processing*, 25(4):1910–1921, 2016. [2](#)
- [40] Y. Yan, W. Ren, Y. Guo, R. Wang, and X. Cao. Image deblurring via extreme channels prior. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. [2](#), [5](#), [6](#), [7](#), [8](#)
- [41] J. Zhang, J. Pan, W.-S. Lai, R. Lau, and M.-H. Yang. Learning fully convolutional networks for iterative non-blind deconvolution. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. [2](#)
- [42] K. Zhang, W. Zuo, S. Gu, and L. Zhang. Learning deep cnn denoiser prior for image restoration. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2017. [2](#)
- [43] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *IEEE International Conference on Computer Vision*, 2011. [5](#)