

Deep Priors inside an Unrolled and Adaptive Deconvolution Model

Hung-Chih Ko, Je-Yuan Chang, and Jian-Jiun Ding

National Taiwan University
{r06942148, r07942158, jjding}@ntu.edu.tw

Abstract. Image deconvolution is an essential but ill-posed problem even if the degradation kernel is known. Recently, learning based methods have demonstrated superior image restoration quality in comparison to traditional methods which are typically based on empirical statistics and parameter adjustment. Though coming up with outstanding performance, most of the plug-and-play priors are trained in a specific degradation model, leading to inferior performance on restoring high-frequency components. To address this problem, a deblurring architecture that adopts (1) adaptive deconvolution modules and (2) learning based image prior solvers is proposed. The adaptive deconvolution module adjusts the regularization weight locally to well process both smooth and non-smooth regions. Moreover, a cascade made of image priors is learned from the mapping between intermediates thus robust to arbitrary noise, aliasing, and artifact. According to our analysis, the proposed architecture can achieve a significant improvement on the convergence rate and result in an even better restoration performance.

1 Introduction

Most of the pictures captured by hand-held devices are easily suffering from camera motion or out-of-focus which leads to blurry observed images. In general, the blurry observed image B is modeled as

$$B = I \otimes K + N \quad (1)$$

where \otimes denotes convolution, I , K , and N refers to the latent clear image, the point spread function (PSF), and noise, respectively. Non-blind deblurring is to retrieve the clear image I when the PSF is already estimated from lens' parameters or motion sensors. Early approaches such as the Wiener filter [1] and the Richardson-Lucy algorithm [2,3] have become a mainstream in the past century; however, they may suffer from ringing artifacts or over-smoothness when dealing with severely interfered cases. In recent years, various image priors are widely adopted to serve as regularization terms in image deconvolution problems. A good image prior not only promotes edge sharpness but also suppresses artifacts. In some early works like total variation (TV) [4] and hyper-Laplacian [5,6], the regularization term is formulated as a norm of derivative of intensities to encourage the gradient to approximate a heavy-tailed distribution. Alternative priors

like K-SVD [7], Gaussian mixture model (GMM) [8] and Markov random field (MRF) [9] have provided spatial support for exploring latent image statistics. Despite coming up with practical solutions, most of the aforementioned methods still highly rely on empirical statistics and knowledge for parameter tuning.

Recently, with the successes in the field of computer vision, neural networks have been used for challenging image restoration tasks including image inpainting, denoising, and deblurring. Several image deblurring works based on discriminative models [10–17] or generative models [18,19] have demonstrated with superiority. Among learning based methods, [10,11,14–17] have focused on improving optimization process through the integration of deep learning networks, instead of conventional end-to-end training procedure. Despite of their surprising performance, models trained on specifically designated degradation *i.e.* a fixed noise level are typically exploited to handle intermediate results within optimization. Therefore, it is inevitable to result in a limited performance when suffering from arbitrary noise or artifacts.

To address this problem, an unrolled deconvolution network (UDN) is proposed. Different from most of the existing works which aim to optimize an image prior within a narrow degradation domain, we further improve the half-quadratic splitting (HQS) [20] optimization process and encourage those priors to be learned from arbitrary intermediate outputs as well as aliasing which possibly emerges. Therefore, UDN can implicitly learn image priors during optimization and be more robust to severely degraded cases.

We also note that conventional deconvolution within HQS may have ignored the spatial variety in an image. When set parameters globally, it can lead the restored image to be over-smoothed in some region and under-smoothed in the others. To address this problem, we establish the adaptive deconvolution module (ADM) which is essentially composed of a small set of convolution layers for edge detection and determines the parameters adaptively. With the aid of increased capacity via the UDN, the models are expected to learn corresponding artifacts possibly different from the conventional scheme. We have combined the results of ADM and UDN by a simple and intuitive way that ensures both smoothness and detail are successfully preserved.

Through our analysis, the proposed methods have outperformed other existing works in PSNR and SSIM. Especially, when suffering from significant noise, the high-frequency components are successfully restored, which is not observed in other works. Furthermore, since the proposed UDN optimizes the image priors within intermediates, a great improvement on the convergence rate is achieved. It leads to a more efficient and faster restoration process.

2 Related Works

In the field of computer vision, non-blind deconvolution has long been an issue that featured with an ill-posed nature. The design on image prior is typically inferred from a Maximum a Posteriori (MAP) perspective which have been widely adopted for inverse problems; the prior always matter during the optimization

on a MAP-based objective function. Among those early works, most of the priors are designed by empirical statistics on natural images. For example, TV regularization [4] and hyper-Laplacian prior [5, 6] have shown success on capturing the heavy-tailed gradient distribution that usually failed when using Gaussian priors. Alternatives like mixture of Laplacians [21] and Gaussians [22] are adopted to approximate the gradient sparsity prior with the increased generality; and a content-aware local gradient prior [23] has also been emphasized to handle the spatial variety. Most of the works are solved by gradient descent, the alternating direction method of multiplier (ADMM) [24], or HQS [20], so that an inevitable increment on computation loads is expected due to the iterative process.

Other learning based methods like [9, 25] has adopted fields of experts (FoE) to approximate the heavy-tailed gradient distribution by learned filters that capture local image statistics and further extended by cascade of shrinkage fields (CSF) [26], regression tree based models [27] and trainable non-linear reaction diffusion (TNRD) [28]. GMM are also developed to fit the image natures like expected patch log-likelihood (EPLL) [8] and its extended version which resorts multi-scale patch-pyramid modeling [29] for performance gains; this kind of works have demonstrated a powerful capacity on noise or blur removal. However, although patch-based methods achieve good results, solving these problems are usually time-consuming as well.

More recently, discriminative models [10–17] have shown significant improvement on restoration quality. An early work that using multi-layer perceptron (MLP) [12] can successfully remove the corresponding artifacts appeared after deconvolution by estimating clear images from an initial estimation by Gaussian prior. A convolution neural network (CNN) [13] model that based on the knowledge of kernel decomposition is trained to suppress various outliers like saturation, boundary effect or compression artifacts. However, both methods require fine-tune over each specific kernel, thus limited by a loss of generality.

Some of the works have decoupled image restoration problem into optimizing data fidelity term and regularization term, so that simpler subproblems are yielded. These methods named as plug-and-play frameworks [30, 31] have exploited the concepts from variable splitting techniques like ADMM and HQS, and demonstrated a flexibility on handling various restoration tasks. Started from this concept, some frameworks integrated with deep learning models have been proposed [11, 14, 15, 17, 18] since it is regraded that models trained on pairs of degraded and latent images can learn the image priors implicitly.

A Fourier deconvolution network (FDN) [11] and a fully convolutional neural network (FCN) [14] are proposed to reduce noise after each deconvolution stage by learning CNN-based priors in gradient domain. A GMM-based method [32] learns a data fidelity term from intermediate results within cascaded shrinkage function models, and a work [17] proposes to simultaneously learn fidelity and regularization term in a task-driven manner. These methods are dedicated for aliasing removal in non-blind deconvolution cases. On the other hand, some of the works have handled these problems in a more generalized perspective. Image restoration CNN (IRCNN) [15] is composed of a set of learned CNN

denoisers that can be plugged into the HQS optimization framework [5, 33]. A denoising prior driven network [34] is proposed for a variety of restoration tasks. A CNN-based model is trained in an adversarial way [18] to encourage the optimal solution lies into natural image domain.

As noted in [10], most of the plugged-in learning based models which learn independently to the form of degradation are adopted to handle those tasks within intermediate optimization process. Although with increased flexibility, the learned priors cannot always react accurately to arbitrary perturbation lasted in the intermediates. For example, the denoising prior is originally trained to remove noise with homogenous power spectrum. However, as iterative deconvolution is applied to a blurry observation, the power spectrum among the outputs is expected to be more heterogenous, which is possibly out of the learned mappings. Besides, [29] has found several learning based methods are biased toward smoothness since the dominance of smooth gradient within natural images, so that the retrieval of coherent high frequency details remains an important challenge. In our investigation on non-blind deblurring, although IRCNN [15] generates outstanding performance for assessment, the restored images have lost a lot of details.

Basically, our proposed work has followed a HQS optimization framework. We first reformulate the deconvolution stage with adaptation (ADM), which only requires neglectable increment on computation complexity. Secondly, a fully convolutional network was proposed to unroll the HQS optimization. The rationale is to enlarge the learning domain with intermediate results, so that the priors can learned to handle the arbitrary deformation within optimization.

3 Unrolled Deconvolution Network

The proposed work is based on HQS optimization for a MAP-based problem that aims to estimate the latent clear image \hat{I} from blurry observed image B by maximizing the posterior probability $P(\cdot)$

$$\hat{I} = \arg \max_I P(I|B, K) = \arg \max_I P(B|I, K)P(I) \quad (2)$$

where $P(B|I, K)$ models the conditional probability distribution of B given I and K and $P(I)$ is the prior for latent clear images. The overall function can be solved in a negative log-likelihood scheme as both of the residual of blurry observation to clear image and the prior term are modeled as Gaussian distributions that can be used to measure the penalty. The objective function derived from (2) can be

$$\hat{I} = \arg \min_I \frac{1}{2\sigma_d^2} \|B - I \otimes K\|_2^2 + \frac{1}{2\sigma_p^2} \Phi(I) \quad (3)$$

$$= \arg \min_I \frac{1}{2} \|B - I \otimes K\|_2^2 + \frac{\lambda}{2} \Phi(I) \quad (4)$$

where the first term is the data fidelity and the second is the image prior term denoted as $\Phi(I)$. Note that $\lambda = \sigma_d^2/\sigma_p^2$, where the constant σ_d^2 and σ_p^2 are the

variances of data term and prior term originated from the Gaussian modeling and can be reduced to a weight on prior term λ that controls the trade-off between fidelity and regularization. Under the HQS optimization scheme, an auxiliary variable Z is introduced to (3) to decouple I from concurrent minimization so that the objective problem becomes

$$\arg \min_{I, Z} \frac{1}{2} \|B - I \otimes K\|_2^2 + \frac{\beta}{2} \|I - Z\|_2^2 + \frac{\lambda}{2} \Phi(Z) \quad (5)$$

where β is a weight that controls the penalty on the fidelity between Z and I . According to the alternating minimization algorithm [5, 33], the optimization process can be separated into I and Z subproblems described in (6) and (7), respectively. β is increased over iterations to encourage convergence.

$$\hat{I} = \arg \min_I \frac{1}{2} \|B - I \otimes K\|_2^2 + \frac{\beta}{2} \|I - Z\|_2^2 \quad (6)$$

$$\hat{Z} = \arg \min_Z \frac{\beta}{2} \|I - Z\|_2^2 + \frac{\lambda}{2} \Phi(Z) \quad (7)$$

As a quadratic regularized least-squared problem, (6) exists a FFT-based closed-form solution. On the other hand, although the solution of (7) is dependent to arbitrary image priors $\Phi(\cdot)$, it can be reduced to the following equation, which is equivalent to the denoising problem with $\sigma = \sqrt{\lambda/\beta}$ described in the CNN-based image restoration algorithms in [15, 34]:

$$\hat{Z} = \arg \min_Z \frac{1}{2(\sqrt{\lambda/\beta})^2} \|I - Z\|_2^2 + \frac{1}{2} \Phi(Z) \quad (8)$$

Hence, as a plug-and-play scheme, the CNN is designed to learn the prior implicitly and benefits the optimization process.

3.1 Adaptive Deconvolution Module

Look back into the I subproblem in (6), we have noted that the essence of β is to balance (a) the fidelity between restored I and observed B and (b) the fidelity of restored I and auxiliary variable Z which is derived from the prior-regulated subproblem (7). Actually, β poses significant influence on restored image quality. In most of the conventional methods, the value is preset empirically or complied with some rules according to the noise level. However, inferior performance can be observed on cases of unnatural images or those interfered with extremely high noise level. Besides, the weight is typically applied in a global manner so that it may have ignored the spatial variety over the entire images.

By intuition, as the value of β is set large, the solution is forced to comply with Z which is a prior-regulated solution, such that a smoothed outcome is expected. On the other hand, when β is set small, since the regularization is diminished, the solution is expected to be closed to that derived from direct inverse filter, which behaves well on restoring high-frequency components but as a trade-off, it also becomes easier to suffer from noise and other artifacts.

We have shown a simulation result in Fig. 1. When deconvolving with a smaller β , the noise becomes dominant while preserves more fine structures, as shown in Fig. 1(b,d). When β is larger, the noise is apparently suppressed with a loss of details, as shown in Fig. 1(a,c). Therefore, β can be set smaller in edge regions to better approximate the steep gradient in ground truth. On the other hand, β can be set larger at flat regions since it demonstrates a good ability for smoothness. Therefore, instead of setting β to be a global scalar, a spatially varying weight is expected to achieve a better performance.

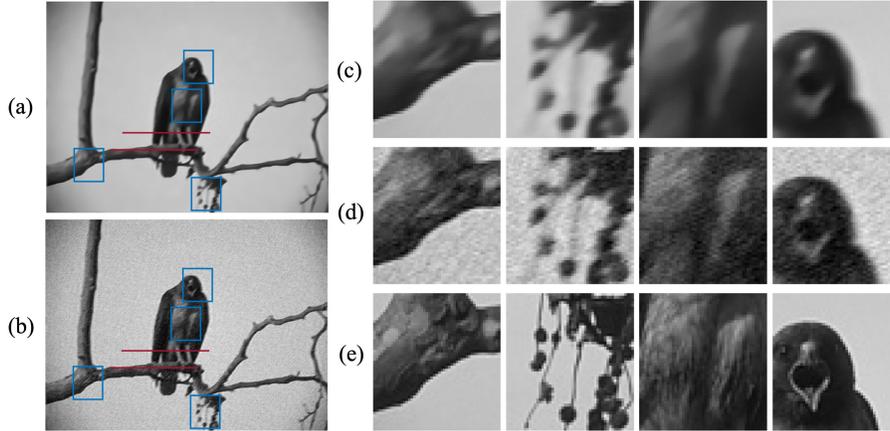


Fig. 1. Deconvolution results. The parameter β adopted in (a) is five-fold larger than that in (b). (c) and (d) are patches corresponding to (a) and (b), respectively. (e) the ground truth.

If we formulate β in (6) as a 2-d spatially varying matrix, one can suffer from difficulties when calculating the closed-form solution. The cause is ascribed to the element-wise multiplication out of the quadratic term lasting convolving operation when solving by FFT. Although it is still possible to solve patch by patch and extend the weight into a Toeplitz matrix for convolution, it seems not practical to implement.

Instead, we propose the ADM, an approximated version with spatially varying weights. To be specific, we solve I in two different fidelity weights and aggregate two solutions by an edge awareness criterion such that

$$I(\hat{x}) = \begin{cases} I_{edge}(\hat{x}), & \text{when } Z(x) \in edge. \\ I_{\sim edge}(\hat{x}), & \text{when } otherwise. \end{cases} \quad (9)$$

where x denotes the pixel index and \hat{I}_n is defined as

$$I_{\{edge, \sim edge\}} = \mathcal{F}^{-1} \left(\frac{\mathcal{F}(k)^* \mathcal{F}(B) + \beta_{\{edge, \sim edge\}} \mathcal{F}(Z)}{\mathcal{F}(k) \mathcal{F}(k)^* + \beta_{\{edge, \sim edge\}}} \right) \quad (10)$$

where the two weights follow a linear relationship $\beta_{edge} = \alpha \beta_{\sim edge}$ and we set $0 < \alpha < 1$, $\beta_{\sim edge} = \beta$; β is determined as described in Section 3.2. The

implementation detail about edge awareness criterion and determination rule for hyperparameters are included in supplementary material.

In practice, Z which is the optimal solution in (7) is sent to the edge detector to provide a binary edge map. α is set in an exponentially descending sequence over iterations since the noise component is usually dominant at early iterations, we regard to encourage the detail preservation in a gradual manner. An edge awareness criterion in (9) is established by a conventional multi-scaled Gaussian-smoothed edge detector. The rationale of Gaussian smoothness is also under the consideration of false detection of noise components as edges. Multi-scaling is designed to increase the detection capacity on edges with various scales.

The comparison on native deconvolution and the ADM is displayed in Table 1. At early stages, the ADM does not achieve a better result as most of the regions are dominated by the amplified noise. When optimization closes to convergence, we surprisingly find ADM leads to a better approximation to the ground truth, especially for some of the spike-liked signals, so that a better solution with lower MSE is achieved (refer to signal visualization in supplementary material). Another evidence is provided in Fig. 2 which records the PSNR of outputs from I and Z subproblems. In I subproblem, though ADM has larger MSE at first, it converges to a more optimal solution than native. On the other hand, in Z subproblem, an apparent performance gain is shown within almost all stages.

In summary, we have found the potential of adaptive deconvolution to retain detailed structures. Although an inferior result can be found in a single deconvolution stage, the performance between two methods are pulled apart as the HQS iteration goes on. A possible rationale to this is the incorporation of high-frequency components *i.e.* the detailed structures has suppressed the over-smoothing effect usually found in the late iterations. Furthermore, the restoration is achieved in an accumulative manner, so that even a slight change in HQS loop can result in totally different performance in the end.

Table 1. The corresponding mean squared error (MSE) over iterations for two 1D signals indicated in the red lines in Fig. 1(a-b). The bold numbers indicate lower MSE.

Native/ADM(10^{-3})	$i = 1$	$i = 10$	$i = 20$	$i = 30$
Signal-1	17.428 /18.695	3.241/ 2.937	2.752/ 2.346	2.634/ 2.146
Signal-2	13.587 /14.589	3.825 /3.933	4.128/ 3.789	4.295/ 3.794

3.2 Learning Image Priors

The proposed UDN is consisted of a number of ADM and Z submodule pairs. The architecture of Z submodule is illustrated in Fig. 3(a) and the overall structure of UDN is sketched in Fig. 3(b). The Z submodules are designed to remove the corresponding noise and artifacts inferred from previous modules. The CNN architecture is basically inspired by the idea from IRCNN [15], but additional skip connections between dilated convolutional layers are introduced, since we find them beneficial for training stability and performance, especially for a deeply stacked scheme like UDN.

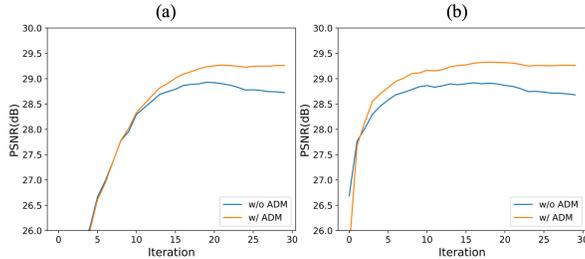


Fig. 2. The convergence process between native deconvolution and ADM on PSNR evaluation for the output of (a) I subproblem and (b) Z subproblem.

Note that the ADM is essentially a learning-free module that only requires presetting on β at each layer. In fact, β is determined implicitly by an aforementioned relationship $\sigma = \sqrt{\lambda/\beta}$. The denoising level σ is set exponentially decreasing from 49 to 5 among iterations. We also find $\lambda \propto \sigma_d^2$ by assuming the variance between latent image to prior is unique among all cases so that λ is related to the squared initial perturbed noise level. Hence, β is set increasing with iterations to encourage the tendency to a regularized solution.

In this paper, the number of pairs inside the UDN is set 8 and 10. While excluding the non-parametric ADMs and batch normalization layers, this model is composed of either $8 \times 7 = 56$ or $10 \times 7 = 70$ layers which easily lead to difficulties to train. However, with all kernels composed of size 3×3 that enlarges the receptive field without parameter overhead and additional skip connections, the gradient of loss can successfully back-propagate to optimize layers.

During training, rather to train the overall model in a simple end-to-end manner, we regard the intermediates after pair of ADM and Z submodule still provide reasonable information for image restoration. Thus, we encourage these intermediates to approximate the ground truth by setting the loss function as

$$loss = \sum_n w_n \|\hat{I}_n - I\|_2^2, w_n = \frac{\beta_n}{\sum_n \beta_n} \quad (11)$$

where \hat{I}_n denotes the restoration after n -th ADM and Z submodule pair, and β_n denotes the preset value on n -th ADM. We have noted β which is an exponentially ascending sequence that complied with an intuitive to put more emphasis on the latter outputs rather than the former, so that it is used as a normalized weighting term over corresponding loss. An advantage derived is a more penetrated structure that the gradient of loss can easily propagate for parameter optimization. During inference, the final output is restored from the last layer. In this scheme, the Z submodule is expected to learn from the mapping between intermediates \hat{I}_n and optimal ADM input \hat{Z}_n . Although with the reduction of flexibility described in existing plug-and-play manner, the learned image priors in Z submodules can accurately reflect to the interior mappings and lead to an efficient convergence and restoration quality.

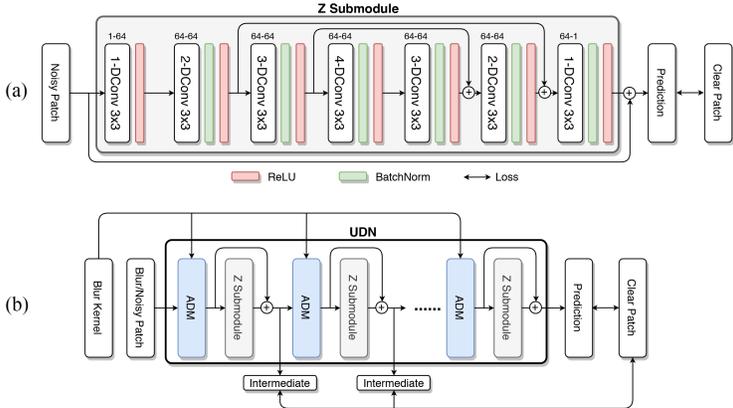


Fig. 3. (a) Training scheme and structure of Z submodule. p -DConv denotes p -dilated convolution. (b) Training scheme and structure of UDN (with ADM integration)

4 Experiments

To generate the training data $\{(I_i, B_i, K_i)\}_{i=1}^M$, we crop 400 grayscale images selected from BSD500 dataset [35] of 64×64 with striding and augment with horizontal and vertical flip to prevent over-repetitiveness, then simulate 200 linear motion kernels of 25×25 with their length uniformly sampled from 5 to 25 at any orientation. The blurry image is synthesized by convolving full images with kernels and introducing additive Gaussian noise term from $N(0, \sigma^2)$. After that, 8-bit quantization is applied and the pixel values are clipped in range $[0, 255]$.

The blurry images are cropped into patches corresponding to that of clean ones, and those near the boundaries are excluded to prevent restoring from missing information. Finally, the training dataset is composed of 69600 triplets in total. Two testing sets are generated in a similar way. The first is synthesized by convolving Set14 dataset [36] which composed of 14 standard testing images with 8 blur kernels from Levin *et al.* [37]. The second is generated from Sun *et al.* dataset [38] which consists of 80 high-resolution images and 8 estimated kernels from Pan *et al.* dataset [39].

To optimize the network parameters in UDN, an Adam optimizer [40] is adopted by setting initial step at 10^{-3} ; the learning rate is halved at 5-th, 15-th and 25-th epoch. The mini-batch size is set to 96, and trained by PyTorch1.0 [41] and Nvidia 2080Ti GPU for 50 epochs (about 15hours).

In addition, we have merged the outputs of ADM and UDN as a final restoration result. Based on the empirical observation that ADM provides enhanced details with smooth structures and UDN performs well on restoring high-frequency components, we combine two outputs in frequency domain by linear combination with a Gaussian mask (see detail description in the supplementary material).

For the following, we utilize a numerical image quality assessment (IQA) with other methods, including a patch-based learning method (EPLL) [8], non-local mean filter based method (DEB-BM3D) [42] and its modified version (IDD-

BM3D) [43], and MLP [12], FCN [14], FDN [11] and IRCNN [15]. Among these methods, EPLL, DED-BM3D and IDD-BM3D are categorized into traditional state-of-the-arts that do not rely on the neural networks. The remaining methods are neural network based, e.g. FCN, FDN and IRCNN are proposed at a similar period, all of these methods are CNN-based and derived from different concepts. The reason that we do not compare our works with those traditional prior-based methods like Krishnan *et al.* [5] or Pan *et al.* [39] is due to the difficulties of setting optimal hyperparameters, so that we cannot ensure these methods reach their best performance. For evaluation fairness, all of the methods are reproduced from their referenced code, as well as the pretrained models. The code corresponds to this paper is released¹.

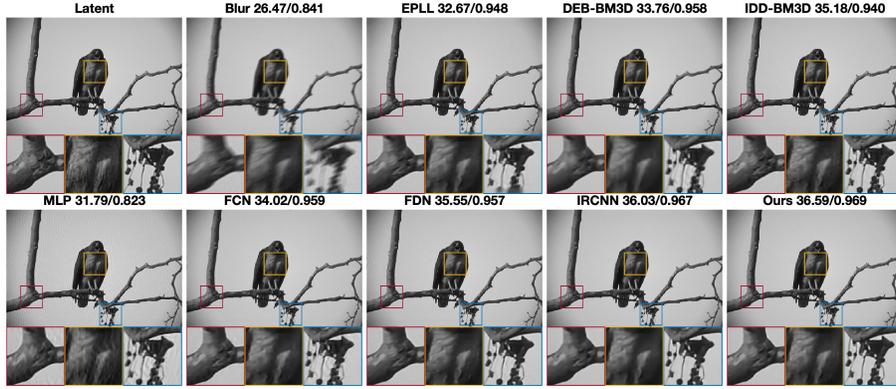
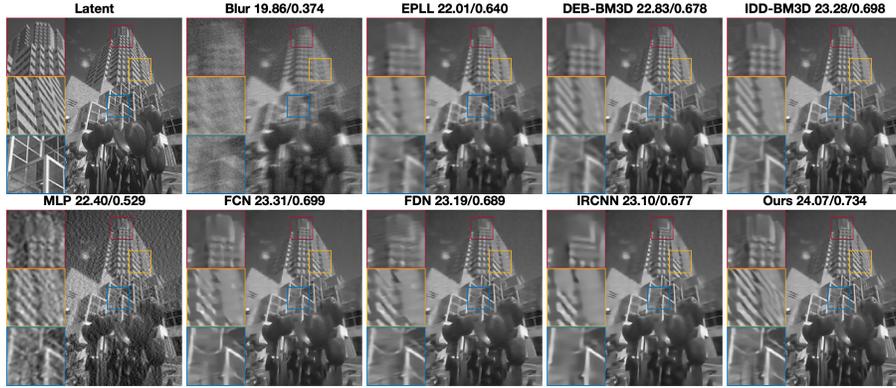
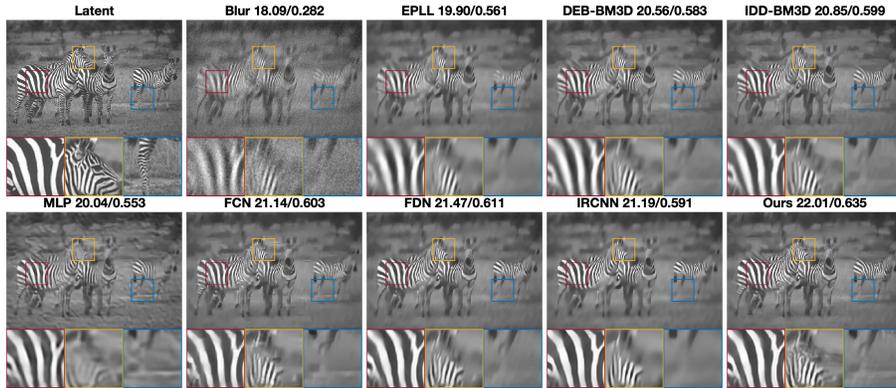
4.1 Non-Blind Deconvolution on Synthesized Datasets

We demonstrate several visualization examples on the BSD500 dataset [35] in exclusive of training patches that convolves with generated linear kernels with additive noise level at 1%, 3%, and 5%, to better discriminate the restoration performance and the robustness of algorithms. More visual and experimental results are included in the supplementary material.

For low noise level conditions, either when $\sigma = 1\%$ or $\sigma = 3\%$, IDD-BM3D shows apparently better results among traditional state-of-the-arts and even surpasses some of the deep learning based methods. In Fig. 4, some of the learning based methods like MLP, FCN and FDN suffer from obvious ringing artifacts near the edge regions. In comparison, IRCNN and proposed methods can successfully preserve the details and structures without aliasing. The latter restores even more high frequency components; for example, displayed in the yellow block, so that the overall PSNR is better than the others. In Fig. 5, IDD-BM3D demonstrates a powerful restoration quality that not only outperforms traditional methods but also shows a competitive performance to some of the learning based methods. We also find that due to the over-smoothness, IRCNN has lost a lot of high-frequency components like the details originally existed in the red or yellow blocks, such that the performance is even worse than a patch-based method IDD-BM3D. Our model demonstrates even better performance on retrieving fine textures lies in the red and yellow blocks that is hard to be aware of in the blurred image. Besides, the restored structures in blue blocks are the most precise among all methods. In Fig. 6, where the noise level is increased to 5%, the FDN owns a powerful restoration quality that surpass all other methods excluding our proposal. We find the aliasing and over-smoothness that usually appear in learning based models can deteriorate the performance significantly. On the other hand, the proposed model has successfully restored the textures in the yellow block which is inappreciable in the blurry observation.

We also numerically evaluate aforementioned 2 testing datasets and summarize quantitative results in Table. 2 and Table. 3. Complying to our visualization

¹ https://github.com/angry77cat/DeepPriorUDN_ADM

Fig. 4. Visual comparison on the images at $\sigma = 1\%$.Fig. 5. Visual comparison on the images at $\sigma = 3\%$.Fig. 6. Visual comparison on the images at $\sigma = 5\%$.

results, the proposed ADM + UDN model generally demonstrates the best performance over the others. Such a surpassing image quality is found at a wide range of perturbed noise level from 1% to 10%, showing the robustness of our proposed models. It is worth to note that the increment on performance are generally proportional to the noise level which implies the significance of restoration quality.

Table 2. Comparison on the Set14 dataset [36]. N/A denotes the absence of corresponding pretrained models in referenced code and bold numbers indicate the best performance.

σ	IQA	EPLL	DEB-BM3D	IDD-BM3D	MLP	FCN	FDN	IRCNN	Ours
1%	PSNR	28.40	30.89	31.76	30.56	29.48	30.02	31.57	31.87
	SSIM	0.845	0.854	0.879	0.848	0.860	0.862	0.876	0.884
3%	PSNR	25.43	26.95	27.79	25.53	26.68	27.16	27.63	28.32
	SSIM	0.725	0.744	0.765	0.637	0.758	0.759	0.760	0.787
5%	PSNR	23.96	25.29	26.00	24.33	25.30	25.88	25.97	26.83
	SSIM	0.663	0.689	0.705	0.657	0.706	0.710	0.697	0.733
10%	PSNR	22.12	23.39	23.80	23.53	N/A	N/A	23.99	24.48
	SSIM	0.581	0.614	0.625	0.614	N/A	N/A	0.624	0.649

Table 3. Comparison on the Sun *et al.* dataset [38]. N/A denotes the absence of corresponding pretrained models in referenced code and bold numbers indicate the best performance.

σ	IQA	EPLL	DEB-BM3D	IDD-BM3D	MLP	FCN	FDN	IRCNN	Ours
1%	PSNR	30.53	31.93	32.65	31.58	31.30	31.42	32.45	33.00
	SSIM	0.867	0.871	0.887	0.858	0.877	0.885	0.880	0.901
3%	PSNR	27.46	27.86	28.73	26.48	28.37	28.51	28.59	29.43
	SSIM	0.749	0.747	0.775	0.707	0.779	0.783	0.759	0.806
5%	PSNR	26.08	26.37	27.08	26.20	26.94	27.28	27.11	27.97
	SSIM	0.689	0.690	0.714	0.691	0.729	0.735	0.704	0.753
10%	PSNR	24.43	24.77	25.22	25.20	N/A	N/A	25.34	26.18
	SSIM	0.622	0.629	0.642	0.641	N/A	N/A	0.640	0.684

4.2 Noise Robustness

Since the ADM acting as a modified deconvolution plugin benefits convergence, the model can be categorized into a plug-and-play framework. On the other hand, the UDNs trained at specific noise levels retrieve more detail information on the sacrifice of training flexibility. As most of the end-to-end training schemes lead to inferior performance when handling different circumstances, the UDNs have shown robustness to a wide range of noise that significantly outperform other models that trained in an end-to-end manner as well. We evaluate such robustness on Set14 dataset [36] by adjusting the level of additive noise. For Fig.

7(b-d), it is worth noting that the ADMs acting as additional plugins have globally improved the restoration quality compared to IRCNN, which demonstrates comprehensive improvements by our proposed adaptive deconvolution.

As shown in a comparison in Fig. 7(a), the UDN trained at a high noise level shows significant improvements at low noise level at which other models originally trained. When comparing with other plug-and-play models², the UDNs do not always outperform IRCNN but generally achieve better performance when perturbed by significant noise like Fig. 7(d) shows. Such a surprising observation has implied a flexibility for a model trained at a specific degradation domain, which is not observed in existing deblurring methods like FDN [11] and FCN [14].

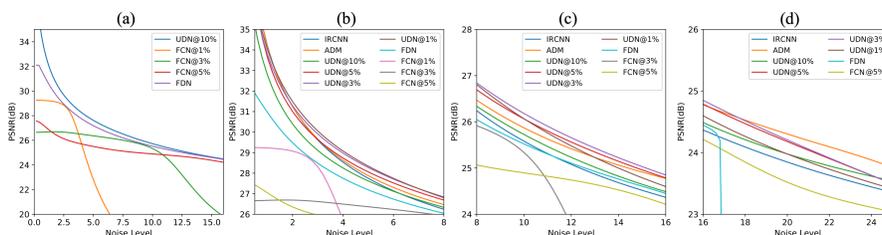


Fig. 7. Comparison on noise robustness for (a) models trained at specific noise levels and (b-d) includes IRCNN and ADM at different noise levels. @ denotes the model trained at which noise level.

4.3 Convergence Efficiency and Runtime Comparison

A good portion of deconvolution methods are based on iterative algorithms, for example, the iterative shrinkage-thresholding algorithm (ISTA) [44], the fast ISTA (FISTA) [45], and alternating direction method of multipliers (ADMM) [46] are useful algorithms for image restoration. Similarly, an important concern is whether the proposed models benefit the convergence efficiency compared to others. We take a close look into an image deconvolution case sampled from the BSD500 dataset [35] and the Levin *et al.* dataset [37] and the PSNR is evaluated at each iteration shown in Fig. 8. One can observe the ADM has benefited the restoration performance especially when the noise level is high. In other words, by deconvolving adaptively, the restoration process can converge to a better solution. Furthermore, the UDN not only reaches the best result but also achieves even more faster convergence speed.

We also compare the runtime. All of the compared methods are implemented by PyTorch1.0 [41] package along with IntelCore i9-9900KF CPU@3.60GHz, 64GB RAM, and Nvidia 2080Ti GPU. Since the originally released code of IRCNN [15] was implemented in Matlab, a PyTorch version is implemented. It is worth noting that the ADM is essentially an edge detector that only accompanied by several convolutional operations. Therefore, it can be parallelized by GPU acceleration. As shown in Table. 4, the runtime of the ADM is inevitably

² FDN is trained at various noise levels ranged from 0.1 to 12.75(5%).

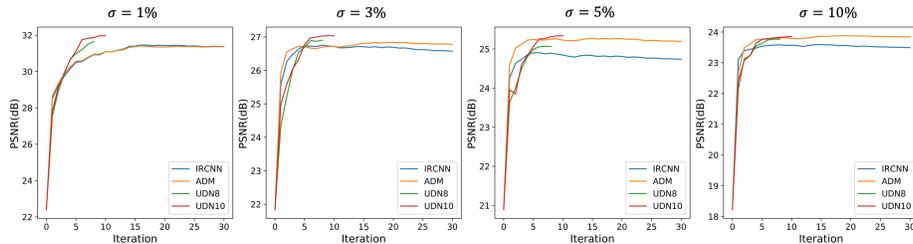


Fig. 8. Comparison on the convergence efficiency and the performance of optimal solution.

longer than that of the IRCNN; however, such overhead can be reduced in fold when handling high-resolution images. On the other hand, with the advantage of fast convergence, the runtime of the UDN8, which composed of 8 pairs of the ADM and Z submodules, is only 1/3 of that of the IRCNN.

Table 4. Runtime (in second) comparison.

Size	IRCNN	ADM	UDN8	UDN10
256×256	0.115	0.361	0.049	0.066
512×512	0.385	0.703	0.131	0.175
1024×1024	1.543	2.194	0.489	0.645

5 Conclusions

In this paper, we propose two learning based methods that are competitive to state-of-the-arts. The first approach is based on the success of incorporating MAP framework with CNN. As a solution to prevent over-smoothness which is usually found in CNN priors, a simple but useful plugin, ADM, is introduced without apparent increment on computation complexity. We have successfully alleviated over-smoothness and improved image quality for a wide range of noise perturbation. The second framework, UDN, is started from a concept to elaborate the great capacity of deep learning model to optimize the iterative optimization process in a gradient descent manner. Thus, a more effective optimization process is observed, which implies significantly less inference time. Furthermore, despite being trained at a specific noise level, UDN has demonstrated robustness for handling a wide range of cases and encouraged detail preservation. We further combine the outputs of two methods in frequency domain, leading to outstanding performance in evaluation benchmarks.

References

1. Wiener, N.: Extrapolation, interpolation, and smoothing of stationary time series. The MIT Press (1964)

2. Lucy, L.B.: An iterative technique for the rectification of observed distributions. *The Astronomical Journal* **79** (1974) 745
3. Richardson, W.H.: Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America* **62** (1972) 55–59
4. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena* **60** (1992) 259–268
5. Krishnan, D., Fergus, R.: Fast image deconvolution using hyper-laplacian priors. In: *Advances in Neural Information Processing Systems*. (2009) 1033–1041
6. Levin, A., Fergus, R., Durand, F., Freeman, W.T.: Image and depth from a conventional camera with a coded aperture. *ACM Transactions on Graphics* **26** (2007) 70–es
7. Aharon, M., Elad, M., Bruckstein, A.: K-svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on signal processing* **54** (2006) 4311–4322
8. Zoran, D., Weiss, Y.: From learning models of natural image patches to whole image restoration. In: *International Conference on Computer Vision*. (2011) 479–486
9. Roth, S., Black, M.J.: Fields of experts. *International Journal of Computer Vision* **82** (2009) 205
10. Gong, D., Zhang, Z., Shi, Q., Hengel, A.v.d., Shen, C., Zhang, Y.: Learning an optimizer for image deconvolution. *arXiv preprint arXiv:1804.03368* (2018)
11. Kruse, J., Rother, C., Schmidt, U.: Learning to push the limits of efficient fft-based image deconvolution. In: *IEEE International Conference on Computer Vision*. (2017) 4586–4594
12. Schuler, C.J., Christopher Burger, H., Harmeling, S., Scholkopf, B.: A machine learning approach for non-blind image deconvolution. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2013) 1067–1074
13. Xu, L., Ren, J.S., Liu, C., Jia, J.: Deep convolutional neural network for image deconvolution. In: *Advances in Neural Information Processing Systems*. (2014) 1790–1798
14. Zhang, J., Pan, J., Lai, W.S., Lau, R.W., Yang, M.H.: Learning fully convolutional networks for iterative non-blind deconvolution. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2017) 3817–3825
15. Zhang, K., Zuo, W., Gu, S., Zhang, L.: Learning deep cnn denoiser prior for image restoration. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2017) 3929–3938
16. Nan, Y., Ji, H.: Deep learning for handling kernel/model uncertainty in image deconvolution. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2020) 2388–2397
17. Ren, D., Zuo, W., Zhang, D., Zhang, L., Yang, M.H.: Simultaneous fidelity and regularization learning for image restoration. *IEEE transactions on pattern analysis and machine intelligence* (2019)
18. Rick Chang, J., Li, C.L., Poczos, B., Vijaya Kumar, B., Sankaranarayanan, A.C.: One network to solve them all—solving linear inverse problems using deep projection models. In: *IEEE International Conference on Computer Vision*. (2017) 5888–5897
19. Ren, D., Zhang, K., Wang, Q., Hu, Q., Zuo, W.: Neural blind deconvolution using deep priors. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2020) 3341–3350
20. Geman, D., Yang, C.: Nonlinear image recovery with half-quadratic regularization. *IEEE transactions on Image Processing* **4** (1995) 932–946

21. Levin, A., Weiss, Y.: User assisted separation of reflections from a single image using a sparsity prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29** (2007) 1647–1654
22. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. *ACM Transactions on Graphics* (2006) 787–794
23. Cho, T.S., Joshi, N., Zitnick, C.L., Kang, S.B., Szeliski, R., Freeman, W.T.: A content-aware image prior. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2010) 169–176
24. Goldstein, T., O’Donoghue, B., Setzer, S., Baraniuk, R.: Fast alternating direction optimization methods. *SIAM Journal on Imaging Sciences* **7** (2014) 1588–1623
25. Roth, S., Black, M.J.: Fields of experts: A framework for learning image priors. In: *IEEE Conference on Computer Vision and Pattern Recognition*. Volume 2. (2005) 860–867
26. Schmidt, U., Roth, S.: Shrinkage fields for effective image restoration. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2014) 2774–2781
27. Schmidt, U., Rother, C., Nowozin, S., Jancsary, J., Roth, S.: Discriminative non-blind deblurring. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2013) 604–611
28. Chen, Y., Pock, T.: Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE transactions on pattern analysis and machine intelligence* **39** (2016) 1256–1272
29. Sun, L., Cho, S., Wang, J., Hays, J.: Good image priors for non-blind deconvolution. In: *European Conference on Computer Vision*. (2014) 231–246
30. Venkatakrishnan, S.V., Bouman, C.A., Wohlberg, B.: Plug-and-play priors for model based reconstruction. In: *IEEE Global Conference on Signal and Information Processing*. (2013) 945–948
31. Chan, S.H., Wang, X., Elgendy, O.A.: Plug-and-play admm for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging* **3** (2016) 84–98
32. Dong, J., Pan, J., Sun, D., Su, Z., Yang, M.H.: Learning data terms for non-blind deblurring. In: *European Conference on Computer Vision*. (2018) 748–763
33. Wang, Y., Yang, J., Yin, W., Zhang, Y.: A new alternating minimization algorithm for total variation image reconstruction. *SIAM Journal on Imaging Sciences* **1** (2008) 248–272
34. Dong, W., Wang, P., Yin, W., Shi, G., Wu, F., Lu, X.: Denoising prior driven deep neural network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **41** (2018) 2305–2318
35. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33** (2010) 898–916
36. Zeyde, R., Elad, M., Protter, M.: On single image scale-up using sparse-representations. In: *International Conference on Curves and Surfaces*. (2010) 711–730
37. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding and evaluating blind deconvolution algorithms. In: *IEEE Conference on Computer Vision and Pattern Recognition*. (2009) 1964–1971
38. Sun, L., Cho, S., Wang, J., Hays, J.: Edge-based blur kernel estimation using patch priors. In: *IEEE International Conference on Computational Photography*. (2013) 1–8

39. Pan, J., Lin, Z., Su, Z., Yang, M.H.: Robust kernel estimation with outliers handling for image deblurring. In: IEEE Conference on Computer Vision and Pattern Recognition. (2016) 2800–2808
40. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
41. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al.: Pytorch: An imperative style, high-performance deep learning library. In: Advances in Neural Information Processing Systems. (2019) 8024–8035
42. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image restoration by sparse 3d transform-domain collaborative filtering. In: SPIE Image Processing: Algorithms and Systems VI. Volume 6812. (2008) 681207
43. Danielyan, A., Katkovnik, V., Egiazarian, K.: Bm3d frames and variational image deblurring. IEEE Transactions on Image Processing **21** (2011) 1715–1728
44. Bioucas-Dias, J.M., Figueiredo, M.A.: A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration. IEEE Transactions on Image processing **16** (2007) 2992–3004
45. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences **2** (2009) 183–202
46. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine learning **3** (2011) 1–122