

## Research Article

# Infrared Image Deblurring Based on Generative Adversarial Networks

Yuqing Zhao , Guangyuan Fu, Hongqiao Wang, Shaolei Zhang, and Min Yue

*Xi'an Research Institute of High-Tech, Shaanxi 710025, China*

Correspondence should be addressed to Yuqing Zhao; [zoe\\_rabbi@126.com](mailto:zoe_rabbi@126.com)

Received 7 March 2021; Accepted 26 April 2021; Published 4 May 2021

Academic Editor: Muhammad Tariq Mahmood

Copyright © 2021 Yuqing Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Blind deblurring of a single infrared image is a challenging computer vision problem. Because the blur is not only caused by the motion of different objects but also by the relative motion and jitter of cameras, there is a change of scene depth. In this work, a method based on the GAN and channel prior discrimination is proposed for infrared image deblurring. Different from the previous work, we combine the traditional blind deblurring method and the blind deblurring method based on the learning method, and uniform and nonuniform blurred images are considered, respectively. By training the proposed model on different datasets, it is proved that the proposed method achieves competitive performance in terms of deblurring quality (objective and subjective).

## 1. Introduction

The main reason of motion blur is that there is rapid relative motion between the camera and the captured object during the exposure time. The blurring of images will reduce the perceptual quality of human beings. It also has a negative impact on advanced visual tasks such as object detection and semantic understanding. Image deblurring is a common and important problem in the field of image processing and computer vision. However, due to the complexity of motion blur processing, most existing methods may not produce satisfactory results when the blur kernel is complex, and the details of the required clear image are abundant. In addition, because the infrared (IR) imaging system is more complex than the natural imaging system, the degradation degree of infrared images is relatively high, such as Gaussian blur, motion blur, and noise pollution. Therefore, infrared image deblurring plays an important role in the IR imaging system. Some researchers are dedicated to hardware-based research for infrared image deblurring. In literature [1], the fluttered shutter is used to solve the problem of infrared image deblurring. Literature [2] uses an ordinary inertial measurement unit (IMU) to estimate the trajectory of the camera

movement during the exposure time. Oswald-Tranta et al. [3] used the parameterized Wiener filter method to blur the infrared images obtained from the infrared detector of the microbolometer. Oswald-Tranta also committed to obtaining accurate temperature measurements by deblurring infrared images [4]. Wang et al. [5] used the iterative Wiener filter to estimate the PSF filter of motion blur in infrared images. The deblurring method based on infrared imaging hardware equipment is more expensive. Therefore, the algorithm-based deblurring of infrared images is more widely used. Luo et al. [6] developed a new infrared blurred image restoration model based on the principle of non-uniform exposure. In order to eliminate the motion blur of the image and restore the image, Jing et al. [7] proposed an infrared target motion deblurring method based on the Haar wavelet transform. Liua et al. [8] proposed a method of using Lp-quasi-linear norm and the overlapping sparse total variation method to blur infrared images.

Inspired by the great progress of traditional blind deblurring methods and learning-based blind deblurring methods recently, we propose a method based on GAN and channel prior discrimination. Specifically, the innovation of this article is summarized as follows:

- (i) A channel-based inverse prior discrimination is proposed. And this method is built into a new framework of the GAN. It improves the blind deblurring performance of infrared images.
- (ii) Different blur types are caused by the motion of the camera or object. In view of this situation, two different methods were used to synthesize two kinds of blurred datasets.
- (iii) In the experimental stage, we conducted extensive experiments which were carried out on two different datasets. The method proposed in this article is compared with the other four advanced methods qualitatively and quantitatively.

## 2. Related Work

**2.1. Image Deblurring.** The solutions to deblurring problems are mainly divided into two types: blind deblurring and nonblind deblurring. The early related work is mainly nonblind deblurring, that is, the ambiguity function is assumed to be known. Most of these algorithms rely on the Lucy–Richardson algorithm and Wiener or Tikhonov filter which are sensitive to noise to perform deconvolution operation and obtain IS estimation. However, in reality, ambiguity functions are often uncertain. It is unrealistic to find the ambiguity function for each pixel. Therefore, a lot of recent works are focused on blind deblurring. The first modern bold attempt was Fergus et al.’s [9] variational Bayesian method to eliminate uniform camera shake. In the past decade, many methods [10–20] have solved the blur caused by camera shake by considering the uniform blur on the image. This kind of algorithm first estimates camera motion according to the induced blur kernel and then reverses the effect by performing deconvolution operation. Unfortunately, these algorithms are usually unable to eliminate nonuniform motion blur.

In fact, due to the camera rotation, radial camera motion, depth of field change, or rapid movement of objects, images taken in the field may experience more complex heterogeneous blur. Therefore, most existing nonuniform blind deblurring methods [21–26] are based on specific motion models. For example, Gupta et al. [27] proposed to model camera motion as a motion density function. The blurring kernel of spatial variables can be derived directly from it. By specifying a prior of sparsity and compactness in density, an optimization problem is formulated, and the density function and deblurred image can be solved iteratively. A new projection motion path model is proposed in [28, 29]. Another method to eliminate spatial variation ambiguity is to estimate through block-by-block blurring kernel [30–32]. Segmented blurring estimation [24, 33] also considers the spatial variation blur caused by the object movement.

In recent years, some methods based on the convolutional neural network (CNN) have appeared [23, 34–42]. Schuler et al. [39] made the first heuristic attempt, focusing on unified blind deblurring, including modules for feature extraction, blurring kernel estimation, and clear image estimation. Sun et al. [40] used the CNN to estimate the

blurring kernel. Chakrabarti [43] put forward another advanced method. This method learnt to predict the plural Fourier coefficients of the deconvolution filter of the input patches of blurred images and then used the traditional optimization strategy to estimate the global blurring kernel from the restored patches. And Gong et al. [34] used the fully convolved network movement flow to estimate. All these methods use the CNN to estimate unknown ambiguity functions. Recently, Noroozi et al. [23] and Nah et al. [44] adopted the kernel-free end-to-end method, using the multiscale CNN to directly remove images. Tao et al.’s latest work [42] expands the multiscale CNN from [37] to scale the recursive CNN to realize image deblurring, and the effect is impressive. Ramakrishnan et al. [38] used a combination of pix2pix framework [45] and densely connected convolution network [46] to perform blind kernel-free image deblurring. These methods can deal with different sources of blur. Since Ramakrishnan et al., the success of GAN in image restoration has also affected the deblurring of single image. Ramakrishnan et al. [38] firstly solved the problem of image deblurring by referring to the idea of image translation [45]. Recently, Kupyn et al. [36] introduced DeblurGAN; it is developed by Wasserstein GAN [47] with gradient penalty and perceived loss.

**2.2. GAN.** Generative adversarial network, commonly known as GAN, was proposed by Goodfellow [48] and inspired by the zero-sum game in game theory. This game has achieved many exciting results in image restoration [49]. After style conversion [45, 50, 51], it can even be used in other fields. The system includes a generator  $G$  and a discriminator  $D$ ; they constitute a minmax game for two. The generator tries to capture the potential actual data distribution and outputs new data samples, while the discriminator tries to distinguish whether the input data come from the real data distribution. The minmax game with the value function  $V(G, D)$  is represented by the following formula [1]. Both generator and discriminator can be constructed based on the CNN and trained based on the above ideas.

$$\min_G \max_D V(G, D) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_x(z)} [\log(1 - D(G(z)))], \quad (1)$$

where  $p_{\text{data}}(x)$  is the real data distribution,  $p_{(x)}(z)$  is the model distribution, and the input  $z$  is a sample from a simple noise distribution.

GAN is known for its ability to preserve textural details in images, create solutions that are close to the real image, and be perceptually persuasive. Literature [51] was further developed; it is based on the conditional GAN [52] and trains a cyclic consistency goal. This target generates a more realistic image in the task of image migration. Inspired by this idea, Isola [45] put forward the earliest idea of image deblurring based on the GAN. Recently, great progress has been made in the related fields of image super-resolution [53] and image restoration [54] by applying the GAN.

**2.3. Dark Channel Prior Algorithm.** He et al. [55] proposed a defogging algorithm (DCP) based on the dark channel prior. DCP is based on the assumption that most nonsky patches of outdoor fog-free images contain some pixels. These pixels have very low intensity in at least one color channel. For any image  $I$ , its dark channel  $I_{\text{dark}}(x)$  is given by the following formula:

$$I_{\text{dark}}(x) = \min_{x \in \Omega(x)} \left( \min_{c \in \{r, g, b\}} I^c(x) \right), \quad (2)$$

in which  $\Omega(x)$  represents a local color block centered on  $x$  and  $I^c$  is the  $c$ -th color channel of  $i$ . The optical channel proposed in a similar article [56] is based on the assumption that the most blurred image block contains some pixels with very bright intensity in at least one color channel. For any image  $I$ , its optical channel  $I_{\text{bright}}(x)$  is as follows:

$$I_{\text{bright}}(x) = \max_{x \in \Omega(x)} \left( \max_{c \in \{r, g, b\}} I^c(x) \right). \quad (3)$$

Many methods use dark channels and bright channels to complete image defogging [55, 56], and they are also used to estimate the blurring kernel in conventional blind image deblurring [15, 57]. In [15], Pan et al. proposed to use the regularization term based on  $L_0$  additionally on the dark channel image to improve the gradient-based  $L_0$ -minimization blind deblurring method [11]. In [57], Yan et al. further combined and used  $L_0$ -based regularization in both dark and bright channel images.

### 3. Method

In this work, the purpose of the infrared image deblurring model is to restore a clear image when only the blurred infrared image is given. In this paper, the architecture, proposed in [51], is used to build two sets of GAN models. The generators are  $G_{B2S}$ :  $I_B \rightarrow I_S$  and  $G_{S2B}$ :  $I_S \rightarrow I_B$ .  $G_{B2S}$  restores clear images from blurred images, while  $G_{S2B}$  generates blurred images from clear images. The discriminators are  $D_B$  and  $D_S$ .  $D_B$  tries to distinguish whether the input is a blurred image, while  $D_S$  tries to distinguish whether the input is sharp. The architecture of the proposed method is shown in Figure 1. The input in the method is the blurred image and clear image. The clear image is sent to the generator  $G_{S2B}$  to generate the corresponding blurred image. The generated blurred image is sent to the generator  $G_{B2S}$  to generate a deblurred image. The generated deblurred image and the real clear image are sent to the discriminator  $D_S$  together to identify true and fake. The real blurred image is input into the generator  $G_{B2S}$  to generate a blurred image. The generated deblurred image is sent to the generator  $G_{S2B}$  to synthesize the blurred image. The synthesized blurred image and the real blurred image are sent to the discriminator  $D_B$  to determine the authenticity. Through continuous iteration, the generator can generate more realistic deblurred images. The algorithm flow is summarized as Algorithm 1.

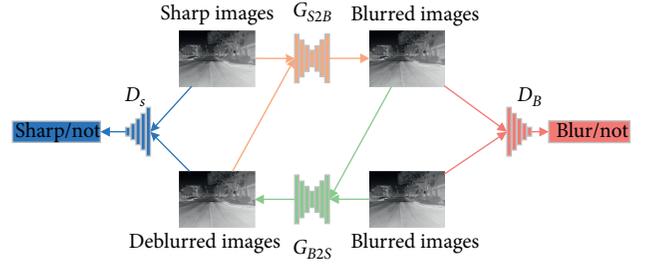


FIGURE 1: The overall structure of the proposed method.

**3.1. Model Architecture.** The method proposed by us includes two pairs of GAN. The model architecture of one pair is shown in Figure 2; it includes two deep convolutional neural network (DCNN) modules. The generator is similar to that proposed by Johnson et al. [50], including two step convolution blocks with a step size of 0.5, nine residual blocks, and two transposed convolution blocks. An instantiation standardization layer (IN) is added after the convolution layer of each convolution module except the ResBlocks. The network structure of the discriminator is the same as that of [45]. It includes five convolution modules; except the last module, each convolution layer is followed by an IN layer and a LeakyReLU layer.

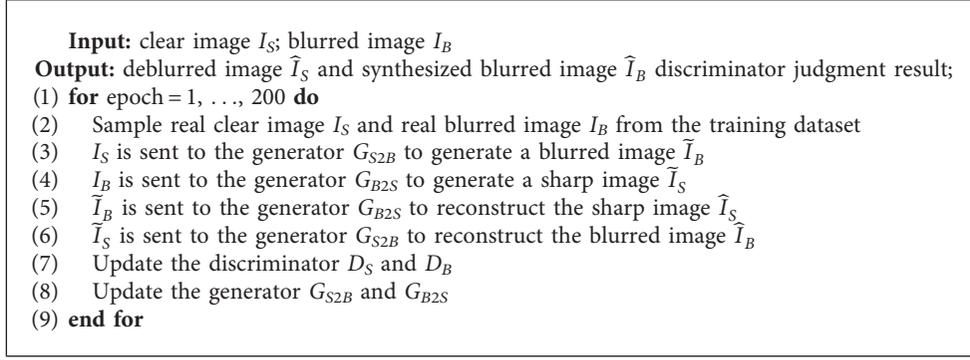
As we all know, both BN and IN layers use a batch of mean and variance to normalize features during training and use the estimated mean and variance of the whole training dataset during testing. One of the potential motivations for applying BN or IN is to accelerate the training of deep neural networks (DNNs). However, recent work [58] on single-image super-resolution points out that the BN layer will bring artifacts in training and testing stages. Especially, these artifacts are more likely to occur with the deepening of the network and training under the framework of the GAN. When turned to blind deblurring, the above empirical discussion shows that the IN layer will bring similar artifacts, that is, irregular block color shift. Therefore, no IN or BN layer is introduced in the residual block, as shown in Figure 3. The network configuration of the generator and discriminator is shown in Tables 1 and 2.

### 3.2. Loss Function

**3.2.1. Adversarial Loss.** Adversarial loss includes generator adversarial loss and discriminator adversarial loss, where generator adversarial loss is defined as follows:

$$\mathcal{L}_{G_{\text{adv}}} = \sum_{n=1}^N -\log[D_B(\hat{I}_B)] + \sum_{n=1}^N -\log[D_S(\hat{I}_S)]. \quad (4)$$

Among them, the first item is the adversarial loss between the reconstructed blurred image  $\hat{I}_B$  and the discriminator  $D_B$ . The second term is the adversarial loss between the reconstructed sharp image  $\hat{I}_S$  and the discriminator  $D_S$ . The least square loss is better than the mean square loss in the image style conversion task. Therefore, the discriminator uses the least square loss as adversarial loss:



ALGORITHM 1: Deblurring method proposed in this work.

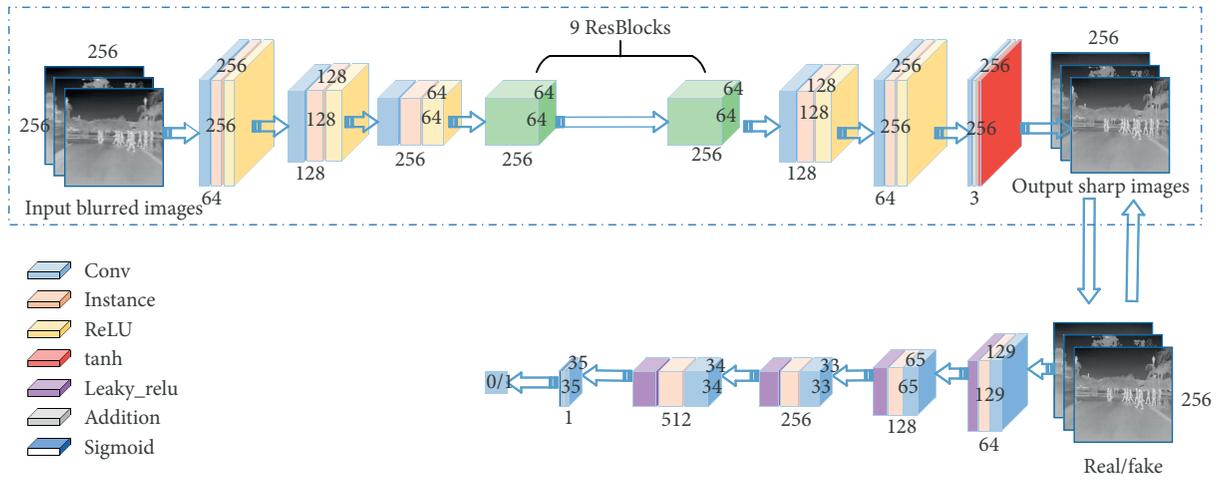


FIGURE 2: The model architecture of the method is proposed. Top: generator model. Bottom: discriminator model.

$$\mathcal{L}_{D_{adv}} = \frac{1}{2} \left\{ \mathbb{E}_{I_B \sim p_{data}(I_B)} [D_B(I_B) - 1]^2 + \mathbb{E}_{\tilde{I}_S \sim p_z(\tilde{I}_S)} [D_B(G_{S2B}(\tilde{I}_S))]^2 \right\} + \frac{1}{2} \left\{ \mathbb{E}_{I_S \sim p_{data}(I_S)} [D_S(I_S) - 1]^2 + \mathbb{E}_{\tilde{I}_B \sim p_z(\tilde{I}_B)} [D_S(G_{B2S}(\tilde{I}_B))]^2 \right\}. \quad (5)$$

Among them, the first term is the loss function of the discriminator  $D_B$  error identification, and the second term is the loss function of the discriminator  $D_S$  error identification.

**3.2.2. Loss of Circular Perception Consistency.** For the general GAN, it is necessary to compare the reconstructed image and the original image in the training stage with a certain metric as content loss. The common choice of content loss is pixel-space loss, and the simplest is L1 or L2 loss. Because this kind of loss often produces excessively smooth pixel-space output, this leads to blurring artifacts on the generated image. This brings negative factors to the deblurring task, so the circular perception consistency loss suggested in [58] is adopted. The purpose of circular

perception consistency loss is to preserve the original image structure by looking at the combination of high-level and low-level features extracted from the second and fifth pooling layers of the VGG-16 system [59]. Under the constraints of generator  $G_{B2S}: I_B \rightarrow I_S$  and generator  $G_{S2B}: I_S \rightarrow I_B$ , the following formula of circular perception consistency loss is given:

$$\mathcal{L}_{cycle\_perceptual} = \mathcal{L}_{cycle\_perceptual1} + \mathcal{L}_{cycle\_perceptual2}, \quad (6)$$

$$\mathcal{L}_{cycle\_perceptual1} = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I_S)_{x,y} - \phi_{i,j}(G_{B2S}(I_B))_{x,y})^2, \quad (7)$$

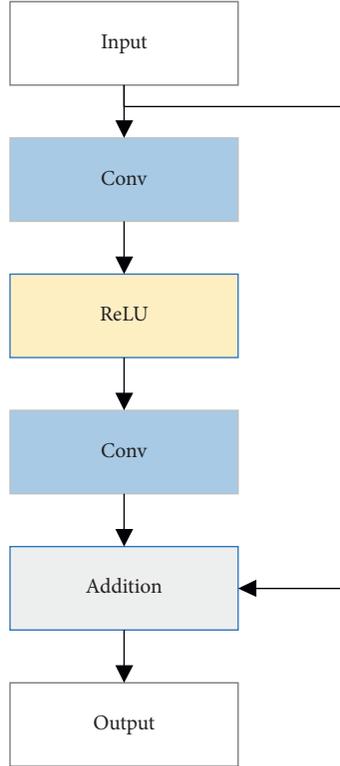


FIGURE 3: The ResBlocks with no IN and BN layers are not introduced.

TABLE 1: The layer structure and parameters of the generator.

Layer (type)	Output shape	Parameters
ReflectionPad2d-1	[-1, 3, 262, 262]	0
Conv2d-2	[-1, 64, 256, 256]	1,792
InstanceNorm2d-3	[-1, 64, 256, 256]	0
ReLU-4	[-1, 64, 256, 256]	0
Conv2d-5	[-1, 128, 128, 128]	73,856
InstanceNorm2d-6	[-1, 128, 128, 128]	0
ReLU-7	[-1, 128, 128, 128]	0
Conv2d-8	[-1, 256, 64, 64]	295,168
InstanceNorm2d-9	[-1, 256, 64, 64]	0
ReLU-10	[-1, 256, 64, 64]	0
ReflectionPad2d-11	[-1, 256, 66, 66]	0
Conv2d-12	[-1, 256, 64, 64]	590,080
ReLU-13	[-1, 256, 64, 64]	0
ReflectionPad2d-14	[-1, 256, 66, 66]	0
Conv2d-15	[-1, 256, 64, 64]	590,080
ResidualBlock-16	[-1, 256, 64, 64]	0
ConvTranspose2d-65	[-1, 128, 128, 128]	295,040
InstanceNorm2d-66	[-1, 128, 128, 128]	0
ReLU-67	[-1, 128, 128, 128]	0
ConvTranspose2d-68	[-1, 64, 256, 256]	73,792
InstanceNorm2d-69	[-1, 64, 256, 256]	0
ReLU-70	[-1, 64, 256, 256]	0
ReflectionPad2d-71	[-1, 64, 262, 262]	0
Conv2d-72	[-1, 3, 256, 256]	1,731
Tanh-73	[-1, 3, 256, 256]	0

TABLE 2: The layer structure and parameters of the discriminator.

Layer (type)	Output shape	Parameters
Conv2d-1	[-1, 64, 128, 128]	1,792
LeakyReLU-2	[-1, 64, 128, 128]	0
Conv2d-3	[-1, 128, 64, 64]	73,856
InstanceNorm2d-4	[-1, 128, 64, 64]	0
LeakyReLU-5	[-1, 128, 64, 64]	0
Conv2d-6	[-1, 256, 32, 32]	295,168
InstanceNorm2d-7	[-1, 256, 32, 32]	0
LeakyReLU-8	[-1, 256, 32, 32]	0
Conv2d-9	[-1, 512, 31, 31]	1,180,160
InstanceNorm2d-10	[-1, 512, 31, 31]	0
LeakyReLU-11	[-1, 512, 31, 31]	0
Conv2d-12	[-1, 1, 30, 30]	4,609

$$\mathcal{L}_{\text{cycle\_perceptual2}} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I_B)_{x,y} - \phi_{i,j}(G_{S2B}(I_S))_{x,y})^2. \quad (8)$$

Among them,  $\mathcal{L}_{\text{cycle\_perceptual1}}$  is the cycle perception consistency loss of the generator  $G_{B2S}$ ;  $\mathcal{L}_{\text{cycle\_perceptual2}}$  is the cycle perception consistency loss of the generator  $G_{S2B}$ . The goal is to make the reconstructed image and the input image as close as possible.  $\phi_{i,j}$  is the feature map obtained by the VGG-16 network from the  $i$ -th largest pooling layer after the  $j$ -th convolutional layer.  $W_{i,j}$  and  $H_{i,j}$  are the corresponding dimensional feature maps.

**3.2.3. Prior Loss Based on the Dark Channel and Bright Channel.** Using the bright channel and dark channel presented in formulas (2) and (3), the following two different energies are defined:

$$\text{Energy}_{\text{dark}}(I) = \left( \frac{\sum_x I_{\text{dark}}^2(x)}{M * N} \right)^{1/2}, \quad (9)$$

$$\text{Energy}_{\text{bright}}(I) = \left( \frac{\sum_x I_{\text{bright}}^2(x)}{M * N} \right)^{1/2}, \quad (10)$$

in which  $M$  and  $N$  are channel sizes.  $I_{\text{dark}}(x)$  is defined by formula (2).  $I_{\text{bright}}(x)$  is defined by formula (3). It is verified by He et al. and Xu et al. [55, 56] that clear images have lower dark energy and higher bright energy. In order to test the resolvability of (9) and (10) between the infrared clear image  $I_S$  and the corresponding blurred image  $I_B$ , the images of the FLIR\_ADAS\_1\_3 dataset are calculated. The results of 8862 clear and blurring image pairs show that  $\text{Energy}_{\text{dark}}(I_S) < \text{Energy}_{\text{dark}}(I_B)$  and  $\text{Energy}_{\text{bright}}(I_S) > \text{Energy}_{\text{bright}}(I_B)$ . In order to visualize the calculation results, 200 images were randomly selected, and the sum curves were provided, as shown in Figure 4.

Based on this conclusion, it is considered that clear images and blurred images can be distinguished by dark energy and bright energy defined in (9) and (10). In order to improve the GAN from the perspective of domain

knowledge, the prior judgment of the traditional blind image deblurring method is taken as the training loss function:

$$\mathcal{L}_{DCP}(G_{B2S}(I_B)) = \text{Energy}_{\text{dark}}(G_{B2S}(I_B)), \quad (11)$$

$$\mathcal{L}_{BCP}(G_{S2B}(I_S)) = \text{Energy}_{\text{bright}}(G_{S2B}(I_S)). \quad (12)$$

Combining formulas (4)–(12), the final losses adopted in this article are as follows:

$$\mathcal{L}_G = \lambda_1 \mathcal{L}_{G_{\text{adv}}} + \lambda_2 \mathcal{L}_{\text{cycle\_perceptual}} + \lambda_3 (\mathcal{L}_{DCP} + \mathcal{L}_{BCP}), \quad (13)$$

$$\mathcal{L}_D = \mathcal{L}_{D_{\text{adv}}}. \quad (14)$$

In formula (13),  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  are the weights of the loss function. According to the experimental results, they are  $\lambda_1 = 0$ ,  $\lambda_2 = 10$ , and  $\lambda_3 = 10^3$ , respectively.

## 4. Experiment

All models are implemented by the PyTorch deep learning framework. FLIR\_ADAS\_1\_3 dataset and LTIR dataset are used to train on a desktop with 2.20 GHz  $\times$  40 Intel Xeon (r) Silver4114 CPU, GeForce GTX 1080Ti, and 64GiB memory. In this section, the experimental results are introduced and compared with the results of mainstream methods. In addition, qualitative results are provided on real images.

**4.1. Synthetic Blurring Dataset.** There are two types of blurred images: the overall image is blurred due to the movement of the imaging device, and the partial image is blurred due to the movement of the imaging object. In order to verify that our deblurring method is effective for both types of blur, we simulate the two types of image blur through two different schemes.

For the overall image blur caused by the motion of the imaging device, we choose to use a linear blur kernel to create a synthetic blur image. Sun et al. [40] created a composite blurred image by convolving a clear natural image with one of 73 possible linear motion kernels. Xu et al. [60] also used the linear motion kernel to create synthetic blurred images. Chakrabarti [61] created a blurring kernel by

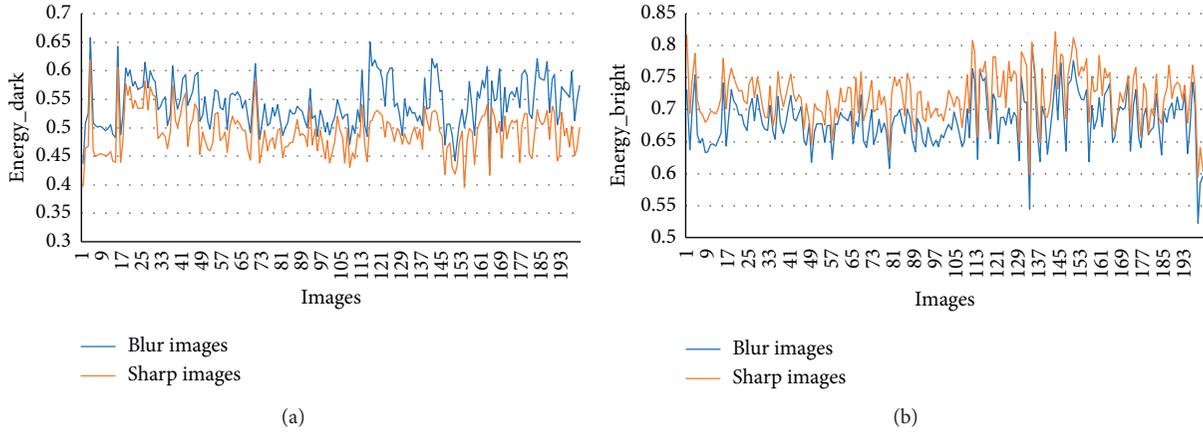


FIGURE 4: Graph of Energy<sub>dark</sub> and Energy<sub>bright</sub> on the test image randomly selected from the FLIR dataset: (a) dark energy and (b) bright energy.

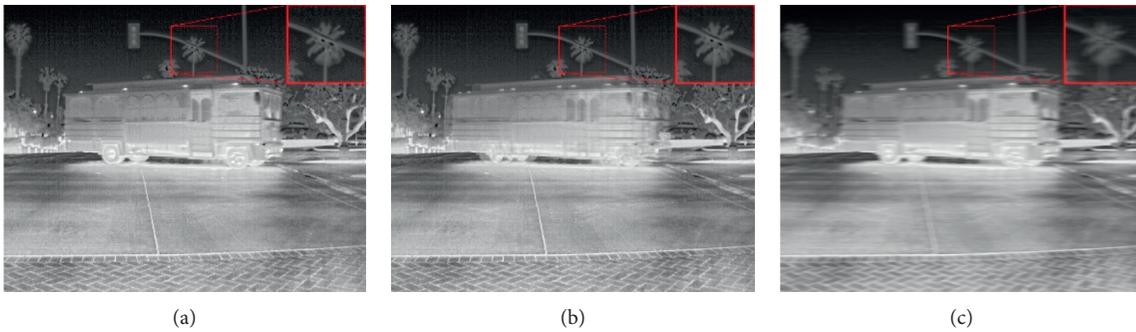


FIGURE 5: Comparison of synthetic blurred images by different methods. (a) Real clear image. (b) Blurred image synthesized by the average frame method. (c) Blurred image synthesized by the blur kernel method.

sampling six random points and fitting splines to them. Levin et al. provided eight blurring kernels [62] that have been used for multiple datasets. However, the maximum blurring kernel size of these eight blurring kernels is  $41 \times 41$ , which is relatively small in practice. Therefore, we follow the algorithm in [63] to generate four uniform blur kernels from  $51 \times 51$  to  $101 \times 101$  by sampling random 6D camera trajectories. Then, a convolution model with 1% Gaussian noise is used to synthesize a blurred image.

For the local image blur caused by the motion of the imaging object, we choose to use the average frame of the video sequence to simulate. This is a typical method of simulating blurred image pairs [23, 37]. This method can create realistic blurred images but only limits the image space to scenes with video sequences; this makes the dataset limited. Figure 5 shows a comparison of two different blur types. The blurred image generated by averaging frames shows the blur caused by moving objects and static background. The car in Figure 5(b) is blurred, but the surrounding trees are clear. The blur kernel method simulates the motion blur of the whole image caused by the motion of the camera. In Figure 5(c), the car and the surrounding trees are blurred. In order to verify the universality of our

algorithm, we use the blur kernel to synthesize blurred images for the LTIR dataset and use two synthetic methods of average frame and the blur kernel for the FLIR dataset to simulate motion blur. The blurred dataset synthesized by the blur kernel method is used as the FLIR-A dataset; the blurred dataset synthesized by the average frame method is used as the FLIR-B dataset.

**4.2. FLIR\_ADAS\_1\_3 Dataset Results.** FLIR\_ADAS\_1\_3 datasets provide annotated thermal imaging datasets and corresponding unannotated RGB images for training and verifying neural networks. Data are acquired by using the RGB camera and thermal imaging camera installed on the vehicle. The dataset contains a total of 14,452 infrared images, of which 10,228 are from multiple short videos, and 4224 are from a video with a length of 144 s. All videos come from streets and highways. The sampling rate of most pictures is two frames per second. The frame rate of the video is 30 frames per second. When there are few targets in a few environments, the sampling rate is 1 frame per second. In the experiment, 8862 8-bit infrared images are divided into 7090 image training sets and 1772 image test sets. Figure 6

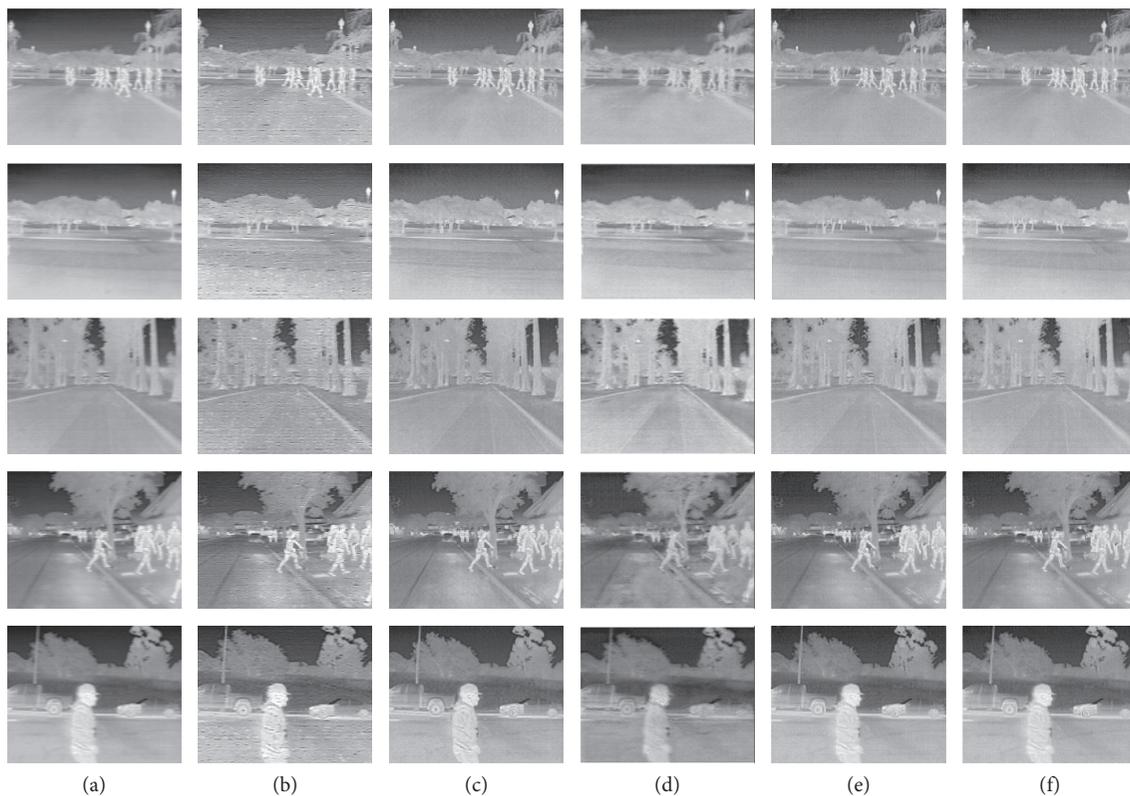


FIGURE 6: Deblurring results of the FLIR-A dataset. The columns from left to right are the original blurred image, results of Schuler et al. [39], Kupyn et al. [36], Zhu et al. [51], Engin et al. [64], and the results of the method proposed in this article. (a) Blurred image. (b) DeepDeblur. (c) DeblurGAN. (d) CycleGAN. (e) Cycle-Dehaze. (f) Ours.

TABLE 3: Comparison of quantitative deblurring performance on FLIR-A datasets.

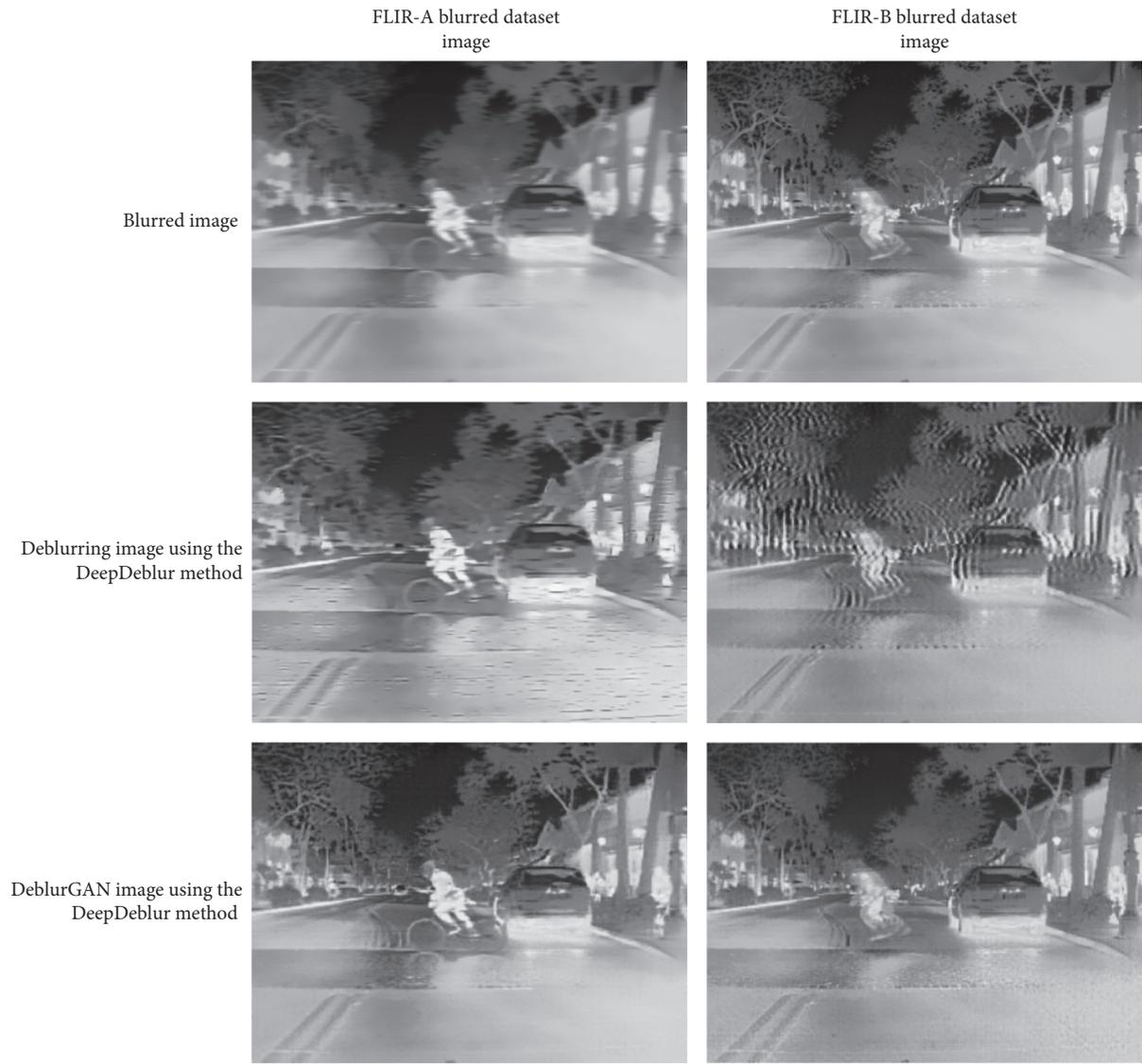
	DeepDeblur	DeblurGAN	CycleGAN	Cycle-Dehaze	Ours
SSIM	0.8916	0.9899	0.9190	0.9788	0.9985
PSNR (dB)	17.48	26.91	20.45	21.03	28.79
Time (s)	40.03	1.05	4.59	7.01	0.14

shows the test images on the FLIR-A blurred dataset, and the quantitative results are shown in Table 3.

In order to further compare the deblurring effects of various methods on different types of blurred images, we compare the deblurring results of FLIR-A and FLIR-B blurred datasets. Figure 7 shows the deblurred images of different methods on the two types of blurred datasets, and the evaluation indicators are shown in Table 4. It can be seen from the subjective and objective results that our method has better deblurring performance than several other methods. This result is particularly obvious on the FLIR-B blurred dataset. For partially blurred images caused by the motion of the imaging object, the deblurring effect of other methods is significantly reduced, the original clear background becomes more blurred, and the blurred area does not achieve the ideal deblurring effect. However, our method can restore the blurred area clearly while keeping the background clear. This

has a lot to do with the idea of channel prior discrimination adopted in our method. The channel prior discrimination algorithm is based on local color patches. This makes our method have better deblurring performance in the local blurred image.

**4.3. LTIR\_v1\_0 Dataset Results.** LTIR dataset is a thermal infrared dataset used to evaluate the tracking of a single object (STSO) in a short time. Currently, only one version is available. Version 1.0 consists of 20 infrared thermal sequences with an average length of 563 frames. This dataset is a subchallenge of the 2015 Visual Object Recognition (VOT) Challenge. In the experiment, 11,262 8-bit images are divided into a training set of 9010 images and a test set of 2252 images. Figure 8 shows the test image on the LTIR dataset. The quantitative results are shown in Table 5.



(a)

FIGURE 7: Continued.

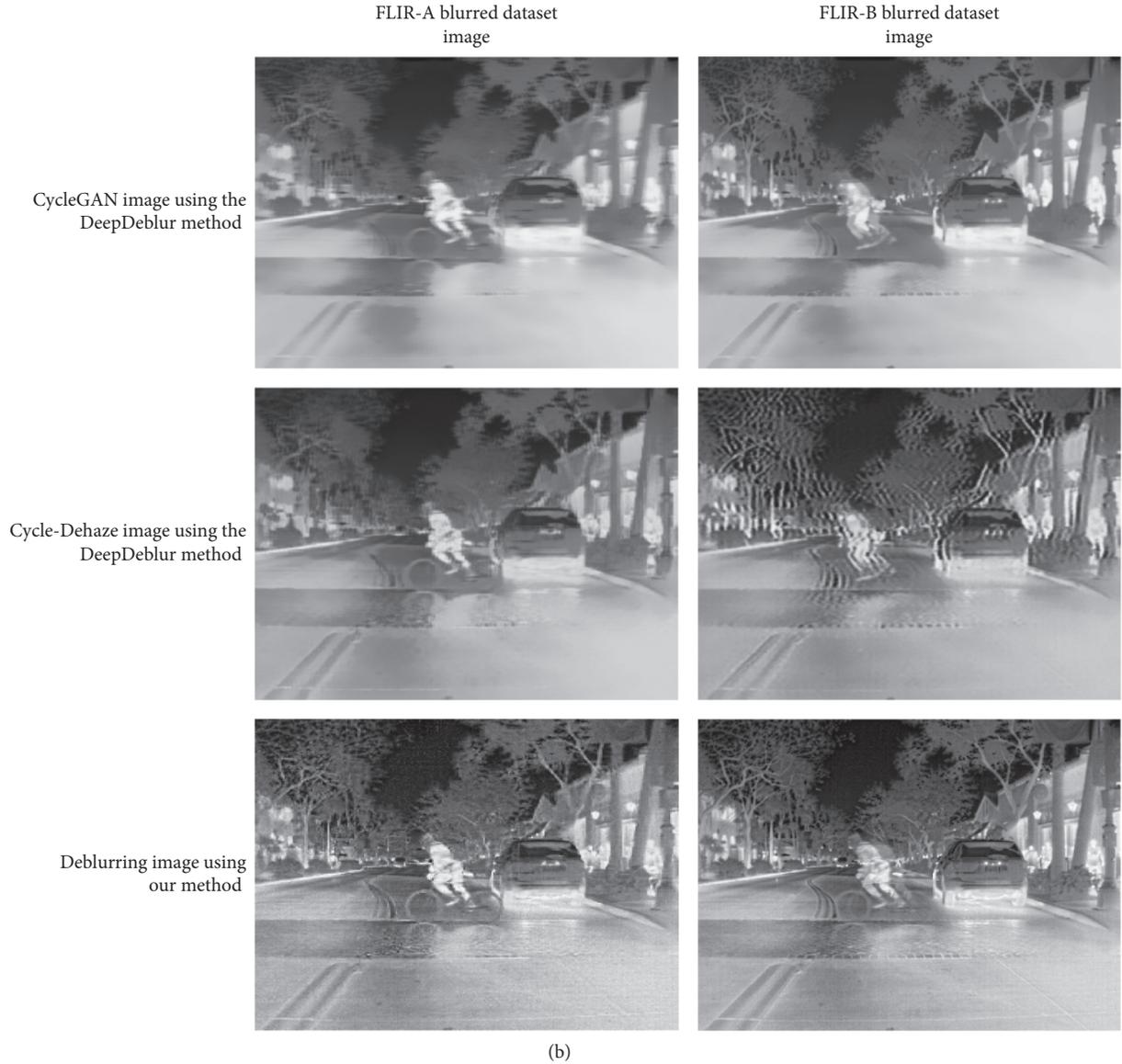


FIGURE 7: Comparison of deblurring results between FLIR-A and FLIR-B blurred datasets.

TABLE 4: Comparison of the deblurring performance of various methods on different types of blurred images.

		DeepDeblur	DeblurGAN	CycleGAN	Cycle-Dehaze	Ours
FLIR-A	SSIM	0.8916	0.9899	0.9190	0.9788	0.9985
	PSNR	17.48	26.91	20.45	21.03	28.79
FLIR-B	SSIM	0.7458	0.8161	0.7997	0.8364	0.9589
	PSNR	16.76	18.47	17.20	19.51	21.22

**4.4. Ablation Research and Analysis.** We conduct ablation research on the effect of the loss function component in the deblurring method proposed in this paper. The results are

summarized in Table 6. We can see that our proposed dark channel and bright channel a priori determination components are steadily improving PSNR and SSIM. In

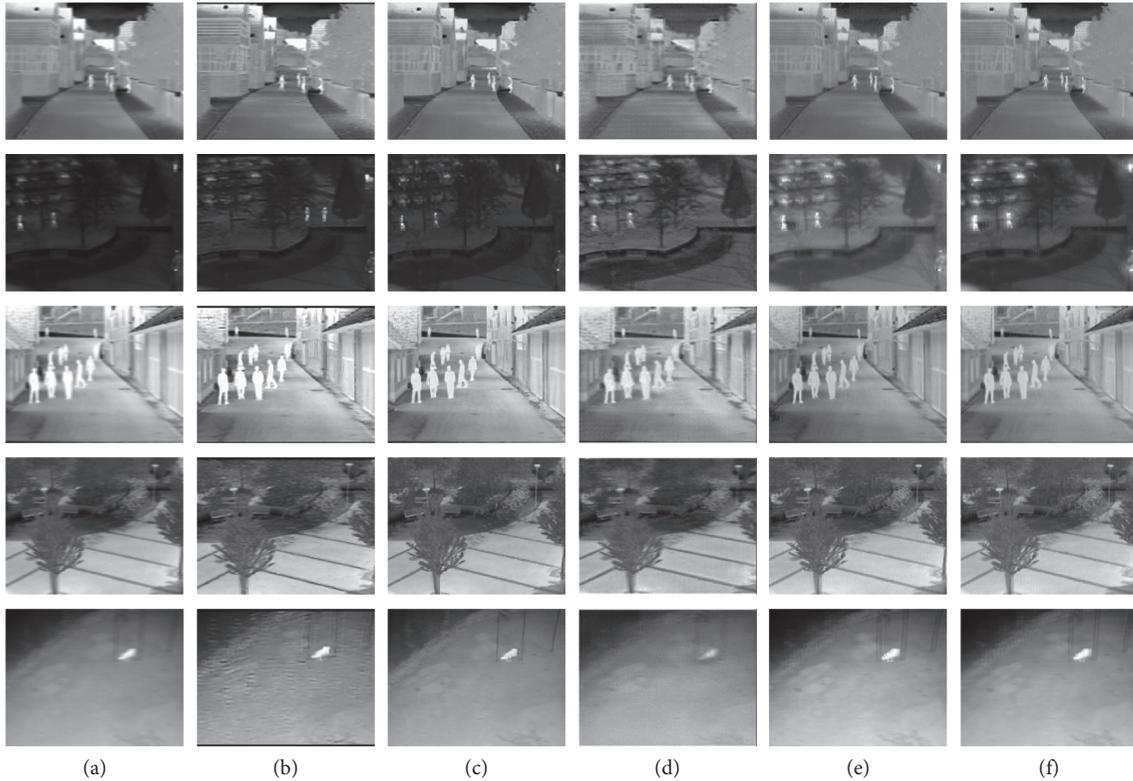


FIGURE 8: Deblurring results of the LTIR dataset. The columns from left to right are the original blurred image, the results of Schuler et al. [39], Kupyn et al. [36], Zhu et al. [51], Engin et al. [64], and the results of the proposed method in this article. (a) Blurred image. (b) DeepDeblur. (c) DeblurGAN. (d) CycleGAN. (e) Cycle-Dehaze. (f) Ours.

TABLE 5: Comparison of quantitative deblurring performance on LTIR datasets.

	DeepDeblur	DeblurGAN	CycleGAN	Cycle-Dehaze	Ours
SSIM	0.7535	0.8576	0.6977	0.7110	0.9697
PSNR (dB)	15.85	22.48	17.51	10.55	25.85
Time (s)	37.62	0.82	3.56	5.74	0.06

TABLE 6: Ablation study of the channel prior loss function.

	SSIM		PSNR (dB)	
	FLIR dataset	LTIR dataset	FLIR dataset	LTIR dataset
Remove the dark channel prior loss function	0.9805	0.7463	21.66	13.95
Remove the bright channel prior loss function	0.9823	0.8818	22.47	22.73
Replace perceptual loss with L1 loss	0.9344	0.9191	19.43	20.20
Replace perceptual loss with L2 loss	0.9421	0.9256	19.25	20.64
Ours	0.9985	0.9697	28.79	25.85

particular, the dark channel a priori determination module contributes the most. When we replace the perceptual loss function with L1 and L2 loss functions, the average SSIM and PSNR both decrease. It can be seen from Figure 9 that the deblurred image generated after replacing the perceptual loss function with the L1 and L2 loss function is too smooth. In summary, in the deblurring task, the perceptual loss function is more suitable than the L1 and L2 loss functions.

*4.5. Use Advanced Vision Tasks to Compare Deblurring Results.* Basic vision tasks, including image deblurring, serve for advanced vision tasks. In order to further verify the effectiveness of our method, we match the deblurred images generated by several methods with real clear images. Scale-Invariant Feature Transformation (SIFT) is a representation of Gaussian image gradient statistics in the field of feature points and is a commonly used image local feature

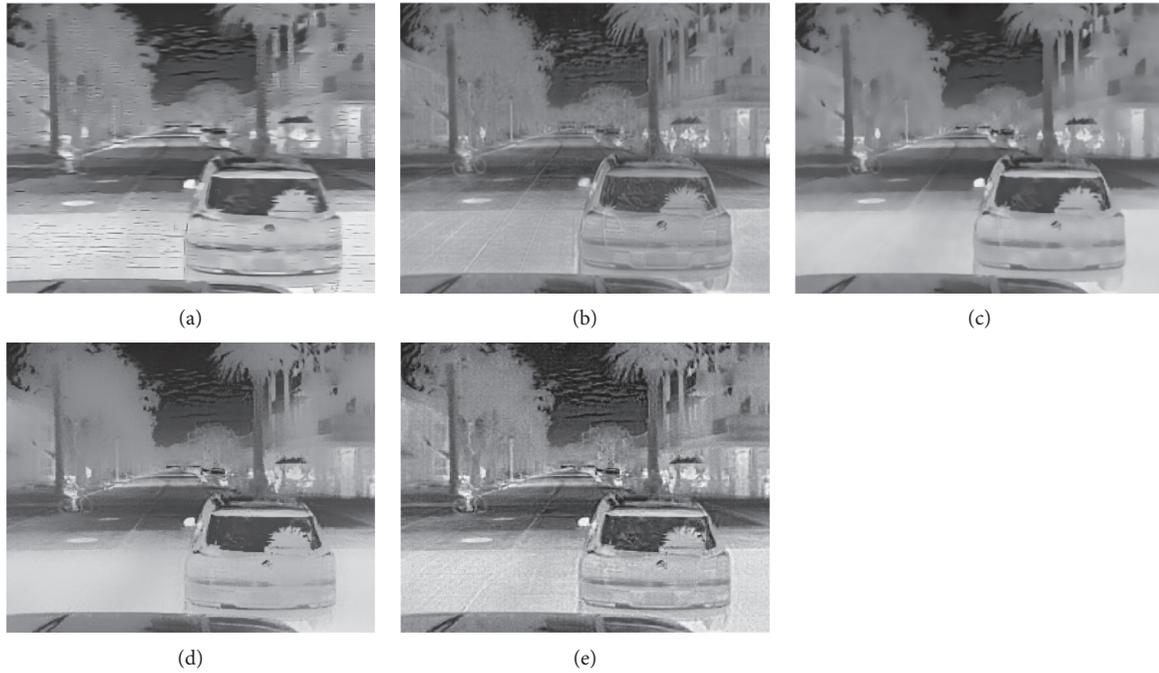


FIGURE 9: Effect picture of the ablation experiment. (a) Remove the dark channel prior loss function. (b) Remove the bright channel prior loss function. (c) Replace perceptual loss with L1 loss. (d) Replace perceptual loss with L2 loss. (e) Ours.

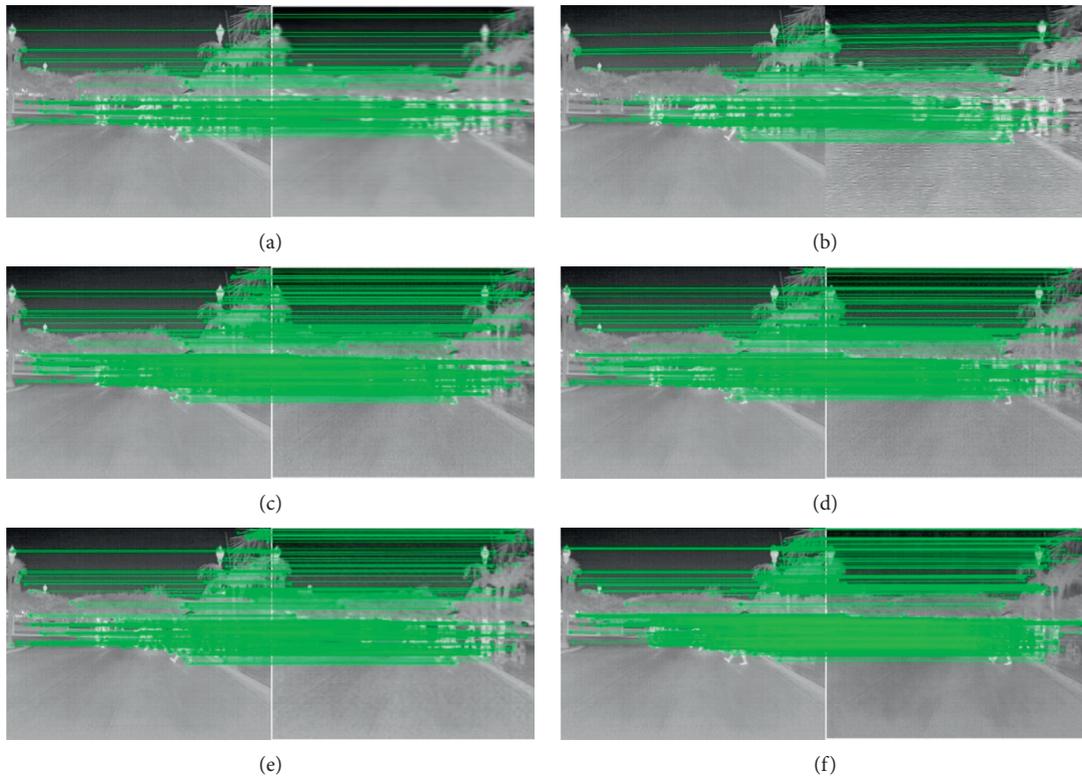


FIGURE 10: Deblurring image matching results (the images on the left are real and clear images). (a) Blurred image matching result. (b) DeepDeblur method deblurred image matching result. (c) DeblurGAN method deblurred image matching result. (d) CycleGAN method deblurred image matching result. (e) Cycle-Dehaze method deblurred image matching result. (f) Our method deblurred image matching result.

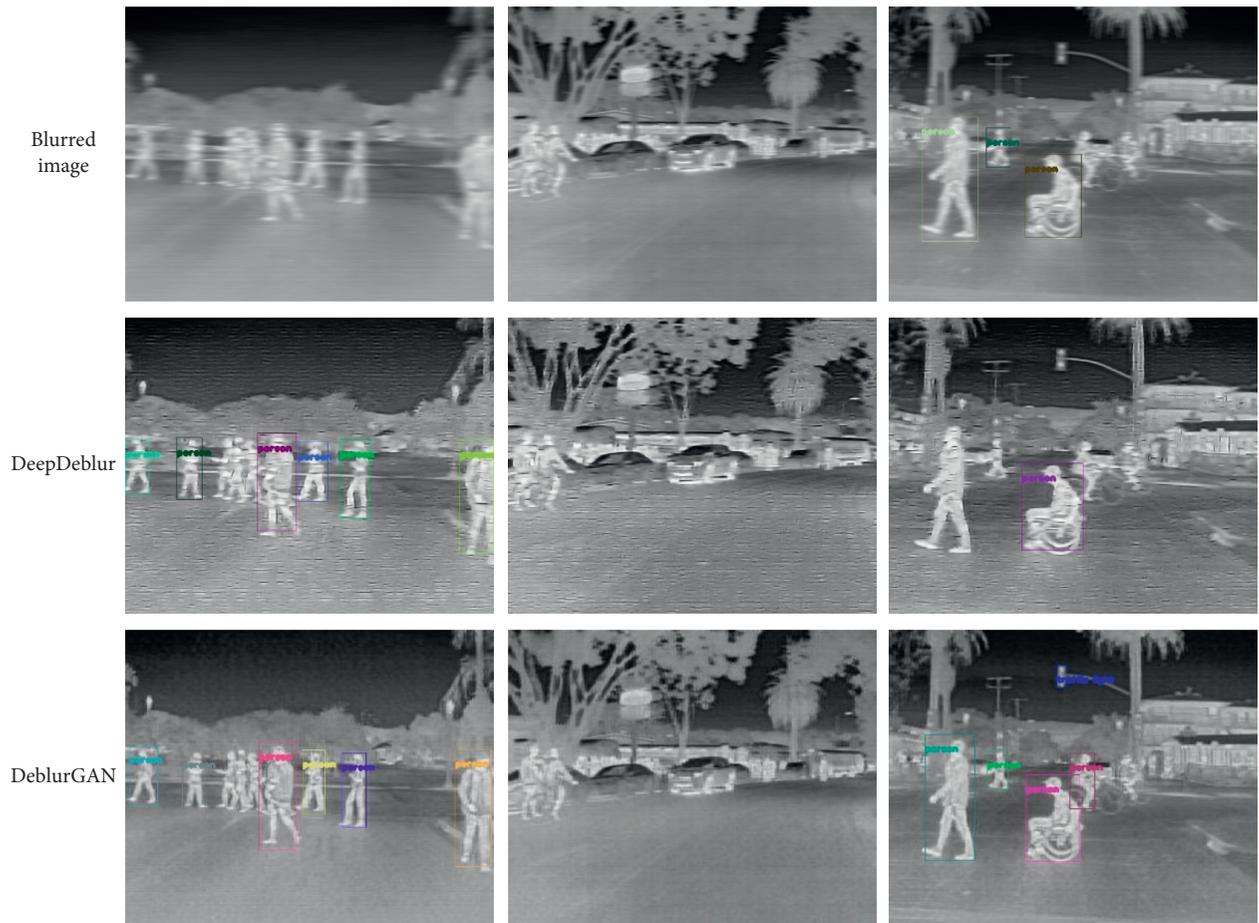


FIGURE 11: Continued.



FIGURE 11: Deblurring image target detection results.

extraction algorithm. In the matching result, the number of matching points can be used as a criterion for matching quality, and the corresponding matching points can also determine the similarity of the local features of the two images. Figure 10 shows the result of matching the deblurred image with the real clear image through the SIFT algorithm. It can be seen from the quantity that the deblurred image produced by our proposed method obtains more correct matching pairs than other methods.

In this experiment, we use the classic YOLO [65] method for deblurring image target detection (Figure 11). As can be seen, the proposed method to generate a blurred image has better detection result, and more targets can be detected.

## 5. Conclusion

Blind deblurring of a single infrared image is still a challenging computer vision problem. In this work, a method based on the GAN and channel prior discrimination is proposed for the problem of infrared image deblurring. Different from the previous deblurring work, we combine traditional blind deblurring and blind deblurring methods based on learning methods. Considering the different types of blur caused by the motion of the imaging device and the

imaging object, extensive experiments were carried out on different public datasets. Experimental results show that the proposed method is more competitive than other popular image deblurring methods in terms of deblurring quality (subjective and objective) and efficiency.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] A. Agrawal, *Motion Deblurring: Motion Deblurring Using Fluttered Shutter*, Cambridge University Press, Cambridge, UK, 2014.
- [2] N. Joshi, S. B. Kang, C. Lawrence Zitnick, and R. Szeliski, "Image deblurring using inertial measurement sensors," *ACM Transactions on Graphics*, vol. 29, no. 4CD, pp. 30–39, 2010.
- [3] B. Oswald-Tranta, M. Sorger, and P. O’Leary, "Motion deblurring of infrared images from a microbolometer

- camera,” *Infrared Physics & Technology*, vol. 53, no. 4, pp. 274–279, 2010.
- [4] B. Oswald-Tranta, “Temperature reconstruction of infrared images with motion deblurring,” *Journal of Sensors and Sensor Systems*, vol. 7, no. 1, pp. 13–20, 2018.
  - [5] N. Wang, W. Jing, Y. Zhang, and X. Sun, “Restoration of the infrared image blurred by motion,” in *Proceedings of the 2016 SPIE Society of Photo-optical Instrumentation Engineers*, Jinhua, China, October 2016.
  - [6] Y. Luo, T. Xu, N. Wang, and F. Liu, “Restoration of non-uniform exposure motion blurred image,” in *Proceedings of the 2014 International Symposium on Optoelectronic Technology & Application*, Beijing, China, November 2014.
  - [7] L. I. Jing, M. Wang, J. Sha, and B. Xujmet, “Research on wavelet transform based motion deblurring method of infrared target,” 2016.
  - [8] X. Liua, Y. Chena, Z. Penga, and J. Wu, “Total variation with overlapping group sparsity and lp quasinorm for infrared image deblurring under salt-and-pepper noise,” *Journal of Electronic Imaging*, vol. 28, no. 4, Article ID 043031, 2018.
  - [9] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, and W. T. Freeman, “Removing camera shake from a single photograph,” *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 787–794, 2006.
  - [10] D. Perrone and P. Favaro, “Total variation blind deconvolution: the devil is in the details,” in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, June 2014.
  - [11] L. Xu, S. Zheng, and J. Jia, “Unnatural L0 sparse representation for natural image deblurring,” in *Proceedings of the 2013 IEEE Conference on Computer Vision & Pattern Recognition*, Portland, OR, USA, June 2013.
  - [12] W. S. Lai, J. J. Ding, Y. Y. Lin, and Y. Y. Chuang, “Blur kernel estimation using normalized color-line priors,” in *Proceedings of the 2015 IEEE Computer Vision & Pattern Recognition*, Boston, MA, USA, June 2015.
  - [13] W. S. Lai, J. B. Huang, Z. Hu, N. Ahuja, and M. H. Yang, “A comparative study for single image blind deblurring,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, June 2016.
  - [14] T. Michaeli and M. Irani, “Blind deblurring using internal patch recurrence,” in *Proceedings of the 2014 European Conference on Computer Vision*, Zurich, Switzerland, September 2014.
  - [15] J. Pan, D. Sun, H. Pfister, and M. H. Yang, “Blind image deblurring using dark channel prior,” in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 2016.
  - [16] J. Pan, H. Zhe, Z. Su, and M. H. Yang, “Deblurring text images via L0-regularized intensity and gradient prior,” in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, June 2014.
  - [17] D. Perrone and P. Favaro, “A logarithmic image prior for blind deconvolution,” *International Journal of Computer Vision*, vol. 117, no. 2, pp. 159–172, 2016.
  - [18] D. Perrone and P. Favaro, “A clearer picture of total variation blind deconvolution,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 6, pp. 1041–1055, 2015.
  - [19] W.-Z. Shao, H.-S. Deng, Q. Ge, H.-B. Li, and Z.-H. Wei, “Regularized motion blur-kernel estimation with adaptive sparse image prior learning,” *Pattern Recognition*, vol. 51, no. C, pp. 402–424, 2016.
  - [20] W. Zuo, D. Ren, D. Zhang, S. Gu, and L. Zhang, “Learning iteration-wise generalized shrinkage-thresholding operators for blind deconvolution,” *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1751–1761, 2016.
  - [21] Z. Hu, L. Xu, and M. H. Yang, “Joint depth estimation and camera shake removal from single blurry image,” in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, June 2014.
  - [22] T. H. Kim and K. M. Lee, “Segmentation-free dynamic scene deblurring,” in *Proceedings of the 2014 Computer Vision & Pattern Recognition*, Columbus, OH, USA, June 2014.
  - [23] M. Noroozi, P. Chandramouli, and P. Favaro, “Motion deblurring in the wild,” in *Proceedings of the 2017 German Conference on Pattern Recognition*, Basel, Switzerland, September 2017.
  - [24] J. Pan, H. Zhe, Z. Su, H. Y. Lee, and M. H. Yang, “Soft-segmentation guided object motion deblurring,” in *Proceedings of the 2016 Computer Vision & Pattern Recognition*, Las Vegas, NV, USA, June 2016.
  - [25] O. Whyte, “Non-uniform deblurring for shaken images: derivation of parameter update equations for blind de-blurring,” 2010.
  - [26] S. Zheng, X. Li, and J. Jia, “Forward motion deblurring,” in *Proceedings of the 2013 IEEE International Conference on Computer Vision*, Sydney, Australia, December 2013.
  - [27] A. Gupta, N. Joshi, C. L. Zitnick, M. F. Cohen, and B. Curless, “Single image deblurring using motion density functions,” in *Proceedings of the 2010 European Conference on Computer Vision*, Heraklion, Crete, Greece, September 2010.
  - [28] Y. W. Tai, P. Tan, and M. S. Brown, “Richardson-lucy deblurring for scenes under a projective motion path,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 8, pp. 1603–1618, 2011.
  - [29] H. Zhang, D. Wipf, and Y. Zhang, “Multi-image blind deblurring using a coupled adaptive sparse prior,” in *Proceedings of the 2013 Computer Vision & Pattern Recognition*, Portland, OR, USA, June 2013.
  - [30] M. Hirsch, C. J. Schuler, S. Harmeling, and B. Schlkopf, “Fast removal of non-uniform camera shake,” in *Proceedings of the 2011 International Conference on Computer Vision*, Barcelona, Spain, November 2011.
  - [31] M. Hirsch, S. Sra, B. Scholkopf, and G. Spemannstrae, “Efficient filter flow for space-variant multiframe blind deconvolution,” in *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, CA, USA, June 2010.
  - [32] H. Ji and K. Wang, “A two-stage approach to blind spatially-varying motion deblurring,” in *Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition*, Providence, RI, USA, June 2012.
  - [33] A. Levin, “Blind motion deblurring using image statistics,” in *Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems*, Vancouver, British Columbia, Canada, December 2006.
  - [34] D. Gong, J. Yang, L. Liu et al., “From motion blur to motion flow: a deep learning solution for removing heterogeneous motion blur,” in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3806–3815, Honolulu, HI, USA, July 2017.
  - [35] M. Hradiš, “Convolutional neural networks for direct text deblurring,” in *Proceedings of the 2015 British Machine Vision Conference*, Swansea, UK, September 2015.
  - [36] O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin, and J. Matas, “DeblurGAN: blind motion deblurring using conditional adversarial networks,” in *Proceedings of the 2018*

- IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, June 2018.
- [37] S. Nah, T. H. Kim, and K. M. Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017.
- [38] S. Ramakrishnan, S. Pachori, A. Gangopadhyay, and S. Raman, "Deep generative filter for motion deblurring," in *Proceedings of the 2017 IEEE International Conference on Computer Vision Workshop (ICCVW)*, Venice, Italy, October 2017.
- [39] C. J. Schuler, M. Hirsch, S. Harmeling, B. Scholkopf, and M. Intelligence, "Learning to deblur," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 7, pp. 1439–1451, 2016.
- [40] J. Sun, W. Cao, Z. Xu, and J. Ponce, "Learning a convolutional neural network for non-uniform motion blur removal," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, June 2015.
- [41] P. Svoboda, M. Hradis, L. Marsik, and P. Zemcik, "CNN for license plate motion deblurring," in *Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, September 2016.
- [42] X. Tao, H. Gao, Y. Wang, X. Shen, J. Wang, and J. Jia, "Scale-recurrent network for deep image deblurring," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, June 2018.
- [43] A. Chakrabarti, "A neural approach to blind motion deblurring," in *Proceedings of the European Conference on Computer Vision*, Amsterdam, The Netherlands, October 2016.
- [44] S. Nah, T. Hyun Kim, and K. Mu Lee, "Deep multi-scale convolutional neural network for dynamic scene deblurring," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, July 2017.
- [45] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, July 2017.
- [46] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, July 2017.
- [47] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," 2017, <http://arxiv.org/abs/1701.07875>.
- [48] I. Goodfellow, J. Pouget-Abadie, M. Mirza et al., "Generative adversarial nets," 2014, <http://arxiv.org/abs/1406.2661>.
- [49] R. A. Yeh, C. Chen, T. Y. Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do, "Semantic image inpainting with deep generative models," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017.
- [50] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-resolution," in *Proceedings of the 2016 European Conference on Computer Vision*, Amsterdam, The Netherlands, October 2016.
- [51] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, October 2017.
- [52] B. Dai, S. Fidler, R. Urtasun, and D. Lin, "Towards diverse and natural image descriptions via a conditional GAN," in *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, October 2017.
- [53] C. Ledig, L. Theis, F. Huszar et al., "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017.
- [54] R. Yeh, C. Chen, T. Y. Lim, M. Hasegawa-Johnson, and M. Do, "Semantic image inpainting with perceptual and contextual losses," 2016, <http://arxiv.org/abs/1607.07539>.
- [55] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2341–2353, 2010.
- [56] Y. Xu, X. Guo, H. Wang, F. Zhao, and L. Peng, "Single image haze removal using light and dark channel prior," in *Proceedings of the 2016 IEEE/CIC International Conference on Communications in China (ICCC)*, Chengdu, China, July 2016.
- [57] Y. Yan, W. Ren, Y. Guo, R. Wang, and X. Cao, "Image deblurring via extreme channels prior," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017.
- [58] X. Wang, K. Yu, and S. Wu, "ESRGAN: enhanced super-resolution generative adversarial networks," 2018, <http://arxiv.org/abs/1809.00219>.
- [59] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <http://arxiv.org/abs/14091556>.
- [60] L. Xu, J. S. J. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in *Proceedings of the 27th International Conference on Neural Information Processing Systems*, pp. 1790–1798, Montreal, Canada, December 2014.
- [61] A. Chakrabarti, *A Neural Approach to Blind Motion Deblurring*, Springer International Publishing, Cham, Switzerland, 2016.
- [62] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman, "Understanding and evaluating blind deconvolution algorithms," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, June 2009.
- [63] U. Schmidt, C. Rother, S. Nowozin, J. Jancsary, and S. Roth, "Discriminative non-blind deblurring," in *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 604–611, Portland, OR, USA, June 2013.
- [64] D. Engin, A. Genc, and H. K. Ekenel, "Cycle-dehaze: enhanced CycleGAN for single image dehazing," in *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, June 2018.
- [65] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, June 2016.