



Unsupervised Knowledge Transfer for Nonblind Image Deconvolution

Zhuojie Chen^a, Xin Yao^a, Yong Xu^{a,b,c}, Junle Wang^d, Yuhui Quan^{a,c},

^aSchool of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China

^bPeng Cheng Laboratory, Shenzhen 510852, China

^cPahzhou Laboratory, Guangzhou 510335, China

^dTuring Lab, Tencent, Shenzhen 518054, China

Article history:

Keywords:

Nonblind Image Deconvolution
Deep Knowledge Transfer
Model Adaption
Image Recovery

ABSTRACT

Nonblind image deconvolution restores the clear image from a blurred one under a known blur kernel, whose recent development has been boosted by supervised deep learning. Motivated by the inaccessibility of ground-truth images for supervised learning in many application domains, such as scientific imaging, this paper studies the unsupervised knowledge transfer problem for nonblind image deconvolution, which aims at adapting a deep model pre-trained on a source domain, to a ground-truth-scare target domain where image contents or blur kernels are distinct from that of the source domain. We propose to conduct the knowledge transfer regarding both images and kernels, by leveraging the model being adapted itself to generate pairs of a pseudo ground-truth image and a blurred image for self training. The proposed method neither accesses source-domain data, which avoids privacy issues, nor accesses target-domain ground-truths, which avoids ground-truth collection. Its effectiveness is demonstrated with the experiments on three deblurring tasks in different domains.

© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

Given a blurred image Y generated by a blur kernel K , nonblind image deconvolution (NID) recovers the latent image X from the following ill-posed linear system:

$$Y = K \otimes X + N, \quad (1)$$

where N is the image noise often assumed to be Gaussian white, and \otimes denotes image convolution. Such a convolutional model covers many types of image blurring. Thus, NID is an important technique in image processing and serves as a critical module for blind deblurring (Liu et al., 2018), with a broad spectrum of applications ranging from hand-shake blur removal and microscopic deconvolution to satellite imagery and remote sensing.

The ill-posedness of (1) originates from both K and N , *i.e.*, convolving with a blur kernel will weaken or erase the high-frequency image parts in Fourier domain, and a naive inversion on (1) will amplify the noise inevitably. Optimization-based methods (Mignotte, 2008; Danielyan et al., 2011; Liu et al., 2013; Quan et al., 2015; Mosleh et al., 2017; Zha et al., 2020; Dong and Pan, 2021) address the ill-posedness by imposing certain statistical image priors on X . Their performance depends on how well the priors fit latent images, whereas it is challenging to design universal image priors for various domains.

In recent years, neural networks (NNs) with deep learning on paired blurred and clear images have emerged as a successful approach for NID. Existing studies mainly focus on the design of effective NN architectures for improving recovery accuracy and generalization; see *e.g.* (Zhang et al., 2017a; Kruse et al., 2017; Jin et al., 2017; Dong et al., 2019; Nan et al., 2020; Eboli et al., 2020; Pronina et al., 2020; Quan et al., 2021; Dong et al., 2021; Mou et al., 2022; Zhai et al., 2022).

No matter what kind of NN architectures is used, the success of those methods relies on the accessibility of large-scale datasets with ground-truth (GT) images. Nevertheless, acquir-

Corresponding author at: School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China.

Zhuojie Chen and Xin Yao contribute equally.

e-mail: cs_zhuojiechen@mail.scut.edu.cn (Zhuojie Chen), 201930344312@mail.scut.edu.cn (Xin Yao), yxu@scut.edu.cn (Yong Xu), jljunlewang@tencent.com (Junle Wang), csyhquan@scut.edu.cn (Yuhui Quan)

ing GT images can be extremely difficult or even impractical in many application domains such as scientific imaging. On those domains (referred to as target domains), calling a model pre-trained from another domain (referred to as source domain) is likely to cause a dramatic performance drop, especially when the images and blur kernels differ a lot between two domains. It is highly desirable to have a method that can transfer a pre-trained model on the source domain so that it performs well on the target domain. This inspires us to study the unsupervised model adaption problem for the knowledge transfer of NID, where the target domain provides no GT images during knowledge transfer. Considering that source-domain training data is routinely unavailable due to privacy issues or transmission limitations, we further assume that the transfer process is blind to (*i.e.*, not accessing) any source-domain data.

In this paper, an unsupervised model adaption method for knowledge transfer is proposed, which enables a model pre-trained on a source domain to be effectively transferred to a target domain. This is done via a self-training task that utilizes the model being adapted as the generator of “quasi” clear images in the target domain, and the generated clear images are re-blurred to form blurred/clear image pairs for self-learning. The idea behind is that, a source-domain pre-trained model is likely to perform reasonably on the target domain, due to its function as a NID solver, even when the two domains differ much from each other. As a result, the target-domain images deblurred by the source-domain model may contain reasonable high-frequency components which can be used as the supervision for the NID in the target domain for improving the performance of the model, and the improved model may further provide better deblurred images of target domain for better supervision.

In other words, we treat target-domain predictions from the intermediately-trained NN as pseudo GT images and use them with transform-based image augmentation and kernel augmentation for continuous training and improvement. The self-training only calls the pre-trained model without the need for accessing the source-domain data during adaptation. Such a feature differs from many knowledge transfer or domain adaptation methods in computer vision (Yuan et al., 2020; Deng et al., 2021; Marnissi et al., 2022) and makes it applicable even when source-domain training data is inaccessible.

The effectiveness of the proposed approach is demonstrated with extensive experiments, where it successfully adapted five representative source models pre-trained on natural image motion deblurring to deblurring Gaussian-blurred natural images, microscopic imaging, and remote sensing imagery respectively, with noticeable improvement achieved in terms of both quantitative and qualitative results. The adapted models perform competitively against those directly trained with a small number of target-domain GT images. To our knowledge, this is one of the very few works that studies the unsupervised knowledge transfer problem for NID, with an effective solution provided.

2. Related Work

2.1. Deep Learning-based NID

Most existing deep learning-based methods for NID are supervised, which leverage a deep NN as a blurry-to-clear map-

ping learned on paired blurred/clear images. These methods have a particular focus on improving NN architectures. Ren et al. (2018) connected an inversion NN with an artifact removal NN to model the mapping from blurred images to the latent ones. Vasu et al. (2018) proposed a deep NN that ensembles multiple deblurred images obtained by a regularized inverse with different regularization weights. Dong et al. (2020) proposed to run Wiener NID on the features generated by a deep NN, with multi-scale progressive refinement. Deep unfolding is a prominent approach to developing effective NNs for NID, which translates an iterative NID solver to the alternation between an inversion process based on the images estimated by previous iterations and a denoising process implemented by a convolutional NN; see *e.g.* (Kruse et al., 2017; Zhang et al., 2017a; Meinhardt et al., 2017; Zhang et al., 2017b; Dong et al., 2019; Jin et al., 2017; Bigdeli et al., 2017; Gong et al., 2020; Nan et al., 2020; Nan and Ji, 2020; Eboli et al., 2020; Chen et al., 2021; Dong et al., 2021; Mou et al., 2022; Zhai et al., 2022). Recently, there is an increasing trend to translate a regularized inverse of NID into a deep NN via replacing the regularizer in the inverse by NN blocks; see *e.g.* (Gilton et al., 2020; Pronina et al., 2020).

2.2. Unsupervised Knowledge Transfer for Image Recovery

There are only a few works on unsupervised knowledge transfer or model/domain adaption for image recovery. Soh et al. (2020) used low-resolution images as GTs and their downsampled versions as input for the self-training of super-resolution. This method is based on scale recurrence of the down-sampling process related to low-resolution images, which is inapplicable to NID. Wang et al. (2021) and Shao et al. (2020) used CycleGAN for super-resolution and dehazing respectively. Such two methods are not for NID and they require the access to the source-domain training data, which is inapplicable to the blind setting adopted in our work. Barbano et al. (2021) proposed a Bayesian NN for unsupervised knowledge transfer of image reconstruction of medical images, but not applicable to NID either. Gilton et al. (2021) studied the model adaption for image recovery which focuses on the adaption to the test data within the same domain, rather than the adaption to another domain. It is not effective when applied to NID, as observed in our experiments.

3. Methodology

We first formulate the problem of unsupervised model adaption. Following the notations in (1), we use

$$\mathbb{D}_{\text{train}}^s := \{(Y_i^s, \mathbf{K}_i^s, X_i^s) | Y_i^s = \mathbf{K}_i^s \otimes X_i^s + N_i^s\}_i \quad (2)$$

to denote the training data in source domain with GT images available. The source model pre-trained on $\mathbb{D}_{\text{train}}^s$ is denoted by

$$\mathcal{M}_\theta : (Y, \mathbf{K}) \rightarrow X, \quad (3)$$

where θ encodes the pre-trained NN’s parameters. Let

$$\mathbb{D}^t := \mathbb{D}_{\text{train}}^t \cup \mathbb{D}_{\text{test}}^t \{(Y_j^t, \mathbf{K}_j^t) | Y_j^t = \mathbf{K}_j^t \otimes X_j^t + N_j^t\}_j \quad (4)$$

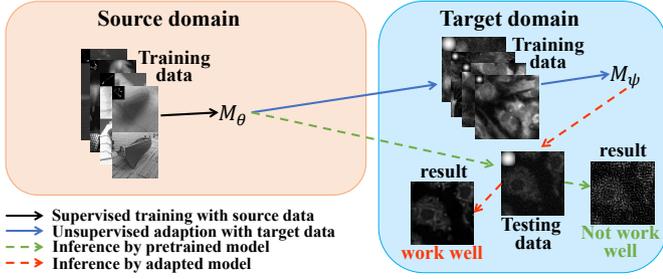


Figure 1. Illustration of unsupervised knowledge transfer by model adaptation for NID.

denote the data in target domain without GT images. The $\mathbb{D}_{\text{train}}^t, \mathbb{D}_{\text{test}}^t$ denote the training and test datasets, respectively.

Recall that convolving with a blur kernel is about attenuating the Fourier coefficients in high frequency bands. The attenuation patterns are statistically different among different types of blur kernels. Accordingly, the noise and artifacts generated in the inversion steps during deconvolution are also statistically different among different blur kernel types. In general, a model trained on specific noise/artifact types cannot generalize well on another type. Therefore, when the blur kernels differ much between the source and target domains, the source model will fail to perform well in the target domain. In addition, when there are statistical differences in image patterns between two domains, the image priors encoded in the source model do not fit well the images in the target domain, and the source model will not work well either.

The unsupervised model adaptation (knowledge transfer) task to study in this paper is illustrated in Figure 1, which aims at transferring \mathcal{M}_θ to a new model \mathcal{M}_ψ via parameter adjustment based on $\mathbb{D}_{\text{train}}^t$:

$$\text{Knowledge Transfer} : (\theta, \mathcal{M}_\theta) \rightarrow (\psi, \mathcal{M}_\psi), \quad (5)$$

so that \mathcal{M}_ψ works well on $\mathbb{D}_{\text{test}}^t$. Due to the non-existence of GT images in \mathbb{D}^t , such an adaptation task is very challenging particularly when image contents or blur kernels are structurally distinct between the source domain and target domain.

Considering source-domain data may be unavailable due to privacy issues or transmission limitations in many scenarios, we further assume the proposed approach is blind to source-domain data $\mathbb{D}_{\text{train}}^s$ and only the source model \mathcal{M}_θ is available for adaptation. In the proposed approach, the unsupervised model adaptation and knowledge transfer of NID is done by fine-tuning \mathcal{M}_θ via the self-training loss below:

$$\mathcal{L}(\psi | \mathbb{D}^t) := \mathbb{E}_{(\tilde{Y}, \tilde{K}) \sim \mathbb{D}^t, \tilde{N} \sim \mathcal{N}, \mathcal{T}_1, \mathcal{T}_2 \sim \mathbb{T}} \left\| \mathcal{T}_1(\mathcal{M}_\psi(\tilde{Y}, \tilde{K})) - \mathcal{M}_\psi(\mathcal{T}_2(\tilde{K}) \otimes \mathcal{T}_1(\mathcal{M}_\psi(\tilde{Y}, \tilde{K})) + \tilde{N}, \mathcal{T}_2(\tilde{K})) \right\|_F^2, \quad (6)$$

where \tilde{N} is the simulated noise set to Gaussian white with noise level randomly drawn from the range of noise levels in the target-domain data, and \mathbb{T} is a set of transforms for augmentation which includes rotations of $0^\circ, 90^\circ, 180^\circ, 270^\circ$ respectively as well as horizontal and vertical reflections. See

While the loss \mathcal{L} looks a bit complicated at the first glance, its basic idea is not complicated. Suppose we omit the transforms

$\mathcal{T}_1, \mathcal{T}_2$ by simply setting them to identity mappings and define $X' = \mathcal{M}_\psi(\tilde{Y}, \tilde{K})$. Then, training with \mathcal{L} is to encourage that

$$\mathcal{M}_\psi(\tilde{K} \otimes X' + \tilde{N}) \approx X'. \quad (7)$$

It can be seen that the model \mathcal{M} is applied to the target-domain samples (\tilde{Y}, \tilde{K}) . Then we use the output X' to simulate a target-domain blurred image using the kernel \tilde{K} in the target domain, and use X' as the supervision for training \mathcal{M} . Since \mathcal{M} has been well trained for deblurring in the source domain, it is likely to output a relatively clear image in the target domain at the first few iterations, and it can become better and better along the iterations, leading to a good adaptation of the model. The transforms $\mathcal{T}_1, \mathcal{T}_2$ act as data augmentation for the self-supervised training for improvement.

The proposed loss can also be viewed as a cycle data consistency loss with augmentations. Consider the plain data consistency loss $\|\tilde{Y} - \tilde{K} \otimes \mathcal{M}_\psi(\tilde{Y}, \tilde{K})\|_F^2$, where the supervision is given by a blurred image \tilde{Y} with attenuated high-frequency components. As a result, \mathcal{M}_ψ is not easy to learn how to restore the high-frequency components of a sharp image. In comparison, the proposed loss utilizes the augmented output image $\mathcal{T}_1(X')$ of the pre-trained model as a latent sharp image to form paired data for training. Since X' is a recovered image which may contain additional high-frequency components over Y , $\mathcal{T}_1(X')$ may provide more effective supervision for the adaption *w.r.t.* kernels $\mathcal{T}_2(\tilde{K})$.

We observed that even the quality of the deblurring results of the target-domain images may be not high, the model being adapted by the proposed method could be continuously improved along the adaption process. The reason is probably that, a source-domain pre-trained model functions as a solver of NID, which has reasonable performance on the target domain, even if two domains contain much diversity. Thus, the estimated high-frequency components of the target-domain images deblurred by the source-domain model contains much useful information, which can be used to supervise the NID in the target domain for improving the models' performance. Then the improved model may further provide better deblurred images of target domain for better supervision.

4. Experiments

The evaluation is done by transferring a model trained on natural image motion deblurring to three scenarios: natural image deblurring with Gaussian kernels, microscopy deconvolution with kernels from the Richards-Wolf (RW) optical model (Pronina et al., 2020), and remote sensing deblurring with kernels from the turbulence degradation model (Gao et al., 2018). Regarding the source models, we select some representative NNs from recent studies, including

- FDN (Kruse et al., 2017): A deep NN for NID that unrolls a CNN-regularized FFT-based deconvolution process with a simple yet effective boundary adjustment scheme to alleviate the problematic circular convolution assumption.
- VEM (Nan et al., 2020): A deep NN constructed by unrolling a variational expectation maximization process of a probabilistic model for NID.

Table 1. Mean PSNR(dB)/SSIM results of source models and adapted models in three NID tasks. The “Diff-Ours” and “Diff-P&P” denote the performance gain from the proposed adaptation method and the P&P method respectively.

Model	Gaussian deblurring			Microscopy deconvolution					Remote sensing deblurring		
	$\sigma = 1\%$	$\sigma = 3\%$	$\sigma = 5\%$	$\sigma = 0.1\%$	$\sigma = 0.5\%$	$\sigma = 1\%$	$\sigma = 5\%$	$\sigma = 10\%$	$\sigma = 1\%$	$\sigma = 3\%$	$\sigma = 5\%$
FDN(S)	34.54/0.93	30.58/0.86	28.47/0.81	39.87/0.94	38.35/0.92	37.17/0.90	33.42/0.83	30.41/0.74	28.13/0.76	26.64/0.71	25.93/0.69
FDN(A)	35.51/0.94	31.92/0.89	30.11/0.86	40.22/0.94	38.97/0.92	37.97/0.91	34.55/0.86	32.65/0.82	28.13/0.77	26.75/0.71	26.08/0.69
Diff-Ours	0.97/0.01	1.34/0.03	1.64/0.05	0.35/0.00	0.62/0.00	0.80/0.01	1.13/0.03	2.24/0.08	0.00/0.01	0.11/0.00	0.15/0.00
Diff-P&P	0.29/0.00	0.33/0.01	0.31/0.02	0.05/0.00	0.10/0.00	0.22/0.00	0.21/0.01	0.22/0.02	-0.02/0.00	0.02/0.00	0.02/0.00
VEM(S)	35.06/0.93	30.42/0.85	28.11/0.78	37.52/0.90	36.51/0.87	36.06/0.88	32.19/0.80	27.01/0.63	27.87/0.74	25.63/0.63	23.83/0.54
VEM(A)	35.99/0.94	32.18/0.89	30.19/0.85	39.81/0.93	39.08/0.93	38.31/0.92	35.08/0.87	33.03/0.84	28.14/0.76	26.75/0.70	26.00/0.66
Diff-Ours	0.93/0.01	1.76/0.04	2.08/0.07	2.29/0.03	2.57/0.06	2.25/0.04	2.89/0.07	6.02/0.21	0.27/0.02	1.12/0.07	2.17/0.12
Diff-P&P	0.27/0.00	0.43/0.01	0.42/0.02	0.38/0.01	0.46/0.02	0.61/0.01	0.53/0.02	0.58/0.03	0.05/0.01	0.27/0.02	0.64/0.04
USRNet(S)	35.55/0.94	31.77/0.88	29.12/0.81	38.55/0.92	38.44/0.92	37.98/0.92	34.52/0.86	31.72/0.80	27.77/0.75	26.00/0.67	24.79/0.62
USRNet(A)	35.85/0.94	32.24/0.89	30.31/0.85	39.10/0.93	38.70/0.93	38.07/0.92	35.08/0.88	33.24/0.84	28.15/0.77	27.00/0.72	26.23/0.69
Diff-Ours	0.30/0.00	0.47/0.01	1.19/0.04	0.55/0.01	0.26/0.01	0.09/0.00	0.56/0.02	1.52/0.04	0.38/0.02	1.00/0.05	1.44/0.07
Diff-P&P	0.08/0.00	0.09/0.00	0.24/0.01	0.09/0.00	0.03/0.00	0.01/0.00	0.10/0.01	0.13/0.01	0.07/0.01	0.23/0.02	0.42/0.02
CPCR(S)	34.90/0.93	30.01/0.81	25.16/0.59	38.78/0.92	37.83/0.91	37.00/0.90	25.52/0.43	20.75/0.23	27.32/0.75	25.76/0.66	22.18/0.45
CPCR(A)	35.25/0.94	31.72/0.89	30.10/0.83	40.15/0.93	39.09/0.93	38.03/0.92	34.80/0.85	32.90/0.80	27.69/0.76	26.59/0.70	26.00/0.67
Diff-Ours	0.35/0.01	1.71/0.08	4.94/0.24	1.37/0.01	1.26/0.02	1.03/0.02	9.28/0.42	12.15/0.57	0.37/0.01	0.83/0.04	3.82/0.22
Diff-P&P	0.08/0.00	0.40/0.02	0.97/0.07	0.26/0.00	0.24/0.01	0.29/0.01	1.85/0.11	1.19/0.11	0.06/0.00	0.19/0.01	1.14/0.07
DWDN(S)	35.89/0.94	32.53/0.90	30.73/0.87	38.94/0.91	38.29/0.91	37.54/0.90	34.67/0.86	32.64/0.84	28.04/0.76	26.78/0.72	26.06/0.69
DWDN(A)	35.90/0.95	32.56/0.91	30.79/0.88	40.72/0.94	39.47/0.93	38.44/0.92	35.16/0.87	33.41/0.85	28.42/0.77	26.95/0.72	26.22/0.69
Diff-Ours	0.01/0.01	0.03/0.01	0.06/0.01	1.78/0.03	1.18/0.02	0.90/0.02	0.49/0.01	0.77/0.01	0.38/0.01	0.17/0.00	0.16/0.00
Diff-P&P	-0.02/0.00	-0.01/0.00	0.00/0.00	0.26/0.01	0.18/0.01	0.26/0.01	0.08/0.00	0.05/0.00	0.08/0.00	0.03/0.00	0.04/0.00

- USRNet (Zhang et al., 2020): an unrolling-based deep NN for image super-resolution which is also applicable to NID by setting downsampling ratio to 1.
- CPCR (Eboli et al., 2020): A deep that unrolls the linear regularized least-squares model of NID using half-quadratic splitting, with Richardson fixed-point iterations pre-conditioned by approximate inverse filters being used for the least-squares updates.
- DWDN (Dong et al., 2020): A deep NN that incorporates Winier-filter-based deconvolution into the intermediate feature processing.

These methods which have diversity among their NN structures. For all methods except DWDN, we use their released models on natural image motion deblurring pre-trained on BSDS500, using motion kernels of (Arbelaez et al., 2010) and white Gaussian noise (WGN) with standard deviation in $[\frac{1}{255}, \frac{14}{255}]$. For DWDN, we use its code for pre-training on the same setting.

Throughout the experiments, model adaptation is done with the same manner: The Adam optimizer is used with the learning rate of $1e^{-4}$ and 600 iterations. We use “Model(S)” to denote the source model directly applied to the target domain, while using “Model(A)” to denote the unsupervisedly-adapted model. Data augmentation using \mathbb{T} is also applied to each image and each kernel in $\mathbb{D}_{\text{train}}^t$, in addition to the training loss \mathcal{L} . The results are summarized in Table 1. For comparison, we select the P&P method (Gilton et al., 2021) as the baseline, which is designed for the adaption to same-domain data rather than cross-domain data. Quantitative results are summarized in Table 1, with performance gain (denoted by “Diff”) listed for facilitating the comparison.

4.1. Performance Evaluation

4.1.1. Adaptation to Gaussian deblurring of Natural Images

We follow (Nan et al., 2020) for preparing the target-domain data. Random cropping is applied to the BSDS500 dataset (Arbelaez et al., 2010) to generate 1500 sharp images of size 256×256 as GTs. Then we generate 50 Gaussian kernels with the size and standard deviation uniformly sampled from $[5, 25]$ and $[0.5, 3.5]$ respectively. The blurred images are generated by first convolving each sharp image with one randomly-picked blur kernel and then corrupting it using WGN with standard deviation randomly chosen from $[\frac{1}{255}, \frac{14}{255}]$. In other words, there is only one blurred version for each shape image, which can better simulate realistic scenarios.

The Levin *et al.*’s dataset (Levin et al., 2011) is used for test. The blurred images are generated by first convolving each sharp image with the five Gaussian-like kernels from (Danielyan et al., 2011) respectively and then adding WGN with standard deviation $\sigma = 1\%, 3\%, 5\%$ respectively. The results listed in Table 1 show that the adaptation does bring improvement. However, the performance gain varies across different models and different noise levels. It is very little for DWDN but noticeable for VEM. The improvement of CPCR is small for $\sigma = 1\%$ but significant for $\sigma = 5\%$. See Figure 2 for a visual inspection. Note that without a specific scheme for cross-knowledge transfer, the models adapted by “P&P” show inferior performance to the ones adapted by the proposed method.

4.1.2. Adaptation to Microscopic Deconvolution

Following Pronina et al. (2020), the cell segmentation dataset (Al-Kofahi et al., 2018) and fluorescence microscopy dataset (Zhang et al., 2019) are used to generate 975 noisy

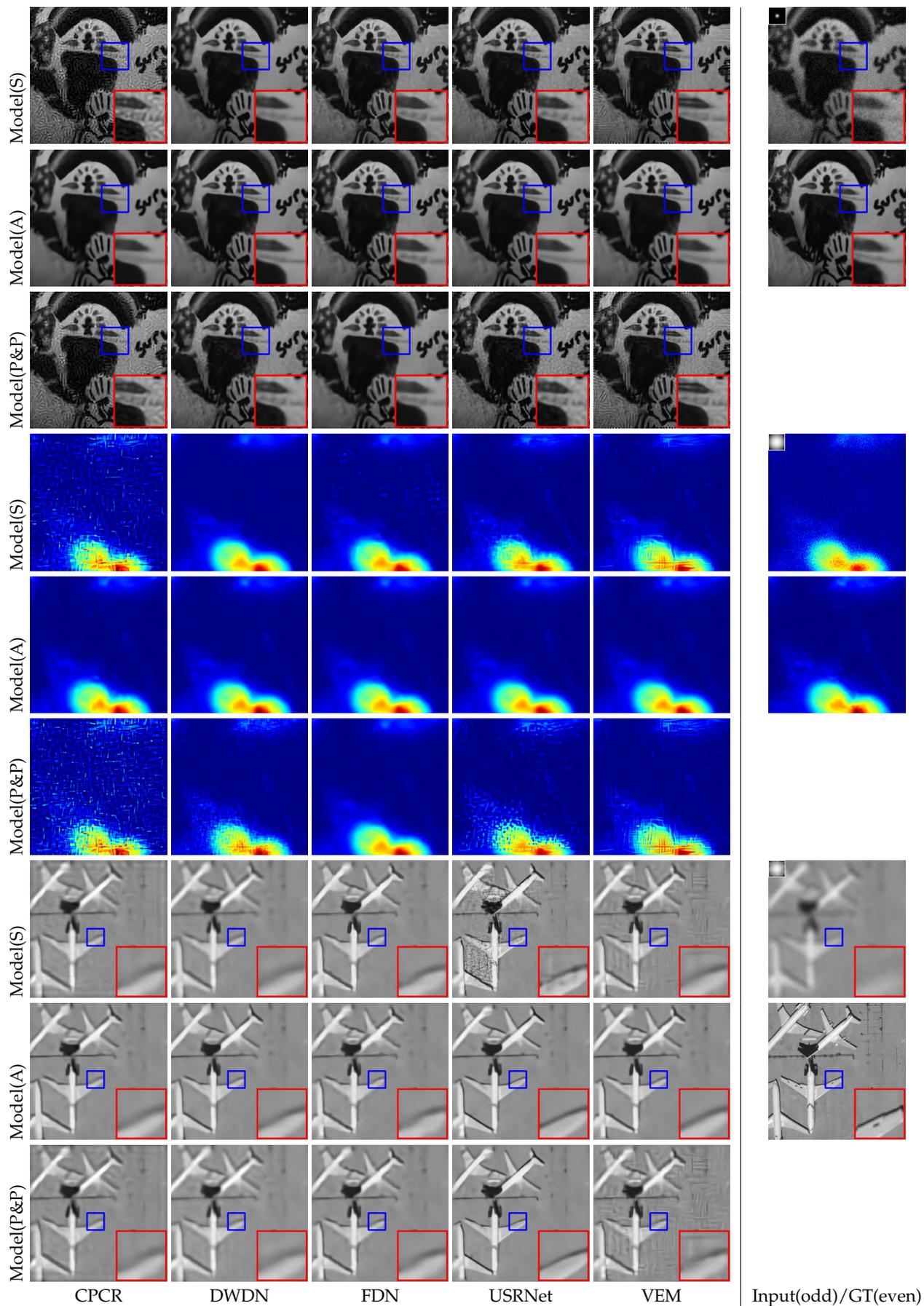


Figure 2. Visual results of selected source/adapted models on Gaussian deblurring, microscopy deconvolution, and remote sensing imagery.

blurred images for adaptation, using 25 optical kernels and WGN with standard deviation randomly selected from $\{0.1\%, 0.5\%, 1\%, 5\%, 10\%\}$. To be realistic, we ensure each blurred image in training corresponds to a different latent image. For evaluation, we generate 1150 blurred images by applying 5 blur kernels (different from the ones in adaptation) to 230 latent images, and then add WGN with standard deviation $\sigma = 0.1\%, 0.5\%, 1\%, 5\%, 10\%$ respectively.

The results in Table 1 show the performance improvement brought by the adaptation. Similar to the previous experiment, the PSNR gain varies across different settings and different models. The least one is 0.09dB on USRNet with $\sigma = 1\%$, while the biggest one is 12.15dB on CPCr with $\sigma = 10\%$. Interestingly, the DWDN which showed little improvement in the previous experiment has noticeable PSNR gain this time. We also note that the overall improvement in this experiment is larger than that in the previous one. This is because, in microscopic deconvolution, both images and kernels differ between the source and target domains, while in Gaussian natural image deblurring only the kernels differ from the source-domain ones. It is worth mentioning that, the source models are pre-trained on the noise levels from $[\frac{1}{255}, \frac{14}{255}]$. When dealing with stronger noise ($\sigma = 10\%$), those source models perform much worse, as shown in the quantitative results in Table 1 and the qualitative results in Figure 2. After adaptation, the models can handle a wider range of noise levels, *i.e.* $[0.1\%, 10\%]$. Again, the models adapted via ‘‘P&P’’ are worse than those via the proposed method, specially in the high noise level *i.e.* 10%. All above results clearly demonstrated the benefits of the proposed method.

4.1.3. Adaptation to Remote Sensing Imagery

Following Gao et al. (2018), 700 remote sensing images of airplanes, baseball diamonds, beaches, buildings, dense residential, storage tanks and tennis courts selected from the Land-Use dataset (Yang and Newsam, 2010) are used to construct the target-domain data for training (570 images) and test (130 images). The turbulence degradation model (Gao et al., 2018) is used to randomly generating 30 blur kernels, with 25 for training and 5 for testing. The blurred images for training are generated by convolving the sharp ones by one of the 25 kernels and then corrupted by WGN with standard deviation sampled from $[\frac{1}{255}, \frac{14}{255}]$. The blurred images for test are generated by convolving each sharp image with 5 testing kernels respectively and then adding WGN with standard deviation $\sigma = 1\%, 3\%, 5\%$.

From the results presented in Table 1 we can come to similar conclusions with the ones in previous experiments: performance gain can be achieved by the adaptation but varies a lot. The overall performance gain is smaller than that in microscopy deconvolution, which is probably due to the relatively-small domain gap in this experiment. Even that, the models adapted by the proposed method show noticeably better performance than that adapted by P&P. See also Figure 2 for a visual comparison, where the adaptation improves the visual quality noticeably.

4.2. Behavior Analysis

4.2.1. Performance versus Iteration Number

Table 1 has demonstrated the effectiveness of the proposed method in unsupervised knowledge transfer of NID models.

For further analysis, Figure 3 shows the mean PSNR value of the recovered images in the target domain versus the number of iterations during the model adaption of CPCr and VEM, respectively. It can be seen that for both the models, their performance in terms of PSNR is increased rapidly in the first 100 iterations, and afterward the performance increase becomes slower but steady. After 700 iterations, the PSNR performance tends to saturate. In other words, the proposed model adaption scheme can lead to effective and fast performance improvement of the model in target domain.

Recall that there is no explicit regularization incorporated in the proposed loss (6) and thus early stopping is expected for avoiding overfitting. But interestingly, as seen in Figure 3, the performance of the adapted models has no noticeable change after many iterations. One probable reason is that the transform-based augmentations on images and kernels induces certain implicit regularization to training which prevents overfitting, as we empirically observed noticeable performance drop along the iteration when no augmentations are used in the loss.

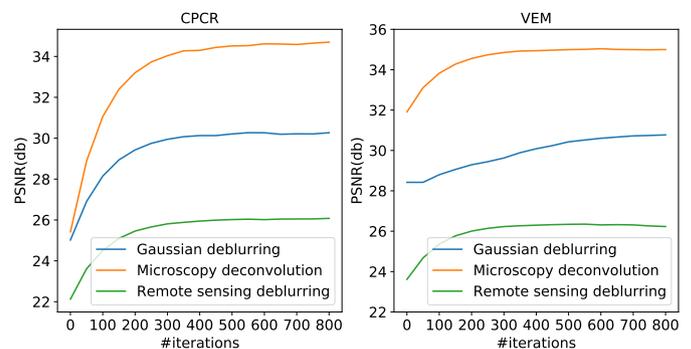


Figure 3. PSNR curves versus number of iterations during model adaption.

4.2.2. Time Cost

Table 2 lists the computing time of the adaption process for different models in three applications. It can be seen that the computational cost of the proposed adaption method is close to that of P&P and acceptable in practice.

Table 2. Computing time (mins) of P&P (first value) and proposed method (second value) for different models in three applications with noise level $\sigma = 1\%$, tested on a NVIDIA GTX 1080Ti GPU.

Applications	FDN	VEM	USRNet	CPCr	DWDN
Natural images	8.2/8.3	10.4/10.5	12.5/12.7	13.9/14.2	10.1/10.3
Microscopy	9.8/10.4	10.8/11.0	11.9/12.8	14.5/15.8	10.1/10.4
Remote sensing	8.0/8.2	9.8/10.1	11.7/12.6	13.6/14.0	8.9/9.3

4.3. Ablation Studies and Comparison to Supervised Training

We remove two key parts in the proposed method, *i.e.*, the augmentation transform $\mathcal{T}_1, \mathcal{T}_2$ by setting to an identity mapping respectively. The results listed in Table 3, denoted by ‘‘w/o \mathcal{T}_1 ’’ and ‘‘w/o \mathcal{T}_2 ’’, show that both of them play important roles, and \mathcal{T}_1 on images leads to more PSNR gain than \mathcal{T}_2 on kernels.

A further analysis is conducted via the comparison to supervised training. We simulate the oracle case where the target domain contains GT images and then train the supervised models on the target domain, which are denoted by "Model(T_I, T_K)". The results on microscopy deconvolution are listed in Table 3. We also include the WF model (Pronina et al., 2020) supervisedly trained on the dataset for comparison. Surprisingly, by transferring source-domain knowledge, the adapted models even outperform the target-domain supervised models when $\sigma = 1\%$. This is probably due to the overfitting of supervised models on the small data size in the target domain. In contrast, the adapted models can exploit both source-domain and target-domain training data, which alleviates overfitting.

Further, we use source-domain images and target-domain kernels for supervised training without model adaptation. We generate the blurred natural images using the clear images in BSDS500 and the kernels and noise settings in microscopic deconvolution. This yields many blurred/clear training pairs. The resulting models are denoted by "Model(S_I, T_K)". See Table 3 for the results, where such a supervised training yields noticeably worse results than our adaptation.

Table 3. Quantitative results in comparison to baselines and supervised training in terms of mean PSNR(dB)/SSIM on microscopic deconvolution. Best results among all compared models in each setting are boldfaced.

Model	$\sigma = 1\%$	$\sigma = 5\%$	$\sigma = 10\%$
VEM(S)	36.06/0.88	32.19/0.80	27.01/0.63
VEM(A)	38.31/ 0.92	35.08/0.87	33.03/0.84
VEM(A) w/o \mathcal{T}_1	37.88/0.91	34.46/0.86	32.22/0.83
VEM(A) w/o \mathcal{T}_2	37.94/0.91	34.79/0.86	32.54/0.84
VEM(S_I, T_K)	37.92/ 0.92	34.49/0.87	31.75/0.81
VEM(T_I, T_K)	38.44/0.92	35.18/0.88	33.23/0.84
DWDN(S)	37.54/0.90	34.67/0.86	32.64/0.84
DWDN(A)	38.44/0.92	35.16/0.87	33.41/ 0.85
DWDN(A) w/o \mathcal{T}_1	37.82/0.90	34.84/0.87	33.13/0.84
DWDN(A) w/o \mathcal{T}_2	37.98/0.90	34.99/0.87	33.21/0.84
DWDN(S_I, T_K)	37.77/0.80	34.83/0.87	33.11/0.83
DWDN(T_I, T_K)	38.01/ 0.92	35.28/0.88	33.55/0.85
WF	37.81/ 0.92	34.58/0.87	32.60/0.84

5. Conclusion

This work developed an unsupervised scheme for adapting an NID model trained on one domain to another so as to improve the performance. Such a scheme does not require any GT image in the target domain, while showing effectiveness in the experiments. It has values for practical applications and may inspire other model adaption techniques for image processing. In future, we will give further theoretical and experimental analysis to our approach and seek further improvement.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [grant numbers 61872151, 62072188] and the Natural Science Foundation of Guangdong Province [grant number 2022A1515011755].

References

- Al-Kofahi, Y., Zaltsman, A., Graves, R., Marshall, W., Rusu, M., 2018. A deep learning-based algorithm for 2-d cell segmentation in microscopy images. *BMC Bioinformatics* 19, 1–11.
- Arbelaez, P., Maire, M., Fowlkes, C., Malik, J., 2010. Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 898–916.
- Barbano, R., Kereta, Z., Hauptmann, A., Arridge, S.R., Jin, B., 2021. Unsupervised knowledge-transfer for learned image reconstruction. *arXiv preprint arXiv:2107.02572*.
- Bigdeli, S.A., Jin, M., Favaro, P., Zwicker, M., 2017. Deep mean-shift priors for image restoration, in: *Proc. NeurIPS*, pp. 763–772.
- Chen, L., Zhang, J., Pan, J., Lin, S., Fang, F., Ren, J.S., 2021. Learning a non-blind deblurring network for night blurry images, in: *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, pp. 10542–10550.
- Danielyan, A., Katkovnik, V., Egiazarian, K., 2011. Bm3d frames and variational image deblurring. *IEEE Trans. Image Process.* 21, 1715–1728.
- Deng, W., Su, Z., Qiu, Q., Zhao, L., Kuang, G., Pietikäinen, M., Xiao, H., Liu, L., 2021. Deep ladder reconstruction-classification network for unsupervised domain adaptation. *Pattern Recognit. Lett.* 152, 398–405.
- Dong, J., Pan, J., 2021. Deep outlier handling for image deblurring. *IEEE Trans. Image Process.* 30, 1799–1811.
- Dong, J., Roth, S., Schiele, B., 2020. Deep wiener deconvolution: Wiener meets deep learning for image deblurring. *Proc. NeurIPS* 33.
- Dong, J., Roth, S., Schiele, B., 2021. Learning spatially-variant map models for non-blind image deblurring, in: *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, pp. 4886–4895.
- Dong, W., Wang, P., Yin, W., Shi, G., 2019. Denoising prior driven deep neural network for image restoration. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 2305–2318.
- Eboli, T., Sun, J., Ponce, J., 2020. End-to-end interpretable learning of non-blind image deblurring, in: *Proc. Euro. Conf. Comput. Vision*.
- Gao, Z., Shen, C., Xie, C., 2018. Stacked convolutional auto-encoders for single space target image blind deconvolution. *Neurocomput.* 313, 295–305.
- Gilton, D., Ongie, G., Willett, R., 2020. Neumann networks for linear inverse problems in imaging. *IEEE Trans. Comput. Imaging* 6, 328–343.
- Gilton, D., Ongie, G., Willett, R., 2021. Model adaptation for inverse problems in imaging. *IEEE Trans. Comput. Imaging* 7, 661–674.
- Gong, D., Zhang, Z., Shi, Q., Hengel, A.v.d., Shen, C., Zhang, Y., 2020. Learning deep gradient descent optimization for image deconvolution. *IEEE Trans. Neural Netw. Learn. Syst.*
- Jin, M., Roth, S., Favaro, P., 2017. Noise-blind image deblurring, in: *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, pp. 3834–3842.
- Kruse, J., Rother, C., Schmidt, U., 2017. Learning to push the limits of efficient fft-based image deconvolution, in: *Proc. IEEE/CVF Int. Conf. Comput. Vision*, pp. 4586–4594.
- Levin, A., Weiss, Y., Durand, F., Freeman, W.T., 2011. Efficient marginal likelihood optimization in blind deconvolution, in: *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, pp. 2657–2664.
- Liu, N., Du, Y., Xu, Y., 2018. Qr codes blind deconvolution algorithm based on binary characteristic and l0 norm minimization. *Pattern Recognit. Lett.* 111, 117–123.
- Liu, N., Zheng, X., Sun, H., Tan, X., 2013. Two-dimensional bar code out-of-focus deblurring via the increment constrained least squares filter. *Pattern Recognit. Lett.* 34, 124–130.
- Marnissi, M.A., Fradi, H., Sahbani, A., Amara, N.E.B., 2022. Unsupervised thermal-to-visible domain adaptation method for pedestrian detection. *Pattern Recognit. Lett.* 153, 222–231.
- Meinhardt, T., Moller, M., Hazirbas, C., Cremers, D., 2017. Learning proximal operators: Using denoising networks for regularizing inverse imaging problems, in: *Proc. IEEE/CVF Int. Conf. Comput. Vision*, pp. 1781–1790.
- Mignotte, M., 2008. A non-local regularization strategy for image deconvolution. *Pattern Recognit. Lett.* 29, 2206–2212.
- Mosleh, A., Sola, Y.E., Zargari, F., Onzon, E., Langlois, J.P., 2017. Explicit ringing removal in image deblurring. *IEEE Trans. Image Process.* 27, 580–593.
- Mou, C., Wang, Q., Zhang, J., 2022. Deep generalized unfolding networks for image restoration, in: *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, pp. 17399–17410.
- Nan, Y., Ji, H., 2020. Deep learning for handling kernel/model uncertainty in image deconvolution, in: *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, pp. 2388–2397.

- Nan, Y., Quan, Y., Ji, H., 2020. Variational-em-based deep learning for noise-blind image deblurring, in: Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit., pp. 3626–3635.
- Pronina, V., Kokkinos, F., Dylov, D.V., Lefkimmiatis, S., 2020. Microscopy image restoration with deep wiener-kolmogorov filters. Proc. Euro. Conf. Comput. Vision .
- Quan, Y., Ji, H., Shen, Z., 2015. Data-driven multi-scale non-local wavelet frame construction and image recovery. J. Sci. Comput. 63, 307–329.
- Quan, Y., Lin, P., Xu, Y., Nan, Y., Ji, H., 2021. Nonblind image deblurring via deep learning in complex field. IEEE Trans. Neural Netw. Learn. Syst. .
- Ren, W., Zhang, J., Ma, L., Pan, J., Cao, X., Zuo, W., Liu, W., Yang, M.H., 2018. Deep non-blind deconvolution via generalized low-rank approximation. Proc. NeurIPS 31, 297–307.
- Shao, Y., Li, L., Ren, W., Gao, C., Sang, N., 2020. Domain adaptation for image dehazing, in: Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit., pp. 2808–2817.
- Soh, J.W., Cho, S., Cho, N.I., 2020. Meta-transfer learning for zero-shot super-resolution, in: Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit., pp. 3516–3525.
- Vasu, S., Maligireddy, V.R., Rajagopalan, A., 2018. Non-blind deblurring: Handling kernel uncertainty with cnns, in: Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit., pp. 3272–3281.
- Wang, W., Zhang, H., Yuan, Z., Wang, C., 2021. Unsupervised real-world super-resolution: A domain adaptation perspective, in: Proc. IEEE/CVF Int. Conf. Comput. Vision, pp. 4318–4327.
- Yang, Y., Newsam, S., 2010. Bag-of-visual-words and spatial extensions for land-use classification, in: Proc. ACM Int. Symp. Adv. Geogr. Inf. Syst., pp. 270–279.
- Yuan, Y., Li, Y., Zhu, Z., Li, R., Gu, X., 2020. Adversarial joint domain adaptation of asymmetric feature mapping based on least squares distance. Pattern Recognit. Lett. 136, 251–256.
- Zha, Z., Yuan, X., Zhou, J., Zhu, C., Wen, B., 2020. Image restoration via simultaneous nonlocal self-similarity priors. IEEE Trans. Image Process. 29, 8561–8576.
- Zhai, S., Ren, C., Wang, Z., He, X., Qing, L., 2022. An effective deep network using target vector update modules for image restoration. Pattern Recognit. 122, 108333.
- Zhang, J., Pan, J., Lai, W.S., Lau, R., Yang, M.H., 2017a. Learning fully convolutional networks for iterative non-blind deconvolution, in: Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit., pp. 6969–6977.
- Zhang, K., Gool, L.V., Timofte, R., 2020. Deep unfolding network for image super-resolution, in: Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit., pp. 3217–3226.
- Zhang, K., Zuo, W., Gu, S., Zhang, L., 2017b. Learning deep cnn denoiser prior for image restoration, in: Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit., pp. 2808–2817.
- Zhang, Y., Zhu, Y., Nichols, E., Wang, Q., Zhang, S., Smith, C., Howard, S., 2019. A poisson-gaussian denoising dataset with real fluorescence microscopy images, in: Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit., pp. 11710–11718.