Blind Image Deconvolution Using Variational Deep Image Prior

Dong Huo, Abbas Masoumzadeh, Rafsanjany Kushol, and Yee-Hong Yang, Senior Member, IEEE

Abstract—Conventional deconvolution methods utilize hand-crafted image priors to constrain the optimization. While deep-learning-based methods have simplified the optimization by end-to-end training, they fail to generalize well to blurs unseen in the training dataset. Thus, training image-specific models is important for higher generalization. Deep image prior (DIP) provides an approach to optimize the weights of a randomly initialized network with a single degraded image by maximum a posteriori (MAP), which shows that the architecture of a network can serve as the hand-crafted image prior. Unlike conventional hand-crafted image priors, which are obtained through statistical methods, finding a suitable network architecture is challenging due to the unclear relationship between images and their corresponding architectures. As a result, the network architecture cannot provide enough constraint for the latent sharp image. This paper proposes a new variational deep image prior (VDIP) for blind image deconvolution, which exploits additive hand-crafted image priors on latent sharp images and approximates a distribution for each pixel to avoid suboptimal solutions. Our mathematical analysis shows that the proposed method can better constrain the optimization. The experimental results further demonstrate that the generated images have better quality than that of the original DIP on benchmark datasets.

Index Terms—Blind image deconvolution, deep image prior, hand-crafted image prior, variational auto-encoder.

I. INTRODUCTION

BLIND image deconvolution is aimed at recovering the latent sharp image based on a single blurred image without knowing the blur kernel. When the blur kernel is spatially invariant, it can be modeled as

\[ I_b = k \otimes I_s + n, \]

(1)

where \( I_b \) denotes the blurred image, \( k \) the blur kernel, \( \otimes \) the convolution operator, \( I_s \) the latent sharp image and \( n \) the additive noise. Most conventional methods utilize maximum a posteriori (MAP) to alternatively solve for \( k \) and \( I_s \), which is formulated as

\[ \arg \max_{I_s,k} P(I_s,k|I_b) = \arg \max_{I_s,k} P(I_b|I_s,k)P(I_s)P(k), \]

(2)

where \( P(I_b|I_s,k) \) is the likelihood term, \( P(I_s) \) and \( P(k) \) are the prior distributions of the latent sharp image and the blur kernel, respectively.

Conventional methods propose various priors to solve the problem \([1],[2],[3],[4],[5],[6],[7],[8],[9],[10],[11],[12],[13],[14],[15],[16],[17],[18]\). Among them, the sparse image prior is one of the most widely used priors in image deconvolution, which includes special cases such as the Gaussian prior \([16]\), the total variational (TV) prior \([3]\), and the hyper-Laplacian prior \([7]\). Fergus et al. \([19]\) illustrate experimentally that the sparse image prior with MAP (sparse MAP) often removes almost all of the gradients. Levin et al. \([10]\) also demonstrate that the sparse MAP is more likely to generate the original blurred image than the latent sharp image when normalizing the blur kernel. In other words, the estimated kernel is more likely to be a delta kernel. Even when the estimated kernel is not a delta kernel, the method is easy to be trapped at a local minimum and hard to escape. Delayed normalization \([4]\) can avoid the delta kernel but still suffers from getting trapped at a local minimum. Edge reweighting \([20]\) and edge-selection \([16]\), which need carefully chosen hyper-parameters, are utilized to address these problems by removing small edges and noise before estimating the kernel. Variational Bayesian (VB) based methods \([6],[19]\) remit the issues of the sparse MAP by considering the standard deviation of images.

Recently, deep-learning-based methods \([21],[22],[23],[24],[25],[26],[27],[28],[29],[30],[31],[32],[33],[34]\) have been applied to this problem, which can implicitly learn the image prior within the network by training on a large dataset. Due to the high dependency on the training datasets, deep-learning-based methods do not generalize well to some image-specific information \([35]\) (e.g., blur kernels and features) which is not encountered during training. Thus, it is necessary to learn an image-specific model.

Deep image prior (DIP) \([36]\) is an appealing approach to optimizing a network using a single degraded image. Indeed, the architecture of a generator network can capture a low-level image prior for image restoration. Ren et al. \([37]\) utilize the DIP to handle blind image deconvolution, and formulate the problem as

\[ \arg \max_{I_s,k,\theta_f,\theta_h} P(I_s,k,\theta_f,\theta_h|I_b) = \arg \max_{I_s,k,\theta_f,\theta_h} P(I_b|I_s,k)P(I_s|\theta_f)P(k|\theta_h)P(\theta_f)P(\theta_h), \]

(3)
where θ₁ and θ₂ denote the parameters of the image generator \( G_I() \) and of the kernel generator \( G_k() \), respectively, \( P(θ₁) \) and \( P(θ₂) \) are, respectively, the priors of these parameters, and \( P(I_{θ₁}) \) and \( P(k|θ₂) \), respectively, are the image and the kernel prior learned by \( G_I() \) and \( G_k() \). Since they assume that \( P(θ₁) \) and \( P(θ₂) \) are constant, there is no constraint on the generated image and the kernel. As a result, it is not surprising that the outputs are suboptimal. One solution is to apply the sparse image prior to constrain \( P(I_{θ₁}) \), but the method still suffers from the problems of the sparse MAP similar to that of conventional methods.

To solve the above mentioned issues of the DIP, we attempt to adopt VB-based methods to the DIP, so that not only the optimization is constrained but also the problems of the sparse MAP can be avoided. Conventional VB-based methods [6], [19], [38] utilize a trivial (e.g., Gaussian) distribution to directly approximate the posterior distribution (the left term of (2) and (3)) by minimizing the Kullbach–Leibler (KL) divergence [39] instead of using MAP. Although the accurate posterior distribution is hard to obtain, the approximated one is good enough and much more robust than the result of MAP. In order to combine the DIP with VB, we propose a new variational deep image prior (VDIP) to learn the distributions of all latent variables (sharp images and blur kernels) which is motivated by the idea of variational auto-encoder [40]. More details of the mathematical analysis of why VDIP can perform better than DIP are given in Section III.

Our contributions are summarized as follows:

- We propose a novel variational deep image prior (VDIP) for single image blind deconvolution by integrating the deep image prior and variational Bayes.
- We provide a complete derivation of our final loss function and a mathematical analysis to demonstrate that the proposed method can better constrain the optimization than that of DIP.
- Our experiments show that the proposed VDIP can significantly improve over the DIP in both quantitative results on benchmark datasets and the quality of the generated sharp images.

II. RELATED WORK

A. Blind Image Deconvolution

Some conventional single image blind deconvolution methods focus on the distribution of image gradient for sparse high-frequency information. Fergus et al. [19] propose a heavy-tailed natural image prior, which is approximated by a mixture-of-Gaussian model. Shan et al. [2] demonstrate that the ringing effect on the deblurred image results from the estimation error of the blur kernel and noise. Cho and Lee [16] utilize the bilateral filter and the shock filter to remove noise and to enhance edges. Xu and Jia [20] find that edges smaller than the kernel size are harmful to kernel estimation and propose an r-map to measure the usefulness of edges. Krishnan et al. [41] adopt the ratio of the L₁ norm and the L₂ norm to avoid the scale variant prior, which is much closer to the L₀ norm. Levin et al. [10] prove that MAP with the sparse image prior favors a blurred solution so that they approximate the marginalization of the blur kernel, which has a closed-form solution when using the Gaussian image prior. Babacan et al. [6] exploit the concave conjugate of the super-Gaussian prior and directly estimate the posterior distribution using VB to avoid the issues of the sparse MAP. Dong et al. [15] adopt a piecewise function to mimic the L₀ norm around zero and to smooth out significant outliers, which is similar to the work of Xu et al. [9]. Chen et al. [42] who enhance the sparse prior by combining the L₀ and L₁ norm. Yang et al. [38] introduce a restarting technique to further improve the performance of VB-based methods.

Some other conventional methods utilize properties of images to form priors. Michaeli and Irani [11] find that blur significantly decreases cross-scale patch recurrence. Thus, they constrain the output by minimizing the dissimilarity between nearest-neighbor patches. Lai et al. [43] assume that each local patch contains two primary colors, and the distance between them should be maximized by deconvolution. Pan et al. [12] apply the dark channel prior to handle blind deconvolution and achieve good results. Yan et al. [13] combine the bright and the dark channel priors to overcome the limitation on bright dominant images. Ren et al. [44] derive an enhanced low-rank prior to reduce the number of non-zero singular values of the image. Pan et al. [45] exploit the phase-only image of a blurred image to estimate the start and end point of the blur kernel, which is efficient for linear motion. Bai et al. [18] utilize the downsampled blurred image as the prior and recover the latent sharp image from coarse to fine. Chen et al. [17] calculate the bright channel of the gradient maps for deblurring images without enough dark and bright pixels.

Deep-learning-based methods are also applied to this problem. Chakrabarti [46] trains a network to estimate the Fourier coefficients of blur kernels. Liu et al. [47] and Zhang et al. [24] exploit recursive filters to take advantage of context information. Generative adversarial networks (GANs) are also exploited for faster convergence and better visual quality [22], [26], [30]. Gong et al. [48] adopt a network to learn the motion flow. Xu et al. [49] develop a network to generate sharp gradient maps for kernel estimation. To enhance the network output, some utilize multi-stage strategies, e.g., multi-scale [21], [25], [50], multi-patch [28], [32], [33] and multi-temporal [51]. Asim et al. [52] adopt a well-trained sharp image generator to generate the sharp image closest to the blurred one. Tran et al. [34] develop a sharp image auto-encoder and a blur representation learning network, then two well-trained networks are fixed as a deep generative prior [52]. Li et al. [23] adopt a well-trained classifier (which can distinguish blurred images and sharp images) as an extra constraint of the MAP framework, and optimize the problem with the half-quadratic splitting method similar to that used in conventional methods.

Different from [23], [34], [52], [53], [54] in which priors need to be trained on external datasets, our proposed method is optimized with only one single blurred input image and the whole framework is optimized by gradient descent instead of conventional optimization-based methods [23]. Although Asim et al. [52] also provide a method optimized with a single image, the method degenerates to the DIP [37] with a sparse image...
prior and learnable inputs, which cannot avoid the problems of the sparse MAP. As well, none of the mentioned deep-learning-based methods consider the standard deviation of the image.

B. Deep Image Prior

Ulyanov et al. [36] introduce the concept of the deep image prior (DIP) that the structure of a randomly-initialized network can be used as an image prior for image restoration tasks. Ren et al. [37] adopt the DIP to implicitly learn the image prior and the kernel prior for blind image deconvolution. Early stopping with carefully chosen time, added random noise to the input and to the gradient with fixed noise level are applied to avoid the suboptimal solution of DIP [55]. Neural architecture search (NAS) can help to search for these hyper-parameters heuristically [56], but with the substantial increase in computational cost. Double-DIP [57] can handle the image separation problems, e.g., image segmentation, image dehazing, and transparency separation, but does not perform well for blind image deconvolution [37]. Some methods stabilize the optimization by adding extra priors to the loss function [58], [59]. However, this technique only works when the degradation kernel is known.

C. Variational Auto-Encoder

Kingma et al. [40] introduce the concept of variational auto-encoder (VAE) for image generation. The goal is to learn a model that generates an image $x$ given a sampled latent variable $z$, which can be formulated as $P(z|x) = P(x|z)P(z|x)/P(z)$, where $P(x)$ is constant. Since obtaining the true distribution of $P(z|x)$ is nontrivial, they utilize a Gaussian distribution $Q(z)$ to approximate $P(z|x)$ with a network to learn the expectation and the standard deviation. Thus, the target of VAE can be converted to minimizing the KL divergence between $Q(z)$ and $P(z|x)$. Vahdat et al. [60] further stabilize the training of VAE by partitioning the latent variables into groups. Similar to image generation, the target of image deconvolution is learning a model to generate a blurred image $I_b$ given a sampled latent sharp image $I_s$ and a blur kernel $k$, and the distributions of $P(I_s|I_b)$ and $P(k|I_b)$ are learned by the network. And predefined hand-crafted $P(I_b)$ and $P(k)$ can help to constrain the optimization.

III. PROPOSED METHOD

In this section, we provide the mathematical analysis of the feasibility of our proposed methods. More derivation details are given in the supplementary materials, available online.

A. Super-Gaussian Distribution

Conventional image priors can be formulated as a super-Gaussian distribution

$$P(I_s) = W \exp\left(-\frac{\rho(F_x(I_s)) + \rho(F_y(I_s))}{2}\right),$$

where $W$ is the normalization coefficient, and $\rho(\cdot)$ is the penalty function to constrain the sparsity of $F_x(I_s)$ and $F_y(I_s)$. For sparse image priors, $F_x(\cdot)$ and $F_y(\cdot)$ are gradient kernels $[-1, 1]^T$ and $[-1, 1]$. When $\rho(\cdot)$ is quadratic, $P(I_s)$ degenerates to a Gaussian distribution. Since $\rho(\sqrt{x})$ has to be increasing and concave for $x \in (0, \infty)$ when $x$ follows the super-Gaussian distribution [61], we can decouple $\rho(\cdot)$ and $I_s$ using the concave conjugate of $\rho(\sqrt{F_x(I_s)})$ and $\rho(\sqrt{F_y(I_s)})$ following the strategy of Babacan et al. [6], and the upper bound of $\rho(F_x(I_s))$ and of $\rho(F_y(I_s))$ are represented as

$$\rho(F_x(I_s)) \leq \frac{1}{2}\xi_x(F_x(I_s))^2 - \rho^*(\frac{1}{2}\xi_x),$$

$$\rho(F_y(I_s)) \leq \frac{1}{2}\xi_y(F_y(I_s))^2 - \rho^*(\frac{1}{2}\xi_y),$$

(5)

where $\rho^*(\frac{1}{2}\xi_x)$ and $\rho^*(\frac{1}{2}\xi_y)$ denote the concave conjugates of $\rho(\sqrt{F_x(I_s)})$ and $\rho(\sqrt{F_y(I_s)})$, respectively, and $\xi_x$ and $\xi_y$ are the variational parameters. We replace $\rho(F_x(I_s))$ and $\rho(F_y(I_s))$ in (4) with their upper bounds in $P(I_s)$

$$P(I_s) \geq W \exp\left(-\frac{\xi_x(F_x(I_s))^2 + \xi_y(F_y(I_s))^2}{4}\right) \cdot \exp\left(\frac{\rho^*(\frac{1}{2}\xi_x) + \rho^*(\frac{1}{2}\xi_y)}{2}\right).$$

(6)

Since the right-hand side of each inequality in (5) is a convex quadratic function with a single global minimum, by calculating the derivative with respect to $F_x(I_s)$ and to $F_y(I_s)$, respectively, in (5), equality is attained when

$$\xi_x = \frac{\rho^*(F_x(I_s))}{\|F_x(I_s)\|^2}, \quad \xi_y = \frac{\rho^*(F_y(I_s))}{\|F_y(I_s)\|^2},$$

(7)

where $\rho^*(\cdot)$ is the derivative of $\rho(\cdot)$. As shown in (6), irrespective of the form of $\rho(\cdot)$, $P(I_s|\xi_x, \xi_y)$ becomes a trivial Gaussian distribution when equality is attained, which simplifies the derivation and the implementation because other penalty functions are discontinuous and the integral is too complicated to obtain (e.g., $|x|, \ln |x|$). Besides, a Gaussian distribution is usually utilized to approximate the real distribution in VB-based methods, and the multiplication of two Gaussian distributions is much easier to calculate.

B. Variational Inference

Due to the extra variational parameters $\xi_x$ and $\xi_y$, the problem can be reformulated as

$$\arg\max_{I_s, k, \xi_x, \xi_y} P(I_s, k, \xi_x, \xi_y|I_b)$$

$$= \arg\max_{I_s, k, \xi_x, \xi_y} \frac{P(I_b|I_s, k)P(I_s|\xi_x, \xi_y)P(\xi_x, \xi_y)P(k)}{P(I_b)}.$$  

(8)

Directly calculating $P(I_s, k, \xi_x, \xi_y|I_b)$ is challenging because the true distribution of $I_b$ is difficult to obtain. The most common strategy is to use MAP, which estimates the posterior distribution by maximizing it as shown in (8). However, as mentioned in Section I, MAP with the sparse image prior favors a trivial solution. An alternative strategy is to use VB, which uses a trivial distribution $Q(I_s, k, \xi_x, \xi_y)$ (e.g., Gaussian) to approximate the posterior distribution $P(I_s, k, \xi_x, \xi_y|I_b)$ by minimizing the KL
divergence between these two distributions, which can be written as
\[
D_{KL}(Q(I_s, k, \xi_x, \xi_y) || P(I_s, k, \xi_x, \xi_y | I_b)) = \ln P(I_b) \\
- \int Q(I_s, k, \xi_x, \xi_y) \ln \frac{P(I_s, k, \xi_x, \xi_y, I_b)}{Q(I_s, k, \xi_x, \xi_y)} dI_s dk d\xi_x d\xi_y \\
= \ln P(I_b) - L(I_s, k, \xi_x, \xi_y, I_b),
\]
where \( D_{KL} \) represents the KL divergence, and \( L(I_s, k, \xi_x, \xi_y, I_b) \) is the variational lower bound. Since \( \ln P(I_b) \) is constant and \( D_{KL} \) is non-negative, minimizing \( D_{KL} \) is equivalent to maximizing \( L(I_s, k, \xi_x, \xi_y, I_b) \). By assuming that the \( I_s \) and \( k \) are independent, the variational lower bound can be rewritten as
\[
L(I_s, k, \xi_x, \xi_y, I_b) = \int Q(k) \ln \frac{P(k)}{Q(k)} dk - \int Q(I_s) \ln Q(I_s) dI_s \\
+ \int Q(I_s) Q(\xi_x, \xi_y) \ln P(I_s | \xi_x, \xi_y) dI_s d\xi_x d\xi_y \\
+ \int Q(\xi_s, \xi_y) \ln \frac{P(\xi_x, \xi_y)}{Q(\xi_x, \xi_y)} d\xi_x d\xi_y \\
+ E_{Q(I_s, k)} \ln P(I_b | I_s, k),
\]
where \( P(I_s | \xi_x, \xi_y) \) can be obtained from (6). \( P(k) \) is set as the standard Gaussian distribution \( \mathcal{N}(0, I) \). Based on the mean field theory [6], [62], it is more convenient to simply assume that pixels on images and kernels are all independent. We can further rewrite (10) as
\[
L(I_s, k, \xi_x, \xi_y, I_b) = \frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} (2 \ln S(k(i, j)) - E^2(k(i, j)) - S^2(k(i, j))) \\
+ \frac{1}{2} \sum_{m=1}^{M} \sum_{n=1}^{N} 2 \ln S(I_s(m, n)) \\
- \frac{1}{4} \sum_{m=1}^{M} \sum_{n=1}^{N} E((F_x(I_s)(m, n))^2) E(\xi_x(m, n)) \\
- \frac{1}{4} \sum_{m=1}^{M} \sum_{n=1}^{N} E((F_y(I_s)(m, n))^2) E(\xi_y(m, n)) \\
+ E_{Q(I_s, k)} \ln P(I_b | I_s, k) \\
+ \int Q(\xi_x, \xi_y) \ln \frac{P(\xi_x, \xi_y)}{Q(\xi_x, \xi_y)} d\xi_x d\xi_y \\
+ \frac{1}{2} \sum_{m=1}^{M} \sum_{n=1}^{N} (\rho^2(\frac{1}{2} \xi_x) + \rho^2(\frac{1}{2} \xi_y)) d\xi_x d\xi_y \\
+ \text{Constant},
\]
where \( S() \) and \( E() \) denote the standard deviation and the expectation, respectively, of distribution \( Q() \), \((i, j)\) is the pixel index of \( k \), \((m, n)\) is the pixel index of \( I_s \) and \( \xi \). Since only the expectation of \( \xi_x \) and \( \xi_y \) are related to \( I_s \), we do not need to consider their distributions so that the last three rows in (11) can be ignored. Following Babacan et al. [6], \( E(\xi_x) \) and \( E(\xi_y) \) can be simply calculated by
\[
E(\xi_x(m, n)) = \frac{\rho(v_x(m, n))}{v_x(m, n)}, \\
E(\xi_y(m, n)) = \frac{\rho(v_y(m, n))}{v_y(m, n)},
\]
\[
v_x(m, n) = \sqrt{E((F_x(I_s)(m, n))^2)} , \\
v_y(m, n) = \sqrt{E((F_y(I_s)(m, n))^2)} .
\]
For the sparse image prior, \( F_x(I_s)(m, n) \) and \( F_y(I_s)(m, n) \) can be reformulated as
\[
F_x(I_s)(m, n) = I_s(m, n) - I_s(m - 1, n), \\
F_y(I_s)(m, n) = I_s(m, n) - I_s(m, n - 1),
\]
where \( I_s(0, \cdot) \) and \( I_s(\cdot, 0) \) denote paddings.

Our VDIP can also be extended to the extreme channel prior. For the extreme channel prior, \( F_x(I_s)(m, n) \) and \( F_y(I_s)(m, n) \) can be reformulated as
\[
F_x(I_s)(m, n) = \min_{i \in \Omega(m, n)} \left( \min_{c \in \{r, g, b\}} (I^c_s(i)) \right), \\
F_y(I_s)(m, n) = 1 - \max_{i \in \Omega(m, n)} \left( \max_{c \in \{r, g, b\}} (I^c_s(i)) \right)
\]
where \( \Omega(m, n) \) denotes a local patch centered at \((m, n)\), and \( I^c_s \) is a color channel of \( I_s \).

Further derivation of \( E((F_x(I_s)(m, n))^2) \) and \( E((F_y(I_s)(m, n))^2) \) are shown in the supplementary materials, available online.

C. Variational Deep Image Prior

Conventional variational inference solves (11) by calculating the closed-form expectation with respect to each variable over all the other variables to get the distribution [62], but it is challenging to apply this strategy to deep learning since the networks are highly non-convex. Hence, we use two networks to learn the distribution of the latent sharp image and the blur kernel, respectively, in an unsupervised manner. For simplification, we assume that the standard deviation of the blur kernel \( S(k) \) is constant. We also assume that the additive noise is white Gaussian noise. Then, we only need to learn the expectation of the image \( E(I_s) \), the expectation of the kernel \( E(k) \), and the standard deviation of the image \( S(I_s) \).

We utilize an encoder-decoder as the image generator \( G_I() \), a fully-connected network as the kernel generator \( G_k() \), and random noises \( Z_I \) and \( Z_k \) as inputs. The image generator outputs both \( E(I_s) \) and \( S(I_s) \), and the kernel generator outputs \( E(k) \). We can now approximate \( E_{Q(I_s, k)} \ln P(I_b | I_s, k) \) in (10) and (11) by Monte Carlo estimation using sampling [40]
\[
E_{Q(I_s, k)} \ln P(I_b | I_s, k) \approx \frac{1}{A} \sum_{a=1}^{A} \frac{||I_b - k \otimes I^a_s||^2}{2 \sigma^2},
\]
Algorithm 1: Blind Image Deconvolution Using Variational Deep Image Prior.

**Input:** blurred image $I_b$, image generator $G_I(\cdot)$, kernel generator $G_k(\cdot)$

**Output:** estimated sharp image $I^*_s$ and blur kernel $k^*$

**Initialization:** fixed noise inputs $z_I$ and $z_k$, parameters of two generators $\theta_I^{(0)}$ and $\theta_k^{(0)}$ to be optimized

for $t = 1, 2, \ldots, T$ do
1: generate $E(I_s^{(t)})$, $S(I_s^{(t)})$ by $G_I(z_I, \theta_I^{(t-1)})$ and $E(k^{(t)})$ by $G_k(z_k, \theta_k^{(t-1)})$
2: calculate $E(\xi_x^{(t)})$ and $E(\xi_y^{(t)})$ using (12)
3: sample $\tilde{I}_s^{(t)}$ $A$ times and approximate $E_Q(I_s^{(t)}, k^{(t)})$ using (16)
4: calculate $L(I_s^{(t)}, k, \xi_x, \xi_y, I_b^{(t)})$ using (11)
5: update $\theta_I^{(t-1)}$ and $\theta_k^{(t-1)}$ by maximizing $L(I_s^{(t)}, k, \xi_x, \xi_y, I_b^{(t)})$
end for

$E(I_s^{(T+1)})$, $S(I_s^{(T+1)}) = G_I(z_I, \theta_I^{(T)})$

$E(k^{(T+1)}) = G_k(z_k, \theta_k^{(T)})$

$I^*_s = E(I_s^{(T+1)})$, $k^* = E(k^{(T+1)})$

and $S(I_s)$ in our VDIP, and the target is minimizing the mean square error $\|I_b - E(k) \otimes E(I_s)\|^2_2$. The target of the DIP only focuses on maximizing $P(I_s|I_b, k)$ in (3), so that $P(I_s|\theta_I)$ and $P(I_s|\theta_k)$ are not properly constrained. In contrast, in our proposed method, we apply a Gaussian prior and a sparse image prior to constrain $P(I_s|\theta_I)$ and $P(I_s|\theta_k)$, respectively, as shown in (11). Simply exploiting the additive priors for optimizing (3) can lead to suboptimal solutions of sparse MAP. Thus, we adopt the VB to avoid such a problem by introducing the standard deviation $S(I_s)$ to the optimization target. It is noteworthy that (11) degenerates to the sparse MAP when we fix $S(I_s)$ as zero. It shows the limitation of optimizing the sparse MAP that its solution is difficult to achieve a large variational lower bound, because $\ln S(I_s)$ is negative infinity. The VB can nicely avoid it by considering non-zero $S(I_s)$. Besides, the values of $E(\xi)$ act as the penalty weights of gradients. In particular, small weights for large gradients and large weights for small gradients. Zero $S(I_s)$ may result in over-penalty in regions with small gradients.

IV. EXPERIMENTS

A. Implementation Details

Our proposed method is implemented in PyTorch [63] and evaluated on a single RTX A6000 GPU with 48 GB of memory. The learning rate of the image generator and of the kernel generator are set as $1 \times 10^{-2}$ and $1 \times 10^{-4}$, respectively, and the number of optimization steps $T$ is $5,000$. In (16), the number of samples $A$ is set as 1. We use $\ln |x|$ as our penalty function $\rho(x)$. Note that the architectures of $G_I(\cdot)$ and $G_k(\cdot)$ are the same as those of DIP [37] for fair comparison, except the output layers of $G_I(\cdot)$ are doubled (half for $E(I_s)$ and half for $S(I_s)$). Different from the original DIP [37] that adds additive random Gaussian noise to $Z_I$ and $Z_k$ to avoid the local minima, we do not add additive random noise to the inputs. The source code of our VDIP is available at https://github.com/Dong-Huo/VDIP-Deconvolution.

B. Quantitative Comparison

We first evaluate different versions of DIP including our VDIP for image deconvolution on the synthetic dataset from Lai et al. [64] and compare with several conventional methods including Cho and Lee [16], Levin et al. [65], Krishnan et al. [41], Xu et al. [9], Perrone et al. [66], Michaeli and Irani [11], Pan et al.
TABLE I

| Method           | Manmade | Natural | People | Saturated | Text | Average |
|------------------|---------|---------|--------|-----------|------|---------|
| Cho et al. [16]  | 17.08/0.482 | 21.15/0.615 | 20.96/0.630 | 14.32/0.531 | 16.01/0.522 | 17.91/0.556 |
| Levin et al. [65] | 15.12/0.284 | 18.76/0.419 | 19.55/0.528 | 13.98/0.457 | 14.44/0.372 | 16.37/0.418 |
| Krishnan et al. [9] | 16.32/0.476 | 20.13/0.587 | 22.59/0.709 | 14.41/0.545 | 15.78/0.518 | 17.85/0.567 |
| Xu et al. [9]    | 19.11/0.686 | 22.70/0.754 | 26.42/0.856 | 14.97/0.586 | 20.36/0.769 | 20.75/0.734 |
| Perrone et al. [66] | 18.66/0.676 | 22.78/0.766 | 24.79/0.828 | 14.46/0.531 | 18.55/0.673 | 19.81/0.699 |
| Michaeli et al. [11] | 18.27/0.509 | 21.93/0.614 | 25.74/0.791 | 14.46/0.539 | 16.59/0.503 | 19.40/0.591 |
| Pan et al. [12]  | 20.00/0.714 | 24.47/0.801 | 26.70/0.811 | 17.46/0.680 | 21.13/0.762 | 21.95/0.753 |
| Dong et al. [15] | 18.88/0.567 | 23.42/0.702 | 25.53/0.769 | 16.72/0.611 | 20.05/0.682 | 20.92/0.666 |
| Tao et al. [25]  | 17.11/0.381 | 20.18/0.492 | 22.12/0.651 | 15.41/0.545 | 17.56/0.649 | 18.12/0.508 |
| Kupyn et al. [30] | 17.47/0.414 | 20.71/0.520 | 22.71/0.682 | 15.67/0.565 | 16.22/0.503 | 18.35/0.537 |
| Wen et al. [67]  | 18.06/0.550 | 22.51/0.669 | 25.99/0.769 | 17.79/0.672 | 18.75/0.598 | 20.36/0.652 |
| Zamir et al. [33] | 17.12/0.392 | 20.30/0.506 | 21.50/0.631 | 15.49/0.547 | 17.45/0.415 | 17.83/0.498 |
| Huo et al. [68]  | 17.11/0.380 | 20.27/0.495 | 21.69/0.636 | 15.45/0.545 | 15.84/0.478 | 18.07/0.507 |
| Zamir et al. [69] | 17.19/0.369 | 20.26/0.495 | 21.67/0.636 | 15.52/0.545 | 15.36/0.460 | 18.00/0.505 |
| Chen et al. [70]  | 16.89/0.371 | 20.10/0.484 | 21.51/0.642 | 15.39/0.544 | 14.87/0.401 | 17.79/0.488 |

TABLE II

| Average Kernel Recovery Error on the Synthetic Dataset from Lai et al. [64] |
|------------------|---------|---------|--------|-----------|------|---------|
| Method           | Manmade | Natural | People | Saturated | Text | Average |
| Cho et al. [16]  | 0.00138 | 0.00123 | 0.00145 | 0.00164 | 0.00139 | 0.00141 |
| Levin et al. [65] | 0.00099 | 0.00107 | 0.00117 | 0.00124 | 0.00117 | 0.00113 |
| Krishnan et al. [41] | 0.00125 | 0.00114 | 0.00128 | 0.00134 | 0.00118 | 0.00124 |
| Xu et al. [9]    | 0.00114 | 0.00084 | 0.00073 | 0.00144 | 0.00074 | 0.00098 |
| Perrone et al. [66] | 0.00108 | 0.00091 | 0.00111 | 0.00135 | 0.00102 | 0.00109 |
| Michaeli et al. [11] | 0.00131 | 0.00118 | 0.00102 | 0.00169 | 0.00146 | 0.00134 |
| Pan et al. [12]  | 0.00078 | 0.00060 | 0.00083 | 0.00099 | 0.00072 | 0.00078 |
| Dong et al. [15] | 0.00097 | 0.00078 | 0.00096 | 0.00111 | 0.00082 | 0.00093 |
| Wen et al. [67]  | 0.00113 | 0.00092 | 0.00089 | 0.00074 | 0.00098 | 0.00093 |

TABLE III

| Quantitative Comparison on the Real Blurred Dataset from Lai et al. [64] |
|------------------|---------|---------|--------|-----------|------|---------|
| Method           | N IQE ( ) | BRIQUE ( ) | P IQE ( ) |
| Cho et al. [16]  | 4.0350 | 36.2829 | 48.6227 |
| Levin et al. [65] | 3.6594 | 36.0666 | 46.7037 |
| Krishnan et al. [41] | 3.8696 | 37.9942 | 50.4024 |
| Xu et al. [9]    | 3.9536 | 37.3240 | 49.5436 |
| Perrone et al. [66] | 4.0397 | 39.7997 | 51.7650 |
| Michaeli et al. [11] | 3.9852 | 35.1205 | 46.7085 |
| Pan et al. [12]  | 4.8790 | 36.3792 | 68.9470 |
| Dong et al. [15] | 4.7532 | 37.1199 | 64.1972 |
| Tao et al. [25]  | 3.5612 | 40.1954 | 53.0908 |
| Kupyn et al. [30] | 3.2937 | 35.8382 | 40.0545 |
| Wen et al. [67]  | 4.9210 | 33.1731 | 58.3326 |
| Zamir et al. [33] | 3.7926 | 42.4894 | 52.1181 |
| Huo et al. [68]  | 3.9222 | 40.1037 | 47.0717 |
| Zamir et al. [69] | 3.7401 | 42.9266 | 50.6804 |
| Chen et al. [70]  | 4.6754 | 46.3900 | 74.0267 |

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Fig. 3. Qualitative comparison on the synthetic dataset from Lai et al. [64]. The estimated blur kernels are pasted at the top-left corners of the corresponding deblurred results.

and the comparison of DIP-Extreme and DIP-Sparse follows this observation. Additionally, our VDIP-sparse takes advantage of gradient-based priors without explicitly handling the outliers of saturated images and performs even better than VDIP-Extreme, which shows the effectiveness of utilizing variational Bayes.

To evaluate the estimated kernel, we calculate the average kernel recovery error [72] and report the results in Table II. Note that the compared deep-learning-based methods do not estimate the blur kernels. Although the evaluated kernel of Pan et al. [12] is more accurate than VDIP-Extreme, our VDIP-Extreme
Fig. 4. Qualitative comparison on the real blurred dataset from Lai et al. [64]. The estimated blur kernels are pasted at the top-left corners of the corresponding deblurred results.
Fig. 5. Qualitative comparison of MAP (DIP) and VB (VDIP-Sparse). The estimated blur kernels are pasted at the top-left corners of the corresponding deblurred results where the estimated kernels of MAP are all delta kernels.

Fig. 6. Failure cases.
performs better, which demonstrate that a proper deconvolution method is important even with accurate estimated blur kernels.

We also evaluate the above mentioned methods on the real blurred dataset from Lai et al. [64]. Since there is no ground truth sharp image, we utilize three no-reference image quality assessment metrics, in particular, Naturalness Image Quality Evaluator (NIQE) [73], Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) [75], and Perception based Image Quality Evaluator (PIQE) [75] to quantitatively evaluate the results. As shown in Table III, our method can generate images of the highest quality based on BRISQUE and PIQE among all compared methods. Similar to all of the compared conventional methods and DIP, our proposed method is also designed for spatially invariant (uniform) blur. However, it even performs better than deep-learning-based methods that are trained for spatially variant blur. We think this is because the compared deep-learning-based methods are all trained on synthetic datasets where the blurred images are generated by averaging consecutive frames from a high-frame-rate video. The performance of these methods are limited on the real data with more artifacts because of the domain-shift issue.

C. Optimization Time

To evaluate the relation between the optimization time and the size of images and kernels, we run the optimization with varying image size and fixed kernel size, and then run the optimization with varying kernel size and fixed image size. All of the experiments are run on a single RTX A6000 GPU with 48 GB of memory. As shown in Fig. 2, the optimization time is proportional to the quadratic of image size and kernel size.

D. Qualitative Comparison

Some of the qualitative comparisons are shown in Figs. 3 and 4. Our VDIP-Sparse can generate sharper results with less noise and artifacts than other methods including DIP. Specifically, Pan et al. [12] are able to obtain correct blur kernels in some cases but the deconvolution results are over-smoothed. Dong et al. [15], Wen et al. [67] and DIP [37] are over-enhanced with many artifacts. Since the blur on the real images are spatially variant (non-uniform), obtaining perfect results with uniform deconvolution methods is difficult, if not impossible. But our method still performs better than Kupyn et al. [30] trained on non-uniform blurred datasets [21], showing the limited generalization ability of external training and the importance of image-specific information.

As outlined in Section I, when a sparse image prior is employed in conjunction with Maximum a Posteriori (MAP) estimation, the resulting solution favors a trivial outcome, wherein the generated kernel is a delta kernel. Fig. 5 demonstrates the effectiveness of our improved approach utilizing Variational Bayes (VB) over the MAP method. It displays the trivial solution obtained by MAP, where the estimated kernels collapse to a single white dot (delta kernel). In contrast, VB successfully avoids such solutions, resulting in more accurate estimations.

E. Failure Cases

As shown in Fig. 6, our VDIP does not perform well on small images with complex scenes, due to the lack of enough information to properly optimize the network.

V. Conclusion

In this paper, we propose a new variational deep image prior (VDIP) for blind image deconvolution, which achieves a better performance than that of the DIP. One common issue of optimizing a model using a single image is high inference time compared with methods trained on external datasets, which makes it hard to adopt the method to large testing datasets. Our method is also limited when the single degraded image cannot provide enough information. In our future work, we plan to adopt meta-learning [76] to train the networks on external datasets and fine-tune on each test image, which can take advantage of the information from other images and obtain a image-specific model with only several iterations.

ACKNOWLEDGMENTS

The authors would like to thank Steve Sutphen for his technical support and the Natural Sciences and Engineering Research Council of Canada, the Department of Computing Science, and the University of Alberta for funding. Masoumzadeh would like to thank Huawei for the Doctoral Scholarship.

REFERENCES

[1] N. Joshi, R. Szeliski, and D. J. Kriegman, “PSF estimation using sharp edge prediction,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2008, pp. 1–8.
[2] Q. Shan, J. Jia, and A. Agarwala, “High-quality motion deblurring from a single image,” ACM Trans. Graph., vol. 27, no. 3, pp. 1–10, 2008.
[3] T. F. Chan and C.-K. Wong, “Total variation blind deconvolution,” IEEE Trans. Image Process., vol. 7, no. 3, pp. 370–375, Mar. 1998.
[4] D. Perrone and P. Favaro, “A clearer picture of total variation blind deconvolution,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 6, pp. 1041–1055, Jun. 2016.
[5] F. Sroubek and J. Koterova, “Motion blur prior,” in Proc. IEEE Int. Conf. Image Process., 2020, pp. 928–932.
[6] S. D. Babacan, R. Molina, M. N. Do, and A. K. Katsaggelos, “Bayesian blind deconvolution with general sparse image priors,” in Proc. Eur. Conf. Comput. Vis., 2012, pp. 341–355.
[7] D. Krishnan and R. Fergus, “Fast image deconvolution using hyper-Laplacian priors,” in Proc. 22nd Int. Conf. Neural Inf. Process. Syst., 2009, pp. 1033–1041.
[8] N. Joshi, C. L. Zitnick, R. Szeliski, and D. J. Kriegman, “Image deblurring and denoising using color priors,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2009, pp. 1550–1557.
[9] L. Xu, S. Zheng, and J. Jia, “Unnatural l0 sparse representation for natural image deblurring,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2013, pp. 1107–1114.
[10] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, “Understanding and evaluating blind deconvolution algorithms,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2009, pp. 1964–1971.
[11] T. Michaeli and M. Irani, “Blind deblurring using internal patch recurrence,” in Proc. Eur. Conf. Comput. Vis., 2014, pp. 783–798.
[12] J. Pan, D. Sun, H. Pfister, and M.-H. Yang, “Blind image deblurring using dark channel prior,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 1628–1636.
[13] Y. Yan, W. Ren, Y. Guo, R. Wang, and X. Cao, “Image deblurring via extreme channels prior,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 6978–6986.
[14] D. Yang, X.-J. Wu, and H. Yin, “Blind image deblurring via enhanced sparse prior,” J. Electron. Imag., vol. 30, no. 2, 2021, Art. no. 023031.
S. Kullback, L. Yang and H. Ji, "A variational EM framework with adaptive edge selection for blind motion deblurring," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019, pp. 10159–10168.

[10] J. Xie, X. Sun and D. Tao, "Temporal consistency for blind motion deblurring," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2021, pp. 8877–8886.

[11] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in Proc. Nat. Conf. Mach. Learn., 2019, pp. 8564–8570.

[12] F. Perrone and P. Favaro, "Total variation blind deconvolution: The devil is in the details," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2010, pp. 2657–2664.

[13] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Efficient marginal likelihood optimization in blind deconvolution," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2011, pp. 2657–2664.

[14] T. M. Nimisha, A. Kumar Singh, and A. N. Rajagopalan, "Blur-invariant deep learning for blind-deblurring," in Proc. IEEE Int. Conf. Comput. Vis., 2019, pp. 3060–3068.

[15] J. Kong, J. Pan, Y.-Y. Zhang, and M.-H. Yang, "Motion blur kernel estimation using normalized color-line prior," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2019, pp. 874–882.

[16] S. Cho and S. Lee, "Fast motion deblurring," in Proc. ACM SIGGRAPH Asia Papers, vol. 29, no. 145, pp. 2033–2045, Jul. 2020.

[17] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, "Removing camera shake from a single photograph," in Proc. ACM SIGGRAPH Papers, 2006, pp. 787–794.

[18] L. Xu and J. Jia, "Two-phase kernel estimation for robust motion deblurring," in Proc. Eur. Conf. Comput. Vis., 2010, pp. 150–170.

[19] S. Nah, T. Hyun Kim, and K. Mu Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2017, pp. 257–265.

[20] T. M. Nimisha, A. Kumar Singh, and A. N. Rajagopalan, "Blur-invariant deep learning for blind-deblurring," in Proc. IEEE Int. Conf. Comput. Vis., 2017, pp. 4762–4770.

[21] L. Li, J. Pan, W.-S. Lai, C. Gao, N. Sang, and M.-H. Yang, "Learning a discriminative prior for blind image deblurring," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2018, pp. 6616–6625.

[22] T. M. Nimisha, A. Kumar Singh, and A. N. Rajagopalan, "Blur-invariant deep learning for blind-deblurring," in Proc. IEEE Int. Conf. Comput. Vis., 2019, pp. 10159–10168.

[23] L. Chen, F. Fang, T. Wang, and G. Zhang, "Blind image deblurring with local maximum gradient prior," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019, pp. 1742–1750.

[24] Y. Bai, H. Jia, M. Jiang, X. Liu, X. Xie, and W. Gao, "Single-image blind deblurring using multi-scale latent structure prior," in Proc. IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 7, pp. 2033–2045, Jul. 2020.

[25] K. Purohit and A. Rajagopalan, "Region-adaptive dense network for blind image deblurring," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019, pp. 10217–10226.

[26] L. Chen, F. Fang, S. Lei, F. Li, and G. Zhang, "Enhanced sparse model for blind deblurring," in Proc. Eur. Conf. Comput. Vis., 2020, pp. 631–646.

[27] W.-S. Lai, J.-J. Ding, Y.-Y. Lin, and Y.-Y. Chuang, "Blur kernel estimation using normalized color-line prior," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2015, pp. 1044–1052.

[28] W. Ren, X. Cao, J. Pan, X. Gao, W. Zao, and M.-H. Yang, "Image deblurring via enhanced low-rank prior," IEEE Trans. Image Process., vol. 25, no. 7, pp. 3426–3437, Jul. 2016.

[29] L. Pan, R. Hartley, M. Liu, and Y. Dai, "Phase-only image based kernel estimation for single image blind deblurring," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2015, pp. 6027–6036.

[30] A. Chakrabarti, "A neural approach to blind motion deblurring," in Proc. Eur. Conf. Comput. Vis., 2016, pp. 221–235.

[31] S. Liu, J. Pan, and M.-H. Yang, "Learning recursive filters for low-level vision via a hybrid neural network," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 560–576.

[32] D. Gong et al., "From motion blur to motion flow: A deep learning solution for removing heterogeneous motion blur," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 3806–3815.

[33] X. Xu, J. Pan, Y.-Y. Zhang, and M.-H. Yang, "Motion blur kernel estimation via deep learning," IEEE Trans. Image Process., vol. 27, no. 1, pp. 194–205, Jan. 2018.

[34] J. Liu, W. Sun, and M. Li, "Recurrent conditional generative adversarial network for image deblurring," IEEE Trans. Image Process., vol. 7, no. 6, pp. 6186–6193, 2018.

[35] D. Park, D. U. Kang, J. Kim, and S. Y. Chun, "Multi-temporal recurrent neural networks for progressive non-uniform single image deblurring with incremental temporal training," in Proc. Eur. Conf. Comput. Vis., 2020, pp. 327–343.

[36] M. Asim, F. Shamshad, and A. Ahmed, "Blind image deconvolution using deep generative priors," IEEE Trans. Comput. Imag., vol. 6, pp. 1493–1506, Nov. 2020.

[37] X. Pan, X. Zhan, B. Dai, D. Lin, C. C. Loy, and P. Luo, "Exploiting deep generative priors for versatile image restoration and manipulation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 11, pp. 7474–7489, Nov. 2022.

[38] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte, "Plug-and-play image restoration with deep denoiser prior," IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 10, pp. 6360–6376, Oct. 2022.

[39] Z. Cheng, M. Gadelha, S. Maji, and D. Sheldon, "A Bayesian perspective on the deep image prior," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2019, pp. 5438–5446.

[40] K. Hua, A. Gilberta, H. Jib, and J. Collomosse, "Neural architecture search for deep image prior," Comput. Graph., vol. 98, pp. 188–196, 2021.

[41] J. A. Palmer, K. Kreutz-Delgado, and S. Makeig, "Strong sub-and super-Gaussianity," in Proc. Int. Conf. Latent Variable Anal. Signal Separation, 2010, pp. 303–310.

[42] C. M. Bishop, Pattern Recognition and Machine Learning. Berlin, Germany: Springer, 2006.

[43] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in Proc. 33rd Int. Conf. Neural Inf. Process. Syst., 2019, Art. no. 721.

[44] A. Vahdat and J. Kautz, "NVAE: A deep hierarchical variational autoencoder," in Proc. 54th Int. Conf. Neural Inf. Process. Syst., 2020, Art. no. 1650.

[45] J. A. Palmer, K. Kreutz-Delgado, and S. Makeig, "Strong sub-and super-Gaussianity," in Proc. Int. Conf. Latent Variable Anal. Signal Separation, 2010, pp. 303–310.

[46] C. M. Bishop, Pattern Recognition and Machine Learning. Berlin, Germany: Springer, 2006.

[47] A. Paszke et al., “PyTorch: An imperative style, high-performance deep learning library,” in Proc. 33rd Int. Conf. Neural Inf. Process. Syst., 2019, Art. no. 721.

[48] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, “Efficient marginal likelihood optimization in blind deconvolution,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2011, pp. 2657–2664.

[49] D. Perrone and P. Favaro, “Total variation blind deconvolution: The devil is in the details,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2014, pp. 2909–2916.

[50] F. Wen, R. Ying, Y. Liu, P. Liu, and T.-K. Truong, “A simple local minimal intensity prior and an improved algorithm for blind image deblurring,” IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 8, pp. 2923–2937, Aug. 2021.
[68] D. Huo, A. Masoumzadeh, and Y.-H. Yang, “Blindnon-uniform motion deblurring using atrous spatial pyramid deformable convolution and deblurring-reblurring consistency,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2022, pp. 437–446.

[69] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, “Restormer: Efficient transformer for high-resolution image restoration,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2022, pp. 5728–5739.

[70] L. Chen, X. Chu, X. Zhang, and J. Sun, “Simple baselines for image restoration;” in Proc. 17th Eur. Conf. Comput. Vis., Tel Aviv, Israel, 2022, pp. 17–33.

[71] S. Nah, S. Son, S. Lee, R. Timofte, and K. M. Lee, “NTIRE 2021 challenge on image deblurring,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2021, pp. 149–165.

[72] Y. Zhang, Y. Lau, H.-W. Kuo, S. Cheung, A. Pasupathy, and J. Wright, “On the global geometry of sphere-constrained sparse blind deconvolution,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2017, pp. 4381–4389.

[73] A. Mittal, R. Soundararajan, and A. C. Bovik, “Making a “completely blind” image quality analyzer,” IEEE Signal Process. Lett., vol. 20, no. 3, pp. 209–212, Mar. 2013.

[74] A. Mittal, A. K. Moorthy, and A. C. Bovik, “Blind/referenceless image spatial quality evaluator,” in Proc. Conf. Rec. 45th Asilomar Conf. Signals Syst. Comput., 2011, pp. 723–727.

[75] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani. “Blind image quality evaluation using perception based features,” in Proc. 21st Nat. Conf. Commun., 2015, pp. 1–6.

[76] J. W. Soh, S. Cho, and N. I. Cho, “Meta-transfer learning for zero-shot super-resolution,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2020, pp. 3513–3522.

Dong Huo received the BEng degree in software engineering from the Harbin Institute of Technology, Harbin, China, in 2018. He is currently working toward the PhD degree with the Department of Computing Science, University of Alberta, Canada. His research interests include computer vision, image processing, deep learning, and machine learning, with a focus on image restoration and image segmentation.

Abbas Masoumzadeh received the BSc degree in software engineering and worked as an information system manager for more than 5 years, and the MSc degree in computer science from York University. Currently, he is working toward the PhD degree in computing science with the University of Alberta. He was the recipient of the Vision Science to Applications (VISTA) Masters Scholarship. He also interned as a Machine Learning researcher with the Borealis AI company. He has received multiple awards during his PhD studies including the University of Alberta PhD Recruitment Scholarship, Alberta Graduate Excellence Scholarship (AGES), and Huawei Doctoral Scholarship. He has served as a reviewer for multiple conferences including ICLR and ICML. His broad research interests include computer vision, machine learning, reinforcement learning, and solutions for complex problems.

Rafsanjany Kushol received the BSc Engg and MSc Engg degrees in computer science and engineering (CSE) from the Islamic University of Technology (IUT), Bangladesh, in 2013 and 2018 respectively, and is currently working toward the PhD degree in Computing Science Department, University of Alberta, Canada. He worked as a lecturer with the Department of CSE, IUT, from 2014 to 2019. His research interests include computer vision, machine learning, deep learning, and medical image analysis.

Yee-Hong Yang (Senior Member, IEEE) received the BSc degree (Hons.) from The University of Hong Kong, the MSc degree from Simon Fraser University, and the PhD degree from the University of Pittsburgh. He was a faculty member of the Department of Computer Science, University of Saskatchewan from 1983 to 2001, and served as the graduate chair from 1999 to 2001. While there, in addition to department level committees, he also served on many college and university level committees. Since July 2001, he has been a professor with the Department of Computing Science, University of Alberta, where he served as the associate chair (Graduate Studies) from 2003 to 2005. His research interests cover a wide range of topics from computer graphics to computer vision, which include physically based animation of Newtonian and non-Newtonian fluids, texture analysis and synthesis, human body motion analysis and synthesis, computational photography, stereo and multiple view computer vision, and underwater imaging. He has published more than 160 papers in international journals and conference proceedings in the areas of computer vision and graphics. He serves on the editorial board of the Journal Pattern Recognition. In addition to serving as a reviewer to numerous international journals, conferences, and funding agencies, he has served on the program committees of many national and international conferences. In 2007, he was invited to serve on the expert review panel to evaluate computer science research in Finland.