

# Deep Constrained Least Squares for Blind Image Super-Resolution

Ziwei Luo<sup>1</sup> Haibin Huang<sup>2</sup> Lei Yu<sup>1</sup> Youwei Li<sup>1</sup> Haoqiang Fan<sup>1</sup> Shuaicheng Liu<sup>3,1\*</sup>

<sup>1</sup> Megvii Technology <sup>2</sup> Kuaishou Technology

<sup>3</sup> University of Electronic Science and Technology of China

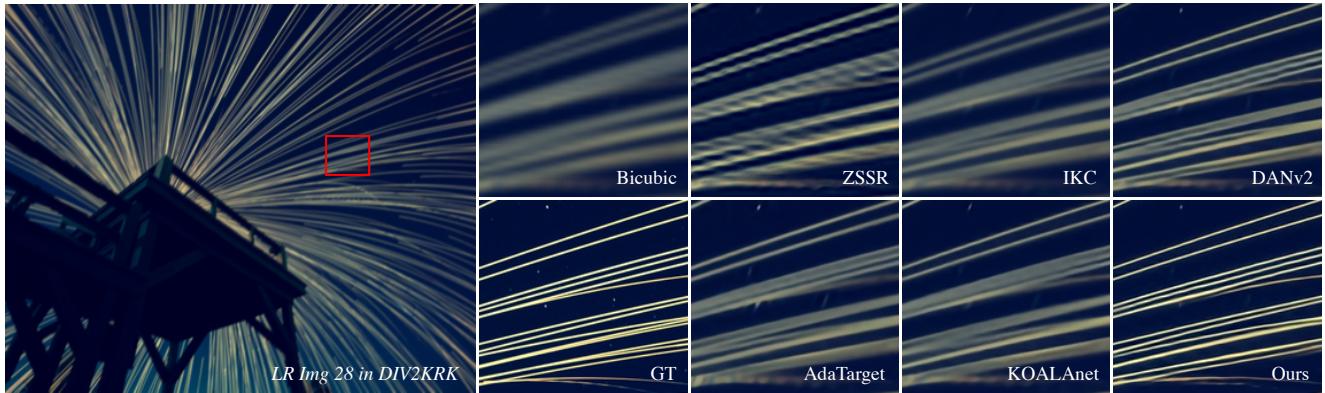


Figure 1. Blind super-resolution of *Img 28* from DIV2KRK [3], for scale factor 4. Based on the proposed deep constrained least squares (DCLS) deconvolution, our method is effective in restoring sharp and clean edges, and outperforms previous state-of-the-art approaches such as KernelGAN [3]+ZSSR [41], IKC [9], DAN [30, 31], AdaTarget [14], and KOALAnet [19].

## Abstract

In this paper, we tackle the problem of blind image super-resolution(SR) with a reformulated degradation model and two novel modules. Following the common practices of blind SR, our method proposes to improve both the kernel estimation as well as the kernel based high resolution image restoration. To be more specific, we first reformulate the degradation model such that the deblurring kernel estimation can be transferred into the low resolution space. On top of this, we introduce a dynamic deep linear filter module. Instead of learning a fixed kernel for all images, it can adaptively generate deblurring kernel weights conditional on the input and yields more robust kernel estimation. Subsequently, a deep constrained least square filtering module is applied to generate clean features based on the reformulation and estimated kernel. The deblurred feature and the low input image feature are then fed into a dual-path structured SR network and restore the final high resolution result. To evaluate our method, we further conduct evaluations on several benchmarks, including Gaussian8 and DIV2KRK. Our experiments demonstrate that the proposed method achieves better accuracy and visual improvements against state-of-the-art methods. Codes and models are

available at <https://github.com/megvii-research/DCLS-SR>.

## 1. Introduction

In this work, we study the problem of image super-resolution,i.e., restoring high-resolution images from low-resolution inputs. Specially, we aim for single image super-resolution (SISR), where only one observation is given which is a more practical setting and with a wide range of downstream applications [6, 8, 10, 17, 22, 26, 28, 48, 57, 59].

Most existing works based on the classical SISR degradation model assuming that the input LR image  $\mathbf{y}$  is a blurred and down-scaled HR image  $\mathbf{x}$  with additional white Gaussian noise  $\mathbf{n}$ , given by

$$\mathbf{y} = (\mathbf{x} * \mathbf{k}_h)_{\downarrow_s} + \mathbf{n}, \quad (1)$$

where  $\mathbf{k}_h$  is the blur kernel applied on  $\mathbf{x}$ ,  $*$  denotes convolution operation and  $\downarrow_s$  denotes downsampling with scale factor  $s$ . Previous blind SR approaches [9, 30] generally solve this problem with a two-stage framework: kernel estimation from LR image and kernel based HR image restoration.

We argue that although such a pipeline demonstrates reasonable performance for SR problem, there are two main drawbacks: First of all, it is difficult to accurately estimate

\*Corresponding author.

blur kernels of HR space directly from LR images due to the ambiguity produced by undersampling step [38, 46]. And the mismatch between the estimated kernel and the real one will cause significant performance drop and even lead to unpleasant artifacts [3, 9, 13, 56]. Secondly, it is also challenging to find a suitable way to fully utilize the information of the estimated HR space kernel and LR space image. A common solution is to employ a kernel stretching strategy [9, 30, 56], where the principal components of the vectorized kernel are preserved and stretched into degradation maps with the same size as the LR input. These degradation maps then can be concatenated with the input image or its features to generate a clean HR image. However, the spatial relation of the kernel is destroyed by the process of vectorizing and PCA (Principal Component Analysis), which causes insufficient usage of the kernel. The subsequent reconstruction network requires a huge effort to harmonize the inconsistent information between LR features and HR-specific kernels, limiting its performance in super-resolving images.

Towards this end, we present a modified learning strategy to tackle the blind SR problem, which can naturally avoid the above mentioned drawbacks. Specifically, we first reformulate the degradation model in a way such that the blur kernel estimation and image upsampling can be disentangled. In particular, as shown in Fig. 2, we derive a new kernel from the primitive kernel  $k_h$  and LR image. It transfers the kernel estimation into the LR space and the new kernel can be estimated without aliasing ambiguity. Based on the new degradation, we further introduce the dynamic deep linear kernel (DDLK) to provide more equivalent choices of possible optimal solutions for the kernel to accelerate training. Subsequently, a novel deep constrained least squares (DCLS) deconvolution module is applied in the feature domain to obtain deblurred features. DCLS is robust to noise and can provide a theoretical and principled guidance to obtain clean images/features from blurred inputs. Moreover, it doesn't require kernel stretching strategy and thus preserves the kernel's spatial relation information. Then the deblurred features are fed into an upsampling module to restore the clean HR images. As illustrated in Fig. 1, the overall method has turned out to be surprisingly effective in recovering sharp and clean SR images.

The main contributions are summarized as follows:

- We introduce a new practical degradation model derived from Eq. (1). Such degradation maintains consistency with the classical model and allows us reliably estimate blur kernel from low-resolution space.
- We propose to use a dynamic deep linear kernel instead of a single layer kernel, which provides more equivalent choices of the optimal solution of the kernel, which is easier to learn.

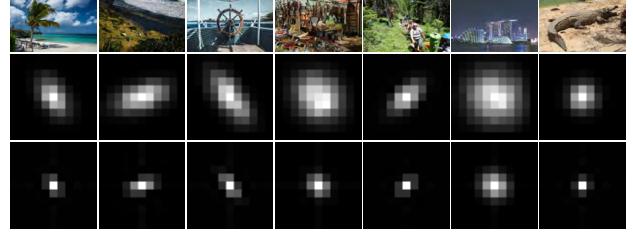


Figure 2. Kernel reformulation examples. The top row and middle row are the LR images and the corresponding primitive kernels. The bottom row is the reformulated kernels.

- We propose a novel deconvolution module named DCLS that is applied on the features as channel-wise deblurring so that we can obtain a clean HR image.
- Extensive experiments on various degradation kernels demonstrate that our method leads to state-of-the-art performance in blind SR problems.

## 2. Related work

**Non-blind SR** Since pioneering work SRCNN [6] proposes to learn image SR with a three-layer convolution network, most subsequent works have focused on optimizing the network architectures [5, 10, 17, 18, 21, 28, 32, 40, 43, 55, 59, 61, 62] and loss functions [15, 22, 29, 47, 48, 52, 58]. These CNN-based methods have achieved impressive performance on SISR with a predefined single degradation setting (e.g., bicubic downsampling). However, they may suffer significant performance drops when the predefined degradation kernel is different from the real one.

Some non-blind SR approaches address the multiple degradation problem by restoring HR images with given the corresponding kernels. Specifically, SRMD [56] is the first method that concatenates LR image with a stretched blur kernel as inputs to obtain a super-resolved image under different degradations. Later, Zhang *et al.* [54, 57] incorporate advanced deblurring algorithms and extend the degradation to arbitrary blur kernels. UDVD [51] improves the performance by incorporating dynamic convolution. Hussein *et al.* [13] introduce a correction filter that transfers blurry LR images to match the bicubically designed SR model. Besides, zero-shot methods [42, 51] have also been investigated in non-blind SR with multiple degradations.

**Blind SR** Under the blind SR setting, HR image is recovered from the LR image degraded with unknown kernel [24, 25, 35]. Most approaches solve this problem with a two stage framework: kernel estimation and kernel-based HR image restoration. For the former, KernelGAN [3] estimates the degradation kernel by utilizing an internal generative adversarial network(GAN) on a single image, and applies that kernel to a non-blind SR approach such as ZSSR to get the SR result. Liang *et al.* [27] improve the kernel

estimating performance by introducing a flow-based prior. Furthermore, Tao *et al.* [44] propose a spectrum-to-kernel network and demonstrate that estimating blur kernel in the frequency domain is more conducive than in spatial domain. For the latter, Gu *et al.* [9] propose to apply spatial feature transform (SFT) and iterative kernel correction (IKC) strategy for accurate kernel estimation and SR refinement. Luo *et al.* [30] develop an end-to-end training deep alternating network (DAN) by estimating reduced kernel and restoring HR image iteratively. However, both IKC and DAN are time-consuming and computationally costly. The modified version of DAN [31] conducts a dual-path conditional block (DPCB) and supervises the estimator on the complete blur kernel to further improve the performance.

### 3. Method

We now formally introduce our method which consists of three main components given a reformation of degradation: A dynamic deep linear kernel estimation module and a deep constrained least squares module for kernel estimation and LR space feature based deblur. A dual-path network is followed to generate the clean HR output. We will first derive the reformulation and then detail each module.

#### 3.1. Degradation Model Reformulation

Ideally, the blur kernel to be estimated and its corresponding image should be in the same low-resolution space such that the degradation can be transformed to the deblurring problem followed by a SISR problem with bicubic degradation [56, 57]. Towards this end, we propose to reformulate Eq. (1) as

$$\mathbf{y} = \mathcal{F}^{-1}(\mathcal{F}((\mathbf{x} * \mathbf{k}_h)_{\downarrow s})) + \mathbf{n} \quad (2)$$

$$= \mathcal{F}^{-1}\left(\mathcal{F}(\mathbf{x}_{\downarrow s}) \frac{\mathcal{F}((\mathbf{x} * \mathbf{k}_h)_{\downarrow s})}{\mathcal{F}(\mathbf{x}_{\downarrow s})}\right) + \mathbf{n} \quad (3)$$

$$= \mathbf{x}_{\downarrow s} * \mathcal{F}^{-1}\left(\frac{\mathcal{F}((\mathbf{x} * \mathbf{k}_h)_{\downarrow s})}{\mathcal{F}(\mathbf{x}_{\downarrow s})}\right) + \mathbf{n}, \quad (4)$$

where  $\mathcal{F}$  denotes the Discrete Fourier Transform and  $\mathcal{F}^{-1}$  denotes its inverse. Then let

$$\mathbf{k}_l = \mathcal{F}^{-1}\left(\frac{\mathcal{F}((\mathbf{x} * \mathbf{k}_h)_{\downarrow s})}{\mathcal{F}(\mathbf{x}_{\downarrow s})}\right), \quad (5)$$

we can obtain another form of degradation:

$$\mathbf{y} = \mathbf{x}_{\downarrow s} * \mathbf{k}_l + \mathbf{n}. \quad (6)$$

In the Eq. (6),  $\mathbf{k}_l$  is derived from the corresponding  $\mathbf{k}_h$  and applied on the downsampled HR image  $\mathbf{x}_{\downarrow s}$ . To ensure numerical stability, we rewrite Eq. (5) with a small regularization parameter  $\epsilon$ :

$$\mathbf{k}_l = \mathcal{F}^{-1}\left(\frac{\overline{\mathcal{F}(\mathbf{x}_{\downarrow s})}}{\overline{\mathcal{F}(\mathbf{x}_{\downarrow s})}\mathcal{F}(\mathbf{x}_{\downarrow s}) + \epsilon} \mathcal{F}((\mathbf{x} * \mathbf{k}_h)_{\downarrow s})\right), \quad (7)$$

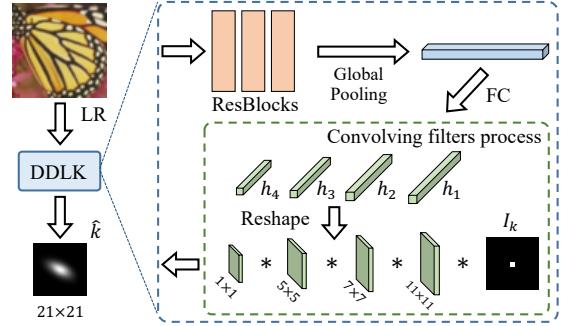


Figure 3. Architecture of the dynamic deep linear kernel.

where  $\overline{\mathcal{F}(\cdot)}$  is the complex conjugate of  $\mathcal{F}$ . Fig. 2 illustrates the results of reformulating kernels by Eq. (7). Based on the new degradation process, our goal is to estimate the blur kernel  $\mathbf{k}_l$  and then restore HR image  $\mathbf{x}$ .

#### 3.2. Dynamic Deep Linear Kernel

Following the reformation, we start our blind SR method from the kernel estimation. A straightforward solution is to adopt a regression network to estimate kernel  $\hat{\mathbf{k}}$  by minimizing the L1 difference w.r.t the new ground-truth blur kernel  $\mathbf{k}_l$  in Eq. (7). We argue such a single layer kernel (all weights of estimated kernel equal to the ground-truth kernel) estimation is in general difficult and unstable due to the highly non-convex of the blind SR problem [3], leading to kernel mismatch and performance drop [9, 30]. Instead, we propose an image-specific dynamic deep linear kernel (DDLK) which consists of a sequence of linear convolution layers without activations. Theoretically, deep linear networks have infinitely equivalent global minimas [3, 16, 39], which allow us to find many different filter parameters to achieve the same correct solution. Moreover, since no non-linearity is used in the network, we can analytically collapse a deep linear kernel as a single layer kernel.

Fig. 3 depicts an example of estimating 4 layers dynamic deep linear kernel. The filters are set to  $11 \times 11$ ,  $7 \times 7$ ,  $5 \times 5$  and  $1 \times 1$ , which make the receptive field to be  $21 \times 21$ . We first generate the filters of each layer based on the LR image, and explicitly sequentially convolve all filters into a single narrow kernel with stride 1. Mathematically, let  $\mathbf{h}_i$  represent the  $i$ -th layer filter, we can get a single layer kernel following

$$\hat{\mathbf{k}} = \mathbf{I}_k * \mathbf{h}_1 * \mathbf{h}_2 * \cdots * \mathbf{h}_r \quad (8)$$

where  $r$  is the number of linear layers,  $\mathbf{I}_k$  is an identity kernel. As an empirically prior, we also constrain the kernel  $\hat{\mathbf{k}}$  sum up to 1. The kernel estimation network can be optimized by minimizing the L1 loss between estimated kernel  $\hat{\mathbf{k}}$  and new ground-truth blur kernel  $\mathbf{k}_l$  from Eq. (7).

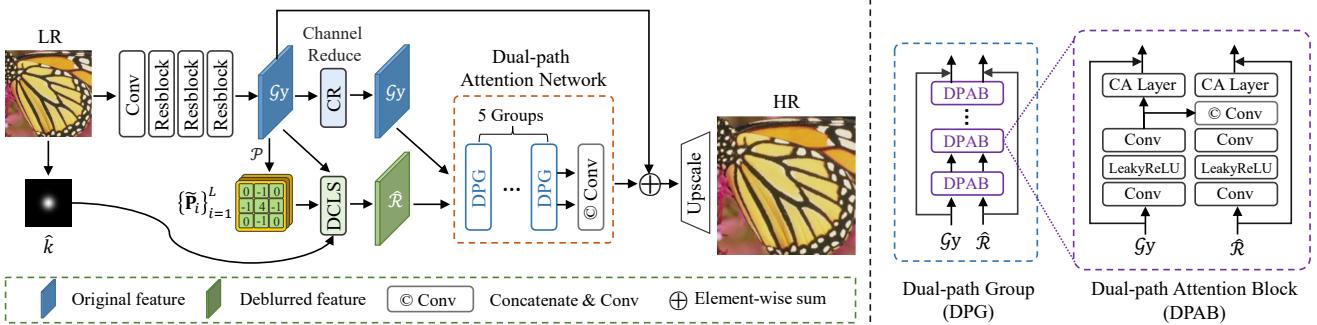


Figure 4. The overview architecture of the proposed method. Given an LR image  $\mathbf{y}$ , we first estimate the degradation kernel  $\hat{\mathbf{k}}$ , and involve it in the deep constrained least squares (DCLS) convolution in the feature domain. The deblurred features  $\hat{\mathcal{R}}$  are then concatenated with primitive features  $\mathcal{G}\mathbf{y}$  to restore the clean HR image  $\mathbf{x}$  through a dual-path attention network (DPAN).

### 3.3. Deep Constrained Least Squares

Our goal is to restore HR image based on LR image and estimated kernel  $\hat{\mathbf{k}}$  according to the new degradation model (Eq. (6)). Considering a group of feature extracting linear layers  $\{\mathcal{G}_i\}_{i=1}^L$  provided to the LR image, we can rewrite Eq. (6) in the feature space, given by

$$\mathcal{G}_i \mathbf{y} = \hat{\mathbf{k}} \mathcal{G}_i \mathbf{x}_{\downarrow_s} + \mathcal{G}_i \mathbf{n}. \quad (9)$$

Let  $\hat{\mathcal{R}}_i$  be the sought after deblurred feature corresponding to  $\mathcal{G}_i \mathbf{x}_{\downarrow_s}$ . To solve Eq. (9), we minimize the following criterion function

$$\mathcal{C} = \|\nabla \hat{\mathcal{R}}_i\|^2, \text{ s.t. } \|\mathcal{G}_i \mathbf{y} - \hat{\mathbf{k}} \hat{\mathcal{R}}_i\|^2 = \|\mathcal{G}_i \mathbf{n}\|^2 \quad (10)$$

where the  $\nabla$  is a smooth filter which can be denoted by  $\mathbf{P}$ . Then we introduce the Lagrange function, defined by

$$\min_{\hat{\mathcal{R}}_i} \left[ \|\mathbf{P} \hat{\mathcal{R}}_i\|^2 + \lambda \left( \|\mathcal{G}_i \mathbf{y} - \hat{\mathbf{k}} \hat{\mathcal{R}}_i\|^2 - \|\mathcal{G}_i \mathbf{n}\|^2 \right) \right], \quad (11)$$

where  $\lambda$  is the Lagrange multiplier. Computing the derivative of Eq. (11) with respect to  $\hat{\mathcal{R}}_i$  and setting it to zero:

$$(\lambda \hat{\mathbf{k}}^\top \hat{\mathbf{k}} + \mathbf{P}^\top \mathbf{P}) \hat{\mathcal{R}}_i - \lambda \hat{\mathbf{k}}^\top \mathcal{G}_i \mathbf{y} = 0. \quad (12)$$

We can obtain the clear features as

$$\hat{\mathcal{R}}_i = \mathcal{H} \mathcal{G}_i \mathbf{y}. \quad (13)$$

where  $\mathcal{H}_i$  denotes the deep constrained least squares deconvolution (DCLS) operator, given by

$$\mathcal{H} = \mathcal{F}^{-1} \left( \frac{\overline{\mathcal{F}(\hat{\mathbf{k}})}}{\overline{\mathcal{F}(\hat{\mathbf{k}})} \mathcal{F}(\hat{\mathbf{k}}) + \frac{1}{\lambda} \overline{\mathcal{F}(\mathbf{P})} \mathcal{F}(\mathbf{P})} \right). \quad (14)$$

Different from in the standard image space (e.g. RGB), smooth filter  $\mathbf{P}$  and variable  $\lambda$  in Eq. (14) might be inconsistent in the feature space. Alternatively, we predict a group

of smooth filters with implicit Lagrange multiplier for different channels through a neural network  $\mathcal{P}$ :

$$\{\tilde{\mathbf{P}}_i\}_{i=1}^L = \{\mathcal{P}(\mathcal{G}_i \mathbf{y})\}_{i=1}^L. \quad (15)$$

Then the feature-specific operator  $\mathcal{H}_i$  can be defined by

$$\mathcal{H}_i = \mathcal{F}^{-1} \left( \frac{\overline{\mathcal{F}(\hat{\mathbf{k}})}}{\overline{\mathcal{F}(\hat{\mathbf{k}})} \mathcal{F}(\hat{\mathbf{k}}) + \overline{\mathcal{F}(\tilde{\mathbf{P}}_i)} \mathcal{F}(\tilde{\mathbf{P}}_i)} \right). \quad (16)$$

Now we can obtain the clear features by Eq. (13) and Eq. (16).

It is worth to note that a deep neural network (DNN) can be locally linear [7, 23, 36], thus we could apply DNN as  $\mathcal{G}_i$  to extract useful features in Eq. (9). In addition, the consequent artifacts or errors can be compensated by the following dual-path attention module.

### 3.4. Dual-Path Attention Network

Unlike previous works [9, 31] in which the dual-path structures are only used to concatenate the stretched kernel with blurred features, we propose to utilize primitive blur features as additive path to compensate the artifacts and errors introduced by the estimated kernel, known as dual-path attention network (DPAN). DPAN is composed of several groups of dual-path attention blocks (DPAB), it receives both deblurred features  $\hat{\mathcal{R}}$  and primitive features  $\mathcal{G}\mathbf{y}$ . The right of Fig. 4 illustrates the architecture of DPAB.

Since the additive path of processing  $\mathcal{G}\mathbf{y}$  is independently updated and used to concatenate with  $\hat{\mathcal{R}}$  to provide primary information to refine the deconvolved features. We can reduce its channels to accelerate training and inference, as the channel reduction (CR) operation illustrated in left of Fig. 4. Moreover, on the deconvolved feature path, we apply the channel attention layer [60] after aggregating original features. In addition, we add a residual connection for each path on all groups and blocks. The pixelshuffle [11] is

Method	Scale	Set5 [4]		Set14 [53]		BSD100 [33]		Urban100 [12]		Manga109 [34]	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic	x2	28.82	0.8577	26.02	0.7634	25.92	0.7310	23.14	0.7258	25.60	0.8498
CARN [2]		30.99	0.8779	28.10	0.7879	26.78	0.7286	25.27	0.7630	26.86	0.8606
Bicubic+ZSSR [41]		31.08	0.8786	28.35	0.7933	27.92	0.7632	25.25	0.7618	28.05	0.8769
Deblurring [37]+CARN [41]		24.20	0.7496	21.12	0.6170	22.69	0.6471	18.89	0.5895	21.54	0.7946
CARN [41]+Deblurring [37]		31.27	0.8974	29.03	0.8267	28.72	0.8033	25.62	0.7981	29.58	0.9134
IKC [9]		37.19	0.9526	32.94	0.9024	31.51	0.8790	29.85	0.8928	36.93	0.9667
DANv1 [30]		37.34	0.9526	33.08	0.9041	31.76	0.8858	30.60	0.9060	37.23	0.9710
DANv2 [31]		<b>37.60</b>	<b>0.9544</b>	<b>33.44</b>	<b>0.9094</b>	<b>32.00</b>	<b>0.8904</b>	<b>31.43</b>	<b>0.9174</b>	<b>38.07</b>	<b>0.9734</b>
DCLS(Ours)		<b>37.63</b>	<b>0.9554</b>	<b>33.46</b>	<b>0.9103</b>	<b>32.04</b>	<b>0.8907</b>	<b>31.69</b>	<b>0.9202</b>	<b>38.31</b>	<b>0.9740</b>
Bicubic	x3	26.21	0.7766	24.01	0.6662	24.25	0.6356	21.39	0.6203	22.98	0.7576
CARN [2]		27.26	0.7855	25.06	0.6676	25.85	0.6566	22.67	0.6323	23.85	0.7620
Bicubic+ZSSR [41]		28.25	0.7989	26.15	0.6942	26.06	0.6633	23.26	0.6534	25.19	0.7914
Deblurring [37]+CARN [41]		19.05	0.5226	17.61	0.4558	20.51	0.5331	16.72	0.5895	18.38	0.6118
CARN [41]+Deblurring [37]		30.31	0.8562	27.57	0.7531	27.14	0.7152	24.45	0.7241	27.67	0.8592
IKC [9]		33.06	0.9146	29.38	0.8233	28.53	0.7899	24.43	0.8302	32.43	0.9316
DANv1 [30]		34.04	0.9199	30.09	0.8287	28.94	0.7919	27.65	0.8352	33.16	0.9382
DANv2 [31]		<b>34.12</b>	<b>0.9209</b>	<b>30.20</b>	<b>0.8309</b>	<b>29.03</b>	<b>0.7948</b>	<b>27.83</b>	<b>0.8395</b>	<b>33.28</b>	<b>0.9400</b>
DCLS(Ours)		<b>34.21</b>	<b>0.9218</b>	<b>30.29</b>	<b>0.8329</b>	<b>29.07</b>	<b>0.7956</b>	<b>28.03</b>	<b>0.8444</b>	<b>33.54</b>	<b>0.9414</b>
Bicubic	x4	24.57	0.7108	22.79	0.6032	23.29	0.5786	20.35	0.5532	21.50	0.6933
CARN [2]		26.57	0.7420	24.62	0.6226	24.79	0.5963	22.17	0.5865	21.85	0.6834
Bicubic+ZSSR [41]		26.45	0.7279	24.78	0.6268	24.97	0.5989	22.11	0.5805	23.53	0.7240
Deblurring [37]+CARN [41]		18.10	0.4843	16.59	0.3994	18.46	0.4481	15.47	0.3872	16.78	0.5371
CARN [41]+Deblurring [37]		28.69	0.8092	26.40	0.6926	26.10	0.6528	23.46	0.6597	25.84	0.8035
IKC [9]		31.67	0.8829	28.31	0.7643	27.37	0.7192	25.33	0.7504	28.91	0.8782
DANv1 [30]		31.89	0.8864	28.42	0.7687	27.51	0.7248	25.86	0.7721	<b>30.50</b>	<b>0.9037</b>
DANv2 [31]		<b>32.00</b>	<b>0.8885</b>	<b>28.50</b>	<b>0.7715</b>	<b>27.56</b>	<b>0.7277</b>	<b>25.94</b>	<b>0.7748</b>	30.45	0.9037
AdaTarget [14]		31.58	0.8814	28.14	0.7626	27.43	0.7216	25.72	0.7683	29.97	0.8955
DCLS(Ours)		<b>32.12</b>	<b>0.8890</b>	<b>28.54</b>	<b>0.7728</b>	<b>27.60</b>	<b>0.7285</b>	<b>26.15</b>	<b>0.7809</b>	<b>30.86</b>	<b>0.9086</b>

Table 1. Quantitative comparison on datasets with *Gaussian8* kernels. The best two results are marked in red and blue colors, respectively.

Method ×4	Noise level	Set5 [4]		Set14 [53]		BSD100 [33]		Urban100 [12]		Manga109 [34]	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bicubic+ZSSR [41]	15	23.32	0.4868	22.49	0.4256	22.61	0.3949	20.68	0.3966	22.04	0.4952
IKC [9]		26.89	0.7671	25.28	0.6483	24.93	0.6019	22.94	0.6362	25.09	0.7819
DANv1 [30]		26.95	0.7711	25.27	0.6490	<b>24.95</b>	<b>0.6033</b>	23.00	0.6407	25.29	0.7879
DANv2 [31]		<b>26.97</b>	<b>0.7726</b>	<b>25.29</b>	<b>0.6497</b>	<b>24.95</b>	0.6025	<b>23.03</b>	<b>0.6429</b>	<b>25.32</b>	<b>0.7896</b>
DCLS(Ours)		<b>27.14</b>	<b>0.7775</b>	<b>25.37</b>	<b>0.6516</b>	<b>24.99</b>	<b>0.6043</b>	<b>27.13</b>	<b>0.6500</b>	<b>25.57</b>	<b>0.7969</b>
Bicubic+ZSSR [41]	30	19.77	0.2938	19.36	0.2534	19.43	0.2308	18.32	0.2450	19.25	0.3046
IKC [9]		25.27	0.7154	24.15	0.6100	<b>24.06</b>	0.5674	22.11	0.5969	23.80	0.7438
DANv1 [30]		25.32	<b>0.7276</b>	24.15	<b>0.6138</b>	24.04	0.5678	22.08	0.5977	23.82	0.7442
DANv2 [31]		<b>25.36</b>	0.7264	<b>24.16</b>	0.6121	<b>24.06</b>	<b>0.5690</b>	<b>22.14</b>	<b>0.6014</b>	<b>23.87</b>	<b>0.7489</b>
DCLS(Ours)		<b>25.49</b>	<b>0.7323</b>	<b>24.23</b>	<b>0.6131</b>	<b>24.09</b>	<b>0.5696</b>	<b>22.37</b>	<b>0.6119</b>	<b>24.21</b>	<b>0.7582</b>

Table 2. Quantitative comparison on various noisy datasets. The best one marks in red and the second best are in blue.

used as the upscale module. We can jointly optimize the SR network and kernel estimation network as follows:

$$\mathcal{L} = l_1(\hat{\mathbf{k}}, \mathbf{k}_l; \theta_k) + l_1(\hat{\mathbf{x}}, \mathbf{x}; \theta_g) \quad (17)$$

where  $\theta_k$  and  $\theta_g$  are the parameters of kernel estimation network and DCLS reconstruction network, respectively.

## 4. Experiments

### 4.1. Datasets and Implementation Details

Following previous works [9, 30], 3450 2K HR images from DIV2K [1] and Flickr2K [45] are collected as

the training dataset. And we synthesize corresponding LR images with specific degradation kernel settings (e.g., isotropic/anisotropic Gaussian) using Eq. (1). The proposed method is evaluated by PSNR and SSIM [49] on only the luminance channel of the SR results (YCbCr space).

**Isotropic Gaussian kernels.** Firstly, we conduct blind SR experiments on isotropic Gaussian kernels following the setting in [9]. Specifically, the kernel sizes are fixed to  $21 \times 21$ . In training, we uniformly sample the kernel width from range [0.2, 2.0], [0.2, 3.0] and [0.2, 4.0] for SR scale factors 2, 3 and 4, respectively. For testing, we use

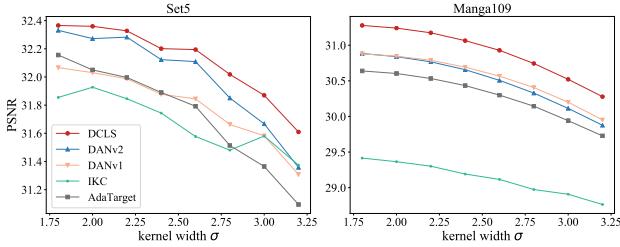


Figure 5. The PSNR performance curves on Set5 and Manga109 of scale factor 4. The kernel width  $\sigma$  are set from 1.8 to 3.2.

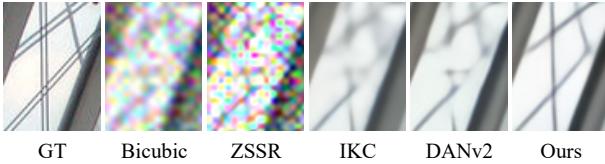


Figure 6. Visual results of *Img 33* from Urban100.

*Gaussian8* [9] kernel setting to generate evaluation dataset from five widely used benchmarks: Set5 [4], Set14 [53], BSD100 [33], Urban100 [12] and Manga109 [34]. *Gaussian8* uniformly chooses 8 kernels from range [0.80, 1.60], [1.35, 2.40] and [1.80, 3.20] for scale factors 2, 3 and 4, respectively. The LR images are obtained by blurring and downampling the HR images with selected kernels.

**Anisotropic Gaussian kernels.** We also conduct experiments on anisotropic Gaussian kernels following the setting in [3]. The kernel size is set to  $11 \times 11$  and  $31 \times 31$  for scale factors 2 and 4, respectively. During training, the anisotropic Gaussian kernels for degradation are generated by randomly selecting kernel width from range (0.6, 5) and rotating from range  $[-\pi, \pi]$ . We also apply uniform multiplicative noise and normalize it to sum to one. For evaluation, we use the DIV2KRK dataset proposed in [3].

**Implementation details.** For all experiments, we use 5 dual-path groups, each containing 10 DPABs with 64 channels. The batch sizes are set to 64 and the LR patch sizes are  $64 \times 64$ . We use Adam [20] optimizer with  $\beta_1 = 0.9$  and  $\beta_2 = 0.99$ . All models are trained on 4 RTX2080Ti GPUs with  $5 \times 10^5$  iterations. The initial learning rate is set to  $4 \times 10^{-4}$  and decayed by half at every  $2 \times 10^{-4}$  iterations. We also augment the training data with random horizontal flips and 90 degree rotations.

## 4.2. Comparison with State-of-the-arts

**Evaluation with isotropic Gaussian kernels.** Following [9], we evaluate our method on datasets synthesized by *Gaussian8* kernels. We compare our method with state-of-the-art blind SR approaches: ZSSR [41] (with bicubic kernel), IKC [9], DANv1 [30], DANv2 [31] and AdaTarget [14]. Following [9], we also conduct comparison with CARN [2] and its variants of performing blind deblurring method [37] before and after CARN. For most methods, we

Method	DIV2KRK [3]			
	$\times 2$		$\times 4$	
	PSNR	SSIM	PSNR	SSIM
Bicubic	28.73	0.8040	25.33	0.6795
Bicubic+ZSSR [41]	29.10	0.8215	25.61	0.6911
EDSR [28]	29.17	0.8216	25.64	0.6928
RCAN [59]	29.20	0.8223	25.66	0.6936
DBPN [10]	29.13	0.8190	25.58	0.6910
DBPN [10]+Correction [13]	30.38	0.8717	26.79	0.7426
KernelGAN [3]+SRMD [56]	29.57	0.8564	27.51	0.7265
KernelGAN [3]+ZSSR [41]	30.36	0.8669	26.81	0.7316
IKC [9]	-	-	27.70	0.7668
DANv1 [30]	32.56	0.8997	27.55	0.7582
DANv2 [31]	<b>32.58</b>	<b>0.9048</b>	<b>28.74</b>	<b>0.7893</b>
AdaTarget [14]	-	-	28.42	0.7854
KOALAnet [19]	31.89	0.8852	27.77	0.7637
DCLS(Ours)	<b>32.75</b>	<b>0.9094</b>	<b>28.99</b>	<b>0.7946</b>

Table 3. Quantitative comparison on DIV2KRK. The best one marks in red and the second best are in blue.

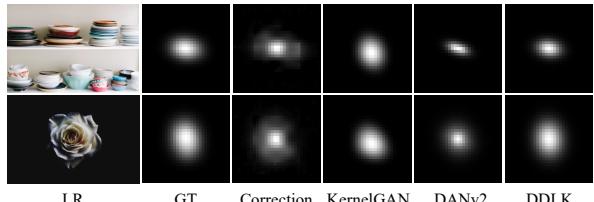


Figure 7. Visual results of estimated kernels of *Img 33* and *Img 43* from DIV2KRK [3] by various kernel estimation methods.

DIV2KRK $\times 4$	KernelGAN	CorrFilter	DANv2	DDLK
LR-PSNR $\uparrow$	41.28	41.35	45.06	<b>45.27</b>
Kernel-MSE $\downarrow$	0.1518	0.1392	0.0817	<b>0.0574</b>

Table 4. Quantitative evaluation on the performance of DDLK.

use their official implementations and pre-trained models.

The quantitative results are shown in Table 1. It is obvious that our method leads to the best performance over all datasets. The bicubic SR model CARN suffers severe performance drop with *Gaussian8* which deviates from the predefined bicubic kernel. Performing deblurring on the super-resolved image can improve the results. ZSSR achieves better performance compared with non-blind SR method but is limited by the image-specific network design (cannot utilize abundant training data). AdaTarget can improve image quality but is still inferior to that of blind SR methods. IKC and DAN are two-step blind SR methods and can largely improve the results. However, both of them predict kernel embedding and directly involve it into the network, which damages the spatial relation of the kernel and thus performs inferior to our method. We also provide the comparison of PSNR values on different datasets with blur kernels width from 1.8 to 3.2 as shown in Fig. 5. DCLS performs the best result over all different kernel widths. The qualitative results shown in Fig. 8 illustrate that DCLS can produce clear and pleasant SR images. Furthermore, we conduct an ex-

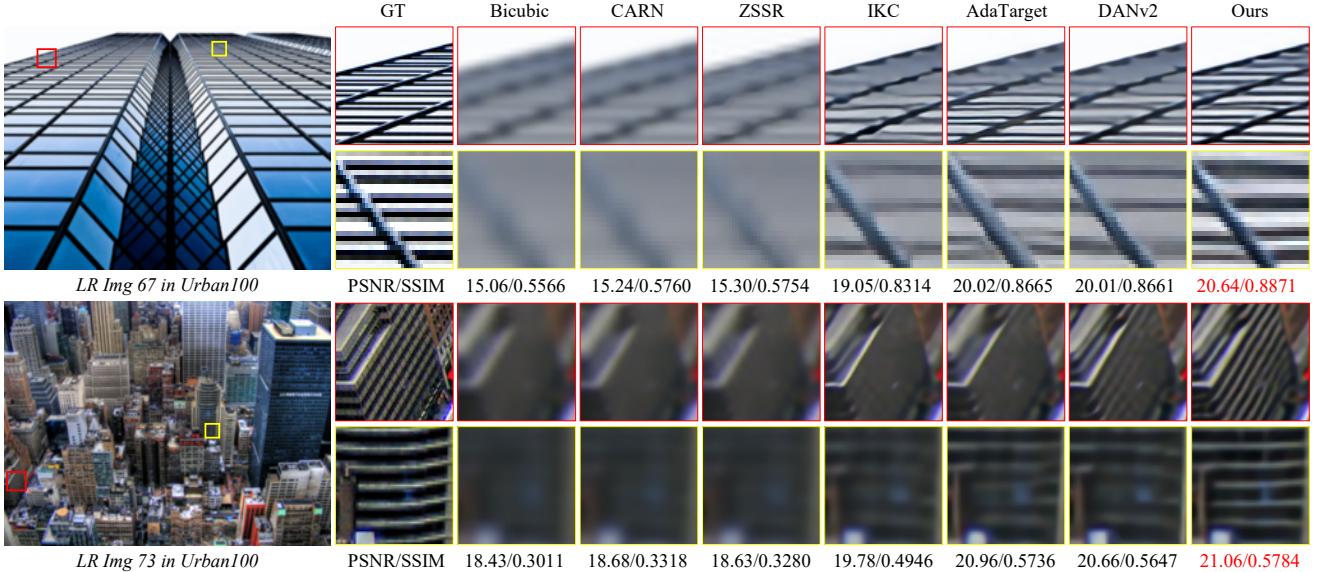


Figure 8. Visual results of *Img 67* and *Img 73* in Urban100 [12], for scale factor 4 and kernel width 2.6. Best viewed in color.

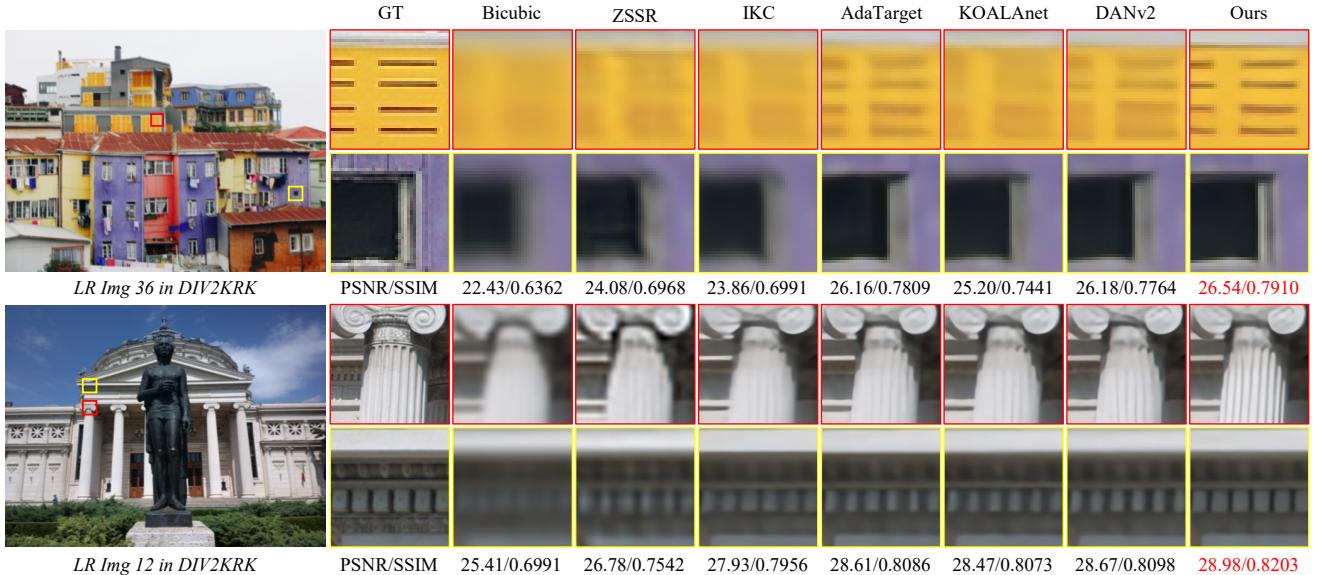


Figure 9. Visual results of *Img 36* and *Img 12* in DIV2KRK [3], for scale factor of 4. Best viewed in color.

periment of super-resolving images with additional noise. As shown in Table 2 and Fig. 6, DCLS still outperforms other methods over all datasets with different noise levels.

**Evaluation with anisotropic Gaussian kernels.** Degradation with anisotropic Gaussian kernels are more general and challenging. Similar to isotropic kernel, we firstly compare our method with SOTA blind SR approaches such as ZSSR [41], IKC [9], DANv1 [30], DANv2 [31], AdaTarget [14] and KOALAnet [19]. We also compare DCLS with some SOTA bicubically designed methods such as EDSR [28], RCAN [59], and DBPN [10]. And we provide Correction [13] for DBPN. In addition, we combine a kernel estimation method (e.g. KernelGAN [3]) with other

non-blind SR methods, such as ZSSR [41] and SRMD [56], as two-step solutions to solve blind SR.

Table 3 shows the quantitative results on DIV2KRK [3]. It can be seen that the proposed DCLS significantly improves the performance compared with other blind SR approaches. Note that ZSSR performs better when combined with KernelGAN, which indicates that good kernel estimation can help a lot. Recent SOTA blind SR methods such as IKC, DAN and KOALAnet can achieve remarkable accuracy in PSNR and SSIM. By applying an adaptive target to finetune the network, AdaTarget can perform comparably with SOTA blind methods. However, all of those methods are still inferior to the proposed DCLS. The visual re-

SLK	DDLK	Stretching Strategy	DCLS	DPAN	DIV2KRK	
			Deconv		PSNR	SSIM
✓	-	✓	-	✓	28.84	0.7921
-	✓	✓	-	✓	28.86	0.7924
✓	-	-	✓	✓	28.94	0.7946
-	✓	-	✓	-	28.94	0.7938
-	✓	-	✓	✓	28.99	0.7964

Table 5. Ablation study on our vital components.

Method	Wiener <sub>Fea</sub> [7]		CLS <sub>Fea</sub>		DCLS <sub>Fea</sub>	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Set5	32.05	0.8878	31.98	0.8862	<b>32.12</b>	<b>0.8890</b>
Set14	28.38	0.7709	28.29	0.7658	<b>28.54</b>	<b>0.7728</b>
BSD100	27.47	0.7238	27.48	0.7216	<b>27.60</b>	<b>0.7285</b>
Urban100	26.07	0.7775	26.03	0.7768	<b>26.15</b>	<b>0.7809</b>
Manga109	30.77	0.9069	30.65	0.9040	<b>30.86</b>	<b>0.9086</b>
DIV2KRK	28.77	0.7886	28.92	0.7921	<b>28.99</b>	<b>0.7947</b>

Table 6. Quantitative comparison on various datasets. Fea means applying deconvolution on the feature space.

DIV2KRK × 4	Wiener <sub>RGB</sub>	CLS <sub>RGB</sub>	DCLS <sub>RGB</sub>	DCLS <sub>Fea</sub>
PSNR	28.91	28.90	28.94	28.99
SSIM	0.7941	0.7935	0.7941	0.7964

Table 7. Quantitative results. RGB and Fea mean applying deconvolution in the RGB space and feature space, respectively.

sults on DIV2KRK are shown in Fig. 9. As we can see, the SR images produced by our method are much sharper and cleaner. We also provide the results of kernel estimation and downsampling HR image with estimated kernel in Fig. 7 and Table 4. Compared with previous image-specific methods such as KernelGAN [3] and Correction Filter [13], the dynamic deep linear kernel (DDLK) is more flexible and capable of producing accurate kernels.

### 4.3. Analysis and Discussions

**Ablation Study.** We conduct ablation studies on vital components of our method: DPAN, DDLK and DCLS deconvolution. The quantitative results on DIV2KRK are exported in Table 5. Note that the baseline model with DPAN eliminates artifacts from kernel and thus improves the result. And the DCLS deconvolution can further make use of the estimated kernel and high-level information from deep features to achieve a higher performance (+0.15dB from baseline).

**Effectiveness of the DCLS deconvolution.** To illustrate the effectiveness of DCLS, we include a comparison of substituting DCLS with other deblurring methods, such as traditional constrained least squares (CLS) and Wiener deconvolution [7, 50] in the RGB space and feature space. The results are presented in Table 6 and Table 7. By applying deconvolution in the RGB space with the reformulated kernel, we can get a clear LR image and thus improve the SR performance. This idea is similar to Correction Filter [13], but with one big difference, in that our estimator is highly

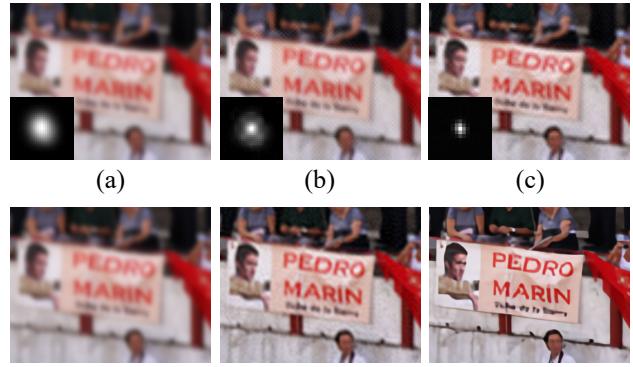


Figure 10. Applying DCLS in the RGB space. (a) Original LR & kernel, (b) corrected LR & estimated kernel by [13], (c) deblurred LR & estimated kernel by the proposed method.



Figure 11. Comparison of historic image [21] for 4× SR.

correlated to the LR image rather than the SR model. The visual example is shown in Fig. 10.

**Performance on Real Degradation** To further demonstrate the effectiveness of our method, we apply the proposed model on real degradation data where the ground truth HR images and the blur kernels are not available. An example of super-resolving historic image is shown in Fig. 11. Compared with LapSRN [21] and DANv2 [31], our DCLS can produce sharper edges and visual pleasing SR results.

### 5. Conclusion

In this work, we have presented a well-principled algorithm to tackle the blind SR problem. We first derive a new form of blur kernel in the low resolution space from classical degradation model. We then propose to estimate and apply that kernel in HR image restoration. Subsequently, a dynamic deep linear kernel (DDLK) module is introduced to improve kernel estimation. We further design a deep constrained least squares (DCLS) deconvolution module that integrates blur kernel and LR image in the feature domain to obtain the clean feature. The clean feature and the primitive feature are then fed into a dual-path network to generate the super-resolved image. Extensive experiments on various kernels and noises demonstrate that the proposed method leads to a state-of-the-art blind SR performance.

**Acknowledgment** This work was supported by the National Natural Science Foundation of China (NSFC) under grants No.61872067 and No.61720106004.

## References

- [1] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *Proc. CVPRW*, pages 126–135, 2017. 5
- [2] Namhyuk Ahn, Byungkon Kang, and Kyung-Ah Sohn. Fast, accurate, and lightweight super-resolution with cascading residual network. In *Proc. ECCV*, pages 252–268, 2018. 5, 6
- [3] Sefi Bell-Kligler, Assaf Shocher, and Michal Irani. Blind super-resolution kernel estimation using an internal-gan. In *Proc. NeurIPS*, pages 284–293, 2019. 1, 2, 3, 6, 7, 8
- [4] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie line Alberi Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In *Proc. BMVC*, pages 135.1–135.10, 2012. 5, 6
- [5] Tao Dai, Jianrui Cai, Yongbing Zhang, Shu-Tao Xia, and Lei Zhang. Second-order attention network for single image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11065–11074, 2019. 2
- [6] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In *European conference on computer vision*, pages 184–199. Springer, 2014. 1, 2
- [7] Jiangxin Dong, Stefan Roth, and Bernt Schiele. Deep wiener deconvolution: Wiener meets deep learning for image de-blurring. *Proc. NeurIPS*, 33:1048–1059, 2020. 4, 8
- [8] Manuel Fritzsche, Shuhang Gu, and Radu Timofte. Frequency separation for real-world super-resolution. In *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pages 3599–3608. IEEE, 2019. 1
- [9] Jinjin Gu, Hannan Lu, Wangmeng Zuo, and Chao Dong. Blind super-resolution with iterative kernel correction. In *Proc. CVPR*, pages 1604–1613, 2019. 1, 2, 3, 4, 5, 6, 7
- [10] Muhammad Haris, Gregory Shakhnarovich, and Norimichi Ukita. Deep back-projection networks for super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1664–1673, 2018. 1, 2, 6, 7
- [11] CK Huang and Hsiao-Hsian Nien. Multi chaotic systems based pixel shuffle for image encryption. *Optics communications*, 282(11):2123–2127, 2009. 4
- [12] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *Proc. CVPR*, pages 5197–5206, 2015. 5, 6, 7
- [13] Shady Abu Hussein, Tom Tirer, and Raja Giryes. Correction filter for single image super-resolution: Robustifying off-the-shelf deep super-resolvers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1428–1437, 2020. 2, 6, 7, 8
- [14] Younghyun Jo, Seoung Wug Oh, Peter Vajda, and Seon Joo Kim. Tackling the ill-posedness of super-resolution through adaptive target generation. In *Proc. CVPR*, pages 16236–16245, 2021. 1, 5, 6, 7
- [15] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European conference on computer vision*, pages 694–711. Springer, 2016. 2
- [16] Kenji Kawaguchi. Deep learning without poor local minima. In *Proc. NeurIPS*, pages 586–594, 2016. 3
- [17] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1646–1654, 2016. 1, 2
- [18] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Deeply-recursive convolutional network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1637–1645, 2016. 2
- [19] Soo Ye Kim, Hyeyonjun Sim, and Munchurl Kim. Koalanet: Blind super-resolution using kernel-oriented adaptive local adjustment. In *Proc. CVPR*, pages 10611–10620, 2021. 1, 6, 7
- [20] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 6
- [21] Wei-Sheng Lai, Jia-Bin Huang, Narendra Ahuja, and Ming-Hsuan Yang. Deep laplacian pyramid networks for fast and accurate super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 624–632, 2017. 2, 8
- [22] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017. 1, 2
- [23] Guang-He Lee, David Alvarez-Melis, and Tommi S Jaakkola. Towards robust, locally linear deep networks. In *Proc. ICLR*, 2018. 4
- [24] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Understanding and evaluating blind deconvolution algorithms. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1964–1971. IEEE, 2009. 2
- [25] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *CVPR 2011*, pages 2657–2664. IEEE, 2011. 2
- [26] Zhen Li, Jinglei Yang, Zheng Liu, Xiaomin Yang, Gwanggil Jeon, and Wei Wu. Feedback network for image super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3867–3876, 2019. 1
- [27] Jingyun Liang, Kai Zhang, Shuhang Gu, Luc Van Gool, and Radu Timofte. Flow-based kernel prior with application to blind super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10601–10610, 2021. 2
- [28] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *Proc. CVPRW*, pages 136–144, 2017. 1, 2, 6, 7
- [29] Andreas Lugmayr, Martin Danelljan, Luc Van Gool, and Radu Timofte. Srfow: Learning the super-resolution space with normalizing flow. In *European Conference on Computer Vision*, pages 715–732. Springer, 2020. 2

- [30] Zhengxiong Luo, Yan Huang, Shang Li, Liang Wang, and Tieniu Tan. Unfolding the alternating optimization for blind super resolution. In *Proc. NeurIPS*, 2020. 1, 2, 3, 5, 6, 7
- [31] Zhengxiong Luo, Yan Huang, Shang Li, Liang Wang, and Tieniu Tan. End-to-end alternating optimization for blind super resolution. *arXiv preprint arXiv:2105.06878*, 2021. 1, 3, 4, 5, 6, 7, 8
- [32] Ziwei Luo, Lei Yu, Xuan Mo, Youwei Li, Lanpeng Jia, Hao- qiang Fan, Jian Sun, and Shuaicheng Liu. Ebsr: Feature enhanced burst super-resolution with deformable alignment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 471–478, 2021. 2
- [33] 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 *Proc. ICCV*, pages 416–423, 2001. 5, 6
- [34] Yusuke Matsui, Kota Ito, Yuji Aramaki, Azuma Fujimoto, Toru Ogawa, Toshihiko Yamasaki, and Kiyoharu Aizawa. Sketch-based manga retrieval using manga109 dataset. *Multimedia Tools and Applications*, 76(20):21811–21838, 2017. 5, 6
- [35] Tomer Michaeli and Michal Irani. Nonparametric blind super-resolution. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 945–952, 2013. 2
- [36] Guido Montúfar, Razvan Pascanu, Kyunghyun Cho, and Yoshua Bengio. On the number of linear regions of deep neural networks. In *Proc. NeurIPS*, pages 2924–2932, 2014. 4
- [37] Jinshan Pan, Deqing Sun, Hanspeter Pfister, and Ming- Hsuan Yang. Deblurring images via dark channel prior. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 40(10):2315–2328, 2017. 5, 6
- [38] Sung Cheol Park, Min Kyu Park, and Moon Gi Kang. Super- resolution image reconstruction: a technical overview. *IEEE signal processing magazine*, 20(3):21–36, 2003. 2
- [39] Andrew M Saxe, James L McClelland, and Surya Ganguli. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. *arXiv preprint arXiv:1312.6120*, 2013. 3
- [40] Wenzhe Shi, Jose Caballero, Ferenc Huszár, Johannes Totz, Andrew P Aitken, Rob Bishop, Daniel Rueckert, and Zehan Wang. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1874–1883, 2016. 2
- [41] Assaf Shocher, Nadav Cohen, and Michal Irani. “zero- shot” super-resolution using deep internal learning. In *Proc. CVPR*, pages 3118–3126, 2018. 1, 5, 6, 7
- [42] Jae Woong Soh, Sunwoo Cho, and Nam Ik Cho. Meta- transfer learning for zero-shot super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3516–3525, 2020. 2
- [43] Ying Tai, Jian Yang, and Xiaoming Liu. Image super- resolution via deep recursive residual network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3147–3155, 2017. 2
- [44] Guangpin Tao, Xiaozhong Ji, Wenzhuo Wang, Shuo Chen, Chuming Lin, Yun Cao, Tong Lu, Donghao Luo, and Ying Tai. Spectrum-to-kernel translation for accurate blind image super-resolution. In *Thirty-Fifth Conference on Neural Information Processing Systems*, 2021. 3
- [45] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming- Hsuan Yang, and Lei Zhang. Ntire 2017 challenge on single image super-resolution: Methods and results. In *Proc. CVPRW*, pages 114–125, 2017. 5
- [46] Patrick Vandewalle, Sabine Süsstrunk, and Martin Vetterli. A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP journal on advances in signal processing*, 2006:1–14, 2006. 2
- [47] Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 606–615, 2018. 2
- [48] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops*, pages 0–0, 2018. 1, 2
- [49] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. on Image Processing*, 13(4):600–612, 2004. 5
- [50] Norbert Wiener et al. *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*, volume 8. MIT press Cambridge, MA, 1964. 8
- [51] Yu-Syuan Xu, Shou-Yao Roy Tseng, Yu Tseng, Hsien-Kai Kuo, and Yi-Min Tsai. Unified dynamic convolutional network for super-resolution with variational degradations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12496–12505, 2020. 2
- [52] Xin Yu and Fatih Porikli. Ultra-resolving face images by discriminative generative networks. In *European conference on computer vision*, pages 318–333. Springer, 2016. 2
- [53] Roman Zeyde, Michael Elad, and Matan Protter. On single image scale-up using sparse-representations. In *International Conference on Curves and Surfaces*, pages 711–730, 2010. 5, 6
- [54] Kai Zhang, Luc Van Gool, and Radu Timofte. Deep unfolding network for image super-resolution. In *Proc. CVPR*, pages 3217–3226, 2020. 2
- [55] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Trans. on Image Processing*, 26(7):3142–3155, 2017. 2
- [56] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Learning a single convolutional super-resolution network for multiple degradations. In *Proc. CVPR*, pages 3262–3271, 2018. 2, 3, 6, 7
- [57] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Deep plug- and-play super-resolution for arbitrary blur kernels. In *Proc. CVPR*, pages 1671–1681, 2019. 1, 2, 3
- [58] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of

- deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018. 2
- [59] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Proc. ECCV*, pages 286–301, 2018. 1, 2, 6, 7
- [60] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *Proc. ECCV*, pages 294–310, 2018. 4
- [61] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image super-resolution. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2472–2481, 2018. 2
- [62] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu. Residual dense network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(7):2480–2495, 2020. 2