

RESEARCH

Dynamic Scene Deblurring and Image De-Raining Based on Generative Adversarial Networks and Transfer Learning for Internet of Vehicle

Bingcai Wei¹, Liye Zhang^{1,2*}, Kangtao Wang¹, Qun Kong¹ and Zhuang Wang¹

*Correspondence:

zhangliye@sdut.edu.cn

¹School of Computer Science and Technology, Shandong University of Technology, Zibo, China

²Shandong Big Data Development and Innovation Laboratory, Zibo, China

Full list of author information is available at the end of the article

Abstract

Extracting traffic information from images plays an important role in Internet of Vehicle (IoV). However, due to the high-speed movement and bumpiness of the vehicle, motion blur will occur in image acquisition. In addition, in rainy days, because the rain is attached to the lens, the target will be blocked by rain, and the image will be distorted. These problems have caused great obstacles for extracting key information from transportation images, which will affect the real-time judgment of vehicle control system on road conditions, and further cause decision-making errors of the system and even cause traffic accidents. In this paper, we propose a motion blurred restoration and rain removal algorithm for IoV based on Generative Adversarial Network (GAN) and transfer learning. Dynamic scene deblurring and image de-raining are both among the challenging classical tasks in low-level vision tasks. For both tasks, firstly, instead of using ReLU in a conventional residual block, we designed a residual block containing three 256-channel convolutional layers, and we used the Leaky-ReLU activation function. Secondly, we used generative adversarial networks for the image deblurring task with our Resblock, as well as the image de-raining task. Thirdly, experimental results on the synthetic blur dataset GOPRO and the real blur dataset RealBlur confirm the effectiveness of our model for image deblurring. Finally, we can use the pre-trained model for the transfer learning-based image de-raining task and show good results on several datasets.

Keywords: Internet of Vehicle; Image deblurring; Image de-raining; Generative Adversarial Network; Transfer learning

1 Introduction

With the rapid development of automobile industry and the increasing number of vehicles, traffic safety and management problems have become increasingly prominent. In order to improve the efficiency of road traffic, ensure the safety of drivers and vehicles, and realize smart city and smart traffic, the interconnection between vehicles has become one of the key technologies. Therefore, the Internet of Vehicle (IoV) technology has been proposed by researchers. In the IoV technology, the information of vehicles, roads and personnel can be collected by sensors such as radar and camera, which can realize real-time monitoring of road traffic conditions, detect vehicle and pedestrian information, and use communication technology to share information with other vehicles. As an important part of the Internet of Things, IoV can use various communication technologies for data interconnection [1, 2]. Finally, the connection of vehicle-road, vehicle-people and vehicle-vehicle is

established to ensure the safety of people and vehicles and the healthy operation of traffic. Among all the data acquisition technologies, computer vision technology has advantages over other technologies in terms of cost, interactivity and security, which has attracted the attention of researchers. Visual-based traffic information extraction has become one of the indispensable capabilities of vehicles. Vision-based information perception capabilities, such as pedestrian and vehicle detection, recognition and instance segmentation, require accurate feature learning of urban street scene images.

Transportation images contain a lot of important information, such as the number of vehicles and pedestrians, vehicle license plate number, traffic signs, etc. These information is of great significance for traffic monitoring and automatic driving. Since the communication rate of the Internet of Things is limited by various conditions, the key information in the image must be extracted to reduce the amount of data transmitted[3, 4, 5]. Due to the high-speed motion of the vehicle, the relative motion between the camera and the object is very large in the short exposure time, and the turbulence of the vehicle causes the camera to vibrate, resulting in blurred motion in the image captured by the camera. Therefore, it is very difficult to extract information from the blurred image. In addition, the vehicle is usually located outside, in rainy weather, the camera will be blocked by rain, which seriously reduces the visual quality of the image and blocks the background objects. These visibility degradations have a negative impact on image feature learning, resulting in the failure of many computer vision systems. These make removing undesirable visual effects from images caused by movement and rain a very desirable technique.

In order to improve the quality of transportation images by image deblurring and de-raining technologies, researchers proposed a variety of algorithms. The previous image deblurring algorithms first estimated the blur kernel of the image from the given image before deblurring. Secondly, assuming that the image has a certain blur feature, and the image prior information is used in the process of image deblurring. Finally, the image deblurring is realized and the clear image is obtained. In recent years, deep learning algorithms represented by Convolutional Neural Network (CNN) have been heavily applied to the field of image blind deblurring. Compared with earlier blind deblurring algorithms based on image prior, deep learning algorithms can achieve better results than them. For example, Xu *et al.* [6] introduced a novel, separable structure of convolutional structure for deconvolution and achieved good deblurring results, and Su *et al.* [7] used CNN for end-to-end training to achieve video deblurring using frame-to-frame information in video. Similarly, due to the rapid development of deep learning algorithm, especially the excellent performance of convolutional neural networks, numerous methods based on deep learning have emerged in the field of image de-raining. The aim of these methods is to learn and obtain a mapping function between rain image and clear image, and thus solve the problem of image de-raining based on this mapping function [8].

However, before the establishment of RealBlur data set [9], the training data set used by image deblurring algorithm based on deep learning is synthetic data set, such as GOPRO dataset [10]. This is because a single imaging device (such as a camera or smartphone) cannot capture a blurred image while capturing a clear image, but the model training of deep learning requires blurred-clear image pairs, with blurred images as the input and clear images as the labels.

In 2014, Goodfellow *et al.* [11] proposed Generative Adversarial Networks (GAN). GAN consists of two competing networks, one called generator and the other called discriminator. The generator is responsible for receiving random noise input and then synthesizing data samples. The synthetic data should be as real as possible in order to "fool" the discriminator. The discriminator is responsible for determining whether the input data is a "fake" sample synthesised by the generator or a real sample, in order to distinguish them. The goal of a good generative countermeasure network is to make the probability of the discriminator close to 0.5, that is, it is impossible to judge whether the sample is generated by the generator.

Due to its powerful performance, GAN was soon applied to the field of image deblurring. The generator in GAN is responsible for receiving and recovering the blurred image, and generating a deblurring image similar to a clear image to deceive the discriminator, and the discriminator receives the original clear image and the generator's deblurring image respectively and tries to distinguish them. GAN have been applied within the field of image deblurring, such as DeblurGAN proposed by Kupyn *et al.* [12], and conditional GAN and Content Loss Function (CLF) are used in this algorithm to eliminate the blurring of motion images. Because the synthetic blurred images obtained by a single neural network on blurred and clear data cannot accurately simulate the blurring process of real scenes, two GANs were used to deblur the motion blurred images, one GAN is used for image blur and the other is used for image deblurring, and finally the real blur and deblurring process is realized [13].

Transfer learning is to apply knowledge or patterns learned in a certain domain or task to different but related fields or problems [14, 15]. With the help of original domain data, transfer learning can reduce the dependence on target domain data. For the two tasks of image deblurring and image de-raining, they have a lot of similarity characteristics. Therefore, after using the blurred data set to train the deep learning algorithm, with the help of the transfer learning algorithm, we can save a lot of time, calculation and data, and achieve better learning effect in image de-raining.

Inspired by the previous research results, the contributions of this paper are as follows. Firstly, we design a residual block which contains three convolution layers and 256 channels. and we use the Leaky Rectified Linear Unit (LReLU) activation function to replace the Rectified Linear Unit (ReLU) in the traditional residual block; Secondly, the proposed residual block is applied to GAN, and the GAN is used to train the image deblurring task. Thirdly, the experimental results on synthetic blur dataset GOPRO and real blur dataset RealBlur verify the effectiveness of our model. Finally, the pre-trained model is used to implement the image de-raining task based on transfer learning, which shows good results on multiple datasets.

2 Related Works

2.1 Image Blurring: synthetic and real

The blurring of images is caused by a number of factors, the process of which can be expressed as follows.

$$I_B = K * I_S + N \quad (1)$$

where I_B and I_S are blurry and sharp images, respectively. K is a known blur kernel (non-blind deblurring) or an unknown blur kernel (blind deblurring) and N is additive noise. And $*$ represents the convolution operation.

Alternatively, blurred images can be generated by frame-by-frame blurring. For blurred spatially varying images, there is no camera response function (CRF) estimation technique[16], and the CRF can be approximated as a known CRF, as shown in Equation 2.

$$g(I_{S[i]}) = I_{S'[i]}^{\frac{1}{\gamma}} \quad (2)$$

In the above formula, γ is a parameter, it is generally considered to be equal to 2.2, and the real clear image $I_{S[i]}$ can be obtained from the observed clear image $I_{S'[i]}$.

The blurred image of the simulation can then be obtained by the following equation.

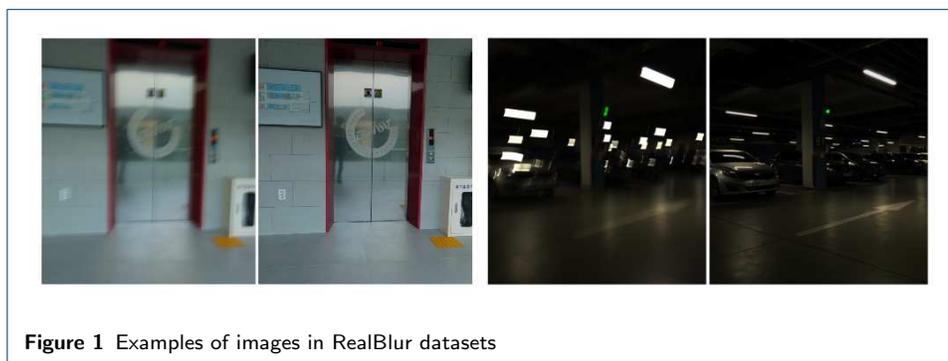
$$I_B \simeq g\left(\frac{1}{M} \sum_{t=1}^M I_{S[t]}\right) \quad (3)$$

where M is the number of sharp frames.

In the real world images are captured during continuous exposure, so a true blur is the integration of multiple clear frames. This can be expressed as:

$$I_B = g\left(\frac{1}{T} \int_{t=0}^T I_{S(t)} dt\right) \quad (4)$$

Before Jaesung Rim *et al.* proposed RealBlur, a real-world blur dataset that can be used for deep learning, deep learning methods in the field of image deblurring were mostly trained and tested on synthetic simulated blur datasets, where non-professional image collection devices could only save blurred images and not the corresponding clear images when taking blurred images. The RealBlur dataset is the first publicly available real blur dataset that can be used for deep learning, using multiple cameras and optical devices to capture both blurred and clear image pairs using professional image collection devices. Figure 1 shows the examples of blurred and clear images in RealBlur datasets.



2.2 Image deblurring

Image deblurring is the operation of recovering, or deblurring, a given blurred image to obtain the corresponding ground truth image.

Non-blind deblurring refers to the deblurring of an image by a given known blur kernel, while the blind deblurring problem refers to the estimation of the original image X and the blur kernel Z from a given noisy image Y .

The blind deblurring process can be expressed as:

$$\{\hat{X}, \hat{Z}\} = \operatorname{argmin} \|Z \oplus X - Y\|_2^2 + \phi(X) + \theta(Z) \quad (5)$$

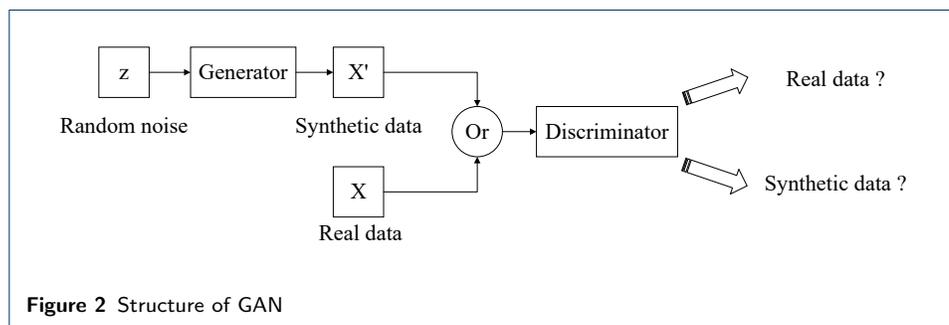
where $\phi(X)$ and $\theta(Z)$ are the regularization terms and possible blur kernels of the expected clear image, respectively.

Image deblurring technology can be divided into traditional optimization-based methods and deep learning-based methods, especially the convolution neural network. Of course, according to whether the blur kernel needs to be estimated, it can also be divided into blind restoration without estimating the blur kernel and non-blind restoration with estimating the blur kernel. In the application scenario, non-blind restoration is more widely used, and most of the traditional deblurring methods require blur kernel, while most of the deblurring methods based on deep learning do not need to estimate the blur kernel. Most traditional optimization-based deblurring methods guide the Maximum A-Posteriori Probability (MAP) process by assuming a priori [17]. For example, Pan *et al.* [18] proposed a image deblurring method using dark channel prior, Xu *et al.* [19] proposed a L0 gradient prior with clear edge information, Kim *et al.* [20] proposed a priori method based on the total variation-L1 model. Due to the rapid development of deep learning in recent years and the strong performance of convolution neural network, which is widely used in the field of computer vision, there are also a large number of methods using deep learning and convolution neural network in the field of image deblurring, which can be roughly divided into the methods of using convolution neural network to estimate the blur kernel. For example, Sun *et al.* [21] used CNN to estimate the probability distribution of the blur kernel. With the end-to-end deblurring method without estimating the blur kernel, most of the deblurring methods using CNN do not need to estimate the blur kernel, and directly realize the image deblurring process, such as Gong *et al.*[22] proposed a fully convolutional neural network, which can estimate the blur kernel at pixel level. Nah *et al.*[10] proposed a multi-scale convolutional neural network, which can use the feature association between images of different sizes, such as 64×64 , 128×128 , 256×256 , to achieve more refined image deblurring. Tao *et al.*[23] also selected the multi-scale structure, and added Long-Short Term Memory(LSTM) into it to propose a multi-scale recurrent neural network. However, since the multi-scale convolutional neural network from coarse to fine takes more time in deconvolution operation, Zhang *et al.*[24] proposed a multi-slice structure, which cuts the input image into multiple slices, and at the same time, multiple encoders and decoders can be arranged in multiple ways to achieve better deblurring performance. The application of image deblurring in IoV, such as filter-deblurGAN proposed by Zhou *et al.*[25], filter-deblurGAN can judge whether the image is blurred, and can be directly applied to the vehicle logo detection (

VLD) task. With the vehicle logo detection method, the motion blurred vehicle logo image can be directly detected.

2.3 GAN

Generative Adversarial Networks (GAN) is a deep learning model, which is one of the most promising unsupervised learning methods on complex distribution in recent years. In the early generative adversarial network, generator and discriminator are not required to be neural networks, and only the corresponding generative and adversarial functions are required.



As shown in figure 2, GAN contains two competing networks generator and discriminator. The idea of confrontation in GAN can be traced back to Nash equilibrium of game theory. The two sides of confrontation are generator and discriminator. The generator is responsible for generating samples as close as possible to the target image and cheating the discriminator, while the discriminator is responsible for distinguishing the generated image from the samples generated by the generator. The objective function of the confrontation can be described as follows:

$$\min_G \max_D V(D, G) = E_{x \sim P_{data(x)}} [\log D(x)] + E_{z \sim P_z(x)} [\log(1 - D(G(z)))] \quad (6)$$

x indicates that the real sample comes from the real data distribution $P_{data(x)}$, $E_{x \sim P_{data(x)}}$ is the expectation of inputting clear images. $D(\cdot)$ represents the output of D , $G(\cdot)$ represents the output of G .

Since its invention, GAN has been the focus of research in the field of deep learning, and there are various variants, such as WGAN[26], which improves GAN in terms of loss function, proposes the wassertein distance loss function and weight clipping measures, further improving the performance of GAN; WGAN-GP[27], which improves on WGAN, proposes the weight penalty (Weight) measures, effectively preventing the gradient disappearance, gradient explosion and the difficulty of weight restriction that may occur in WGAN, etc.

2.4 Image deraining

In simple and broad terms, the band rain image model can be modelled as:

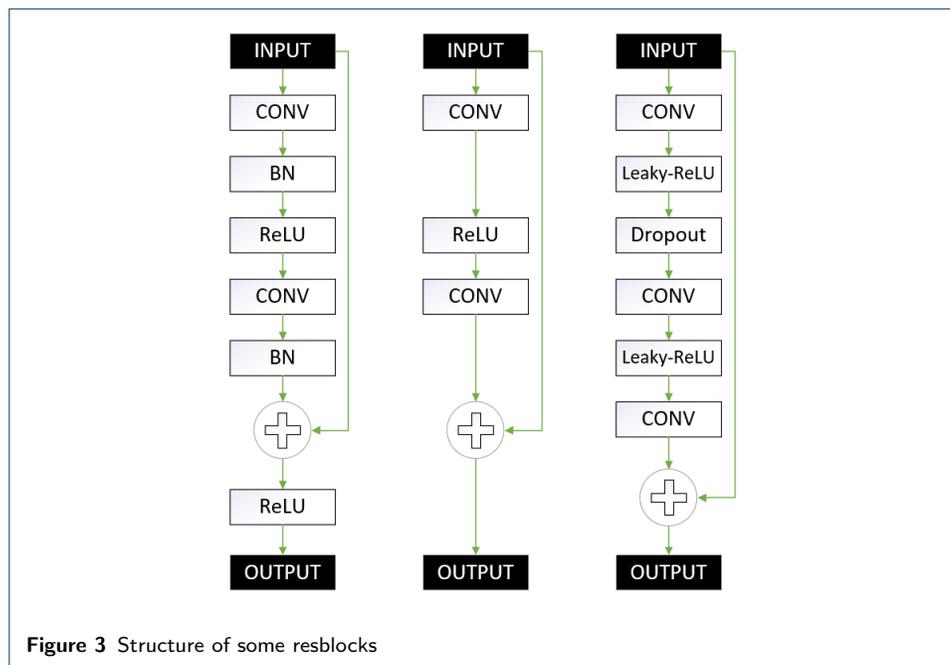
$$Y = X + W \quad (7)$$

where Y represents the rain-bearing image, X represents the original clear image, and W represents rain-streak component. Thus, the goal of a single image de-raining task is to recover the original clear image X from a given rain-bearing image Y . Similar to image deblurring, which is also a low-level vision task, earlier image de-raining methods also used image priors to solve it. For example, sparse coding-based methods[28], GMM-based (Gaussian Mixture Model)[29] based methods and patchrank prior methods[30]. Applications of single image de-raining in IoV, such as Sun *et al.*[31], proposed a convolution neural network with rainy images as input, which can directly recover clear images in the case of atmospheric veiling effects caused by distant rain streak accumulation.

3 Methods

3.1 The proposed Resblocks

The residual block[32] consists of two weight layers, including a ReLU activation function in the middle, then a skip block, and then a ReLU activation function. The jump junction block can realize the cross layer propagation of gradient, which is helpful to overcome the gradient attenuation phenomenon. By adding residual blocks, the problem of gradient vanishing and gradient exploding can be solved with deepening network structure.



Our residual block in this paper consists of three convolutional layers, each with 256 channels, two Leaky-ReLU activation functions are used, which allows for faster convergence, and a Dropout layer with probability 0.5 is added between the first and second convolutional layers, which helps prevent overfitting of the model while speeding up model training. Finally there is a jump connection module which helps to solve the gradient disappearance problem as well as the gradient explosion problem. Also, since BN layers have been shown to increase computational complexity and degrade performance [10, 13], the discriminator in this paper removes the Batch

Normalization (BN) layer, while most of the studies in this research area that use deep learning for deblurring use small batches for training, such as Nah *et al.* [10] with a training batch of 2 and Kupyn *et al.* [12] proposed DeblurGAN, with a training batch of 1, and Zhang *et al.* [13] proposed realistic blur-based deblurring with a batch of 4. The use of small batches for training is not suitable for using batch normalisation layers. The structure of original residual block [32], residual block in [10], and our residual block is shown in figure 3. All the channel of conv layers in the resblock is 256.

3.2 Loss function

In this paper, we use the Wassertein distance from WGAN [26] as the loss function of the discriminator, which is defined as follows.

$$W(P_{data}, P_g) = \inf_{\gamma \sim \Pi(P_{data}, P_g)} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (8)$$

In Equation (8), x and y denote real samples and generated samples respectively, $\Pi(P_{data}, P_g)$ is the set of joint distribution of P_{data} and P_g , $(x, y) \sim \gamma$ denotes the samples in it, \inf denotes the distance expectation of the sampled real samples and generated samples, that is, $E_{(x,y) \sim \gamma} [\|x - y\|]$ takes the lower limit.

Meanwhile, this paper uses Perceptual loss [33] as the loss function of the generator. Perceptual loss is an L2 loss based on the difference between the CNN feature maps of the generated and target images. Unlike the normal L2 loss, the content loss is defined by the output features of one layer of the pre-trained network.

$$L_x = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\Phi_{i,j}(I^S)_{x,y} - \Phi_{i,j}(G_{\theta_G}(I^B))_{x,y})^2 \quad (9)$$

where $\Phi_{i,j}$ represents the feature map extracted by pre-training the convolutional neural network, the pre-trained model used in this paper is VGG19. $H_{i,j}$ and $W_{i,j}$ represents the size of the feature map.

3.3 Network structure

3.3.1 Structure of generator

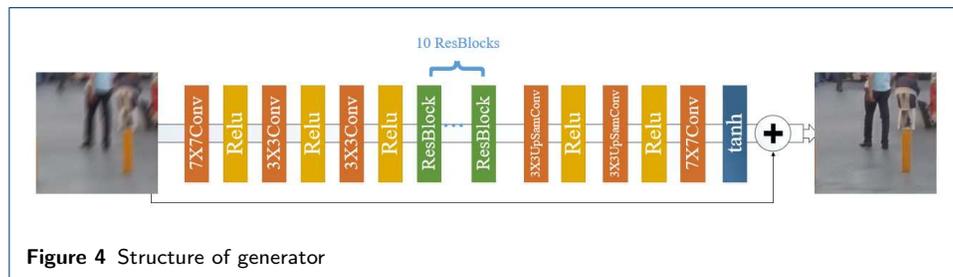


Figure 4 Structure of generator

The network structure is shown in the figure 4. The first part in figure 4 is a convolution layer with convolution core size of 7×7 , and the output results are input into two convolution blocks with step length of 2, and six improved residual blocks

are located at the third layer, then the system connects two transposed convolution blocks with convolution core size of 3×3 with step length of 2. The next layer is a 7×7 convolution layer with a tanh activation function as the activation function. In addition to this layer, the generator’s activation functions are all ReLU activation functions. Finally, there is a jump connection block, and the filling method of all the convolution layers is the same.

3.3.2 Structure of discriminator

PatchGAN [34] is a Markovian discriminator proposed by Phillip Isola *et al.* that can efficiently model images as Markov random fields. the PatchGAN discriminator attempts to classify each $N \times N$ block in an image and uses this to determine whether it is a generator-generated fake sample or a real sample, averaging all the eigenvalues of the final layer output to PatchGAN extends the perceptual field of the bottom convolutional layer to 70×70 by superimposing five convolutional layers. Inspired by the PatchGAN discriminator, the discriminator structure used in this paper is similar with it, but we add two dense layers at the bottom of it. The structure of discriminator is shown in table 1.

Table 1 Structure of discriminator

#	Layer	Size of feature map	Stride	Parameters	Receptive field
1	Input	$256 \times 256 \times 3$	-	-	-
2	Conv	$128 \times 128 \times 64$	2	3136	142
3	Conv	$64 \times 64 \times 64$	2	65600	70
4	Conv	$32 \times 32 \times 128$	2	131200	34
5	Conv	$16 \times 16 \times 256$	2	524544	16
6	Conv	$16 \times 16 \times 512$	1	2097644	7
7	Conv	$16 \times 16 \times 1$	1	8193	4
8	Dense	-	-	263168	-
9	Dense	-	-	1025	-

4 Results and discussion

4.1 Train details

Although GAN has good generative models, the training of the model is not stable enough, resulting in difficulty in convergence and mode collapse. To address these drawbacks, WGAN uses weight pruning to achieve constraints on the discriminator, thus solving the above problems of generative adversarial networks, but it is not completely solved, for example, it is still difficult to converge sometimes. Later, WGAN-GP, which uses a gradient penalty to achieve better performance and further solve the problems of GAN. In contrast to the above mentioned methods of improving GAN’s shortcomings, we found that we can also solve the shortcomings of GAN to some extent by a small trick, which we call the gradient training strategy. Specifically, after a certain stage of training, the trained weights of the generator are saved and training is stopped, but the weights of the discriminator are not saved, and then training is restarted and the generator is loaded with the previously trained weights, while the discriminator is trained again. We use an incremental training strategy throughout the training process, and our experiments show that this training technique can address the shortcomings of the GAN to a certain extent without making changes to the network internals during the training process.

The simulation experiments in this paper were conducted on a server configured with an Nvidia Tesla-P100 using the TensorFlow2 framework with the Adam optimizer, with the initial learning rate set to 10^{-4} . After several iterations of training, the final learning rate decreased linearly to 10^{-7} . During training, images are randomly flipped horizontally and vertically to improve the generalization capability and robustness of the model, while at the same time random Gaussian noise is added to the images during the training process.

Due to the input requirements of the network structure in this paper, the images in the training dataset need to be cropped to a size of 256×256 . The generator is a Fully Convolutional Network (FCN), which can be applied to images of arbitrary size.

4.2 Deblurring on GOPRO Datasets

GOPRO dataset is one of the most common datasets for image deblurring research. GOPRO4 camera is used to shoot the video with 240 frames per second, and then generate blur images to restore real motion blur. We used 2103 pictures as the training set to train the model, and the remaining 1110 pictures as the test set to test the training effect of the model. The comparison of deblurred images obtained by the proposed method and other methods is displayed in figure 5 and 6, and the Peak Signal-to-Noise Ratio (PSNR) and Structural SIMilarity (SSIM) of different algorithms are shown in Table 2. Through the simulation results in Figure 5 and Figure 6, it can be seen that the results obtained by using the proposed algorithm for blurred image restoration are more clear, and the data of PSNR and SSIM also confirm our simulation results. The PSNR of the proposed algorithm reaches 29.19, and the SSIM reaches 0.883.

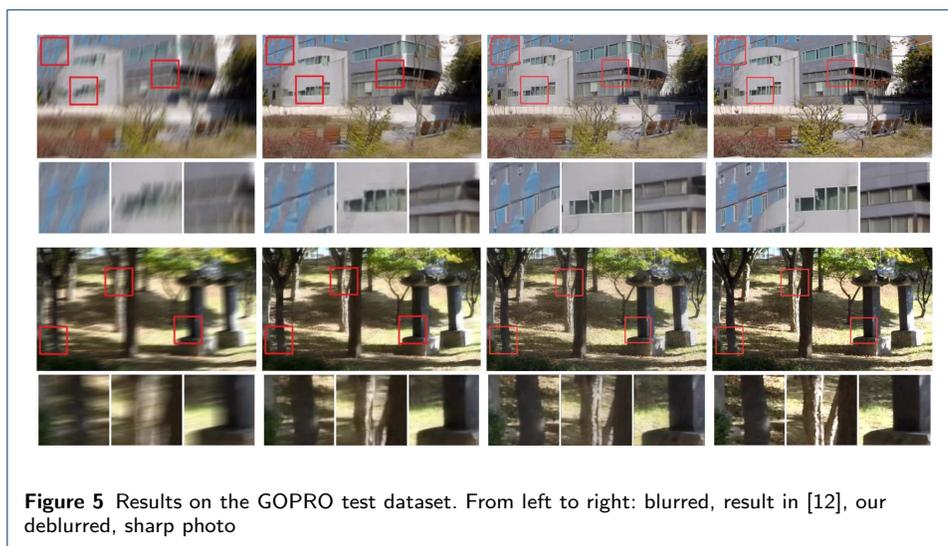


Figure 5 Results on the GOPRO test dataset. From left to right: blurred, result in [12], our deblurred, sharp photo

Table 2 Performance comparison on the GOPRO dataset

Method	Xu[6]	Kim [35]	Sun[21]	Gong[22]	DeblurGAN[12]	Ours
PSNR	21.00	23.64	24.64	26.40	28.70	29.19
SSIM	0.741	0.824	0.843	0.863	0.858	0.883

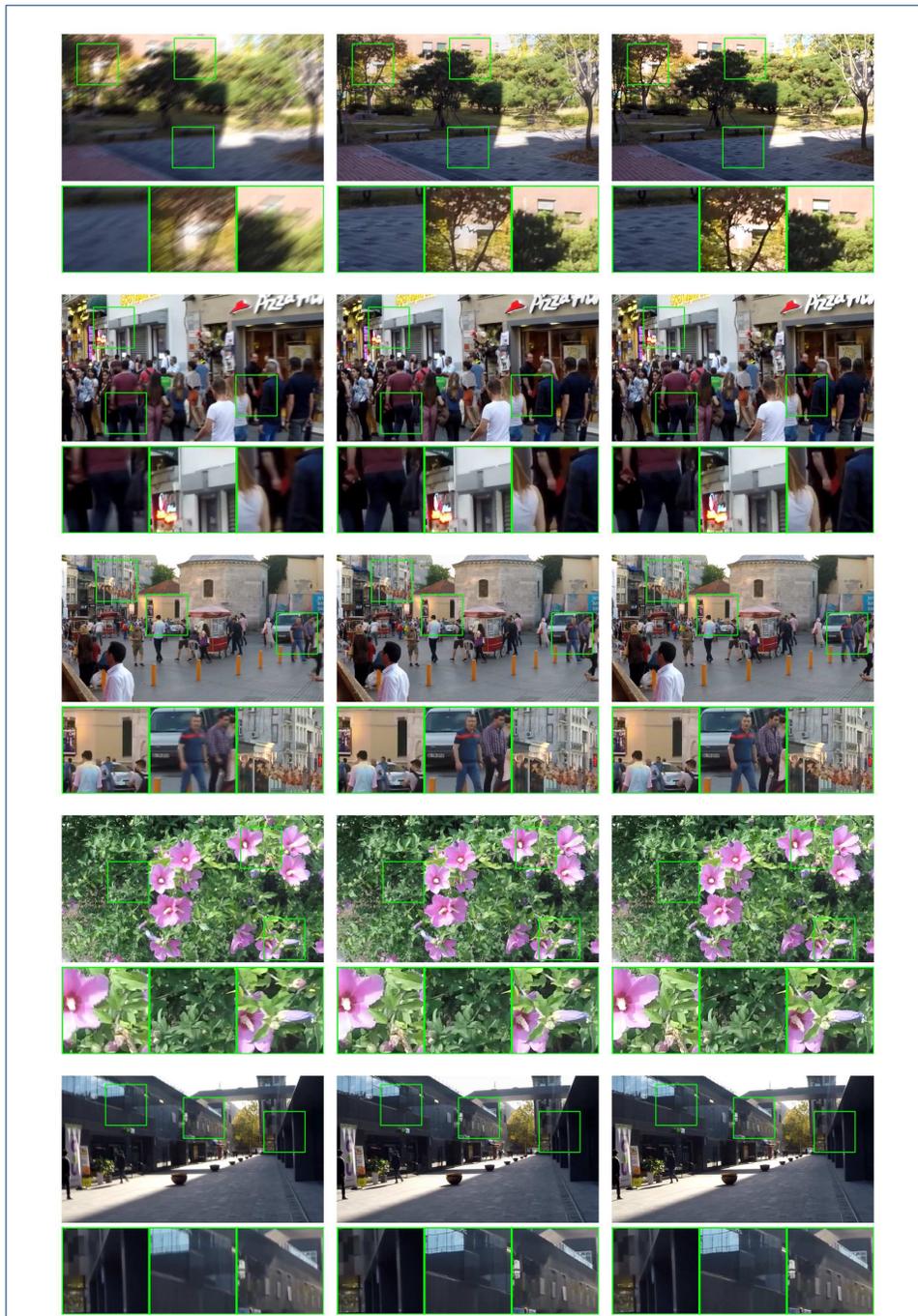


Figure 6 Results on the GOPRO test dataset. From left to right: blurred, our deblurred, sharp photo, all green small areas in the pictures in the above different scenes are randomly selected

4.3 Deblurring on RealBlur Datasets

The RealBlur dataset, the first large blurred clear image pair dataset that can be used as a deep learning training dataset, was created by two cameras in a single photoreceptor. The real blur dataset consists of two parts. One part is RealBlur-J, which is a pair of JPEG images formed after camera shooting; the other part is

RealBlur-R, which is the untransformed image pair captured by the camera. Some examples of the deblurring results trained and tested on the RealBlur dataset are shown in Figures 7 and 8, which show that our model can also achieve good deblurring results on the RealBlur dataset. The PSNR and SSIM results are displayed in Table 3, compared with other algorithms, the proposed algorithm has the best PSNR and SSIM results.

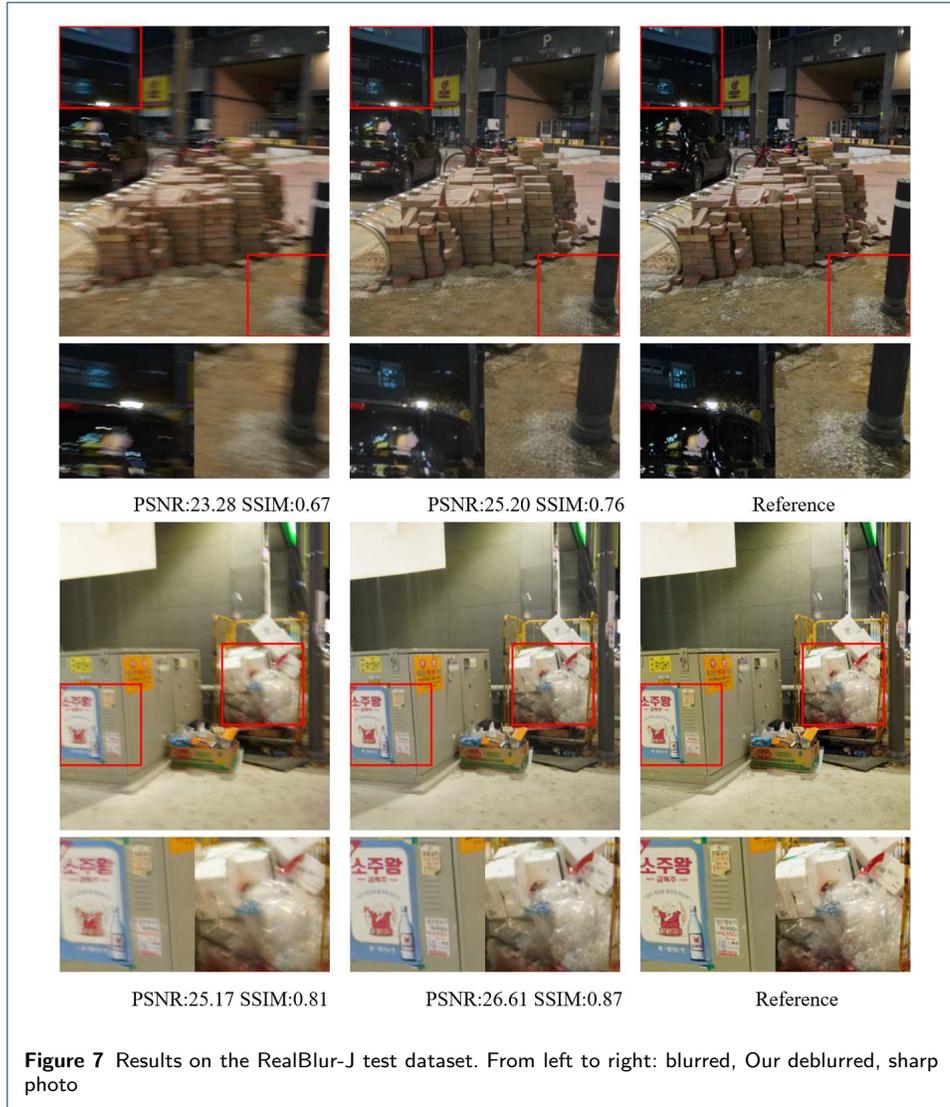
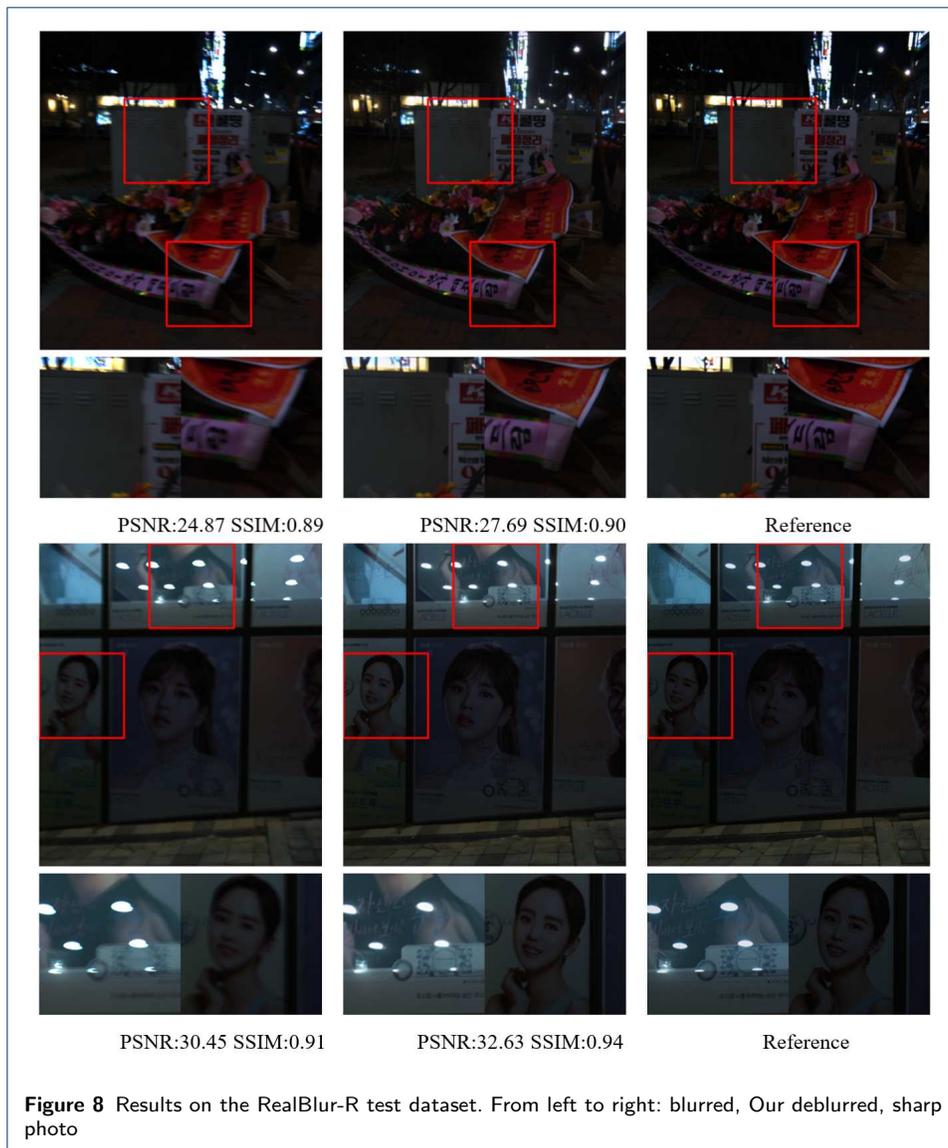


Table 3 Performance comparison on the RealBlur dataset

Method	RealBlur-J		RealBlur-R	
	PSNR	SSIM	PSNR	SSIM
Hu[36]	26.41	0.803	33.67	0.916
Nah[10]	27.87	0.827	32.51	0.841
DeblurGAN[12]	27.97	0.834	33.79	0.903
Pan[18]	27.22	0.790	34.01	0.916
Ours	27.73	0.833	34.32	0.911



4.4 Image de-raining based on transfer learning

In order to validate the transfer learning-based image de-raining capability of our proposed model, we trained and tested it on the following datasets: Rain800[37], Rain100H[38], Rain-100L[38]. All the above datasets are synthetic image de-raining datasets used for deep learning. On a test set of the Rain800 dataset, we used transfer learning to train on the Rain800 training set using a model that had been previously trained for image deblurring. On the Rain100H dataset, the transfer learning-based image deblurring model just trained on the Rain800 training set was used, and the same on the Rain100L dataset.

The de-raining results obtained by our method and other methods are illustrated in Figure 9 and Figure 10, and the data of PSNR and SSIM of different methods are exhibited in Table 4. It is worth noting that the training set in reference [39] consists of 11,200 image pairs, which will cost huge time and computational resources for model training. In comparison, the datasets in reference [37] and reference [38] only

contain 700 and 1,800 image pairs respectively. Based on the deblurring model, the transfer learning method is used to train the trained model in dataset [37], and the image deblurring model can be quickly applied to the de-raining task. The simulated results shown in Table 4, except our method, all the other methods were trained on dataset[37], dataset [38] and dataset [39]. Although our algorithm is not trained on the largest dataset [39], it still achieves the best de-raining effect, which helps us save more than 80% of the training time and resources. The test set of dataset [38] is divided into two parts: Rain100H and Rain100L. However, the proposed model trained only on the dataset [37] can achieve PSNR 18.30 and SSIM 0.476 on Rain100H and SSIM 0.865 and PSNR 26.21 on Rain100L.



5 Conclusions

In IoV, transportation image is one of the important data sources. In order to extract traffic information from motion blurred images and rain images, the image deblurring method and the image de-raining method based on deep learning and



Table 4 De-raining performance comparison on the serval datasets.

Method	Rain800		Rain100L		Rain100H	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DerainNet[40]	22.77	0.810	27.03	0.884	14.92	0.592
SEMI[41]	23.64	0.788	25.03	0.842	16.56	0.486
DIDMDN[8]	22.56	0.818	25.23	0.471	17.35	0.524
Ours	25.50	0.825	29.91	0.934	21.99	0.787

transfer learning are studied in this paper, and a motion deblurring algorithm based on deep residual generation adversarial network with our own designed resblock is proposed to achieve higher accuracy in the image deblurring task. After the model training, in order to reduce the training workload of the image rain removal model, the trained model is applied to the image de-raining task by transfer learning, and satisfactory results are also achieved. We tested the proposed method on multiple

public datasets, and the experimental results demonstrate that the proposed method leads to significantly improvements on image deblurring and image de-raining.

Acknowledgements

This work is supported by School of Computer Science and Technology, Shandong University of Technology.

Funding

This paper is supported by Shandong Provincial Natural Science Foundation, China (grant number ZR2019BF022), and National Natural Science Foundation of China (grant number 62001272).

Abbreviations

IoV: Internet of Vehicle; GAN: Generative Adversarial Network; LReLU: Leaky Rectified Linear Unit; CNN: Convolutional Neural Network; CLF: Content Loss Function; CRF: Camera Response Function; MAP: Maximum A-Posteriori Probability; LSTM: Long-Short Term Memory; WGAN: Wasserstein GAN; WGAN-GP: WGAN Gradient Penalty; GMM: Gaussian Mixture Model; FCN: Fully Convolutional Network; PSNR: Peak Signal-to-Noise Ratio; SSIM: Structural SIMilarity

Availability of data and materials

All the datasets used for training the model of this paper are come from internet.

Ethics approval and consent to participate

Text for this section. . .

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Authors' contributions

The algorithms proposed in this paper have been conceived by L. Zhang and B. Wei. B. Wei and L. Zhang made the analysis and experiment and wrote the paper. K. Wang, Q. Kong and Z. Wang investigated, validated and revised this paper. The authors approved the final manuscript.

Authors' information

Bingcai Wei received the bachelor's degree in software engineering from Qufu Normal University. He is currently pursuing the M.Sc. degree with the School of Computer Science and Technology, Shandong University of Technology. His current research interests include machine learning and image processing.

Liye Zhang received the M.Sc. and Ph.D. degrees in communication engineering from the Harbin Institute of Technology, in 2011 and 2018, respectively. From 2014 to 2015, he was a Visiting Scholar with Department of Electrical and Computer Engineering, University of Toronto, Canada. He is currently a Lecturer with the Shandong University of Technology. His current research interests include Indoor Localization, Computer Vision and Machine Learning.

Kangtao Wang received his bachelor's of Engineering degree from the School of Software of Pingdingshan College in 2020. He is currently pursuing a master's degree in the School of Computer Science and Technology at Shandong University of Technology. His current research interests include machine learning and image stitching.

Qun Kong received the bachelor's degree in computer science and technology from DeZhou University in 2020. She is currently pursuing the M.Sc. degree with the School of Computer Science and Technology, Shandong University of Technology. Her current research interests include machine learning and binocular vision.

Zhuang Wang received the bachelor's degree in computer science and technology from Shandong Youth University for Political Sciences in 2020. He is currently pursuing the M.Sc. degree with the School of Computer Science and Technology, Shandong University of Technology. His current research interests include machine learning and indoor localization.

Author details

¹School of Computer Science and Technology, Shandong University of Technology, Zibo, China. ²Shandong Big Data Development and Innovation Laboratory, Zibo, China.

References

1. F. Li, K. Lam, X. Liu, J. Wang, K. Zhao and L. Wang, "Joint Pricing and Power Allocation for Multibeam Satellite Systems With Dynamic Game Model," in *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2398-2408, March 2018
2. Xin Liu, Xueyan Zhang. NOMA-based Resource Allocation for Cluster-based Cognitive Industrial Internet of Things. in *IEEE Transactions on Industrial Informatics*. Vol.16, no. 8, 2020: 5379-5388.
3. Xin Liu and Xueyan Zhang. Rate and Energy Efficiency Improvements for 5G-Based IoT With Simultaneous Transfer. in *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 5971-5980, Aug. 2019
4. Xin Liu, Xueyan Zhang, Min Jia, et al. 5G-based green broadband communication system design with simultaneous wireless information and power transfer. *Physical Communication*. Vol. 28, 2018: 130-137.
5. Xin Liu, Xiangping Zhai, Weidang Lu and Celimuge Wu, QoS-guarantee Resource Allocation for Multibeam Satellite Industrial Internet of Things with NOMA. in *IEEE Transactions on Industrial Informatics*. vol. 17, no. 3, 2021: 2052-2061.

6. Xu L, Ren J S J, Liu C, et al. Deep convolutional neural network for image deconvolution. *Advances in Neural Information Processing Systems*, 2014, 2: 1790-1798.
7. S. Su, M. Delbracio, J. Wang, G. Sapiro, W. Heidrich and O. Wang, "Deep Video Deblurring for Hand-Held Cameras," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 237-246, doi: 10.1109/CVPR.2017.33.
8. H. Zhang and V. M. Patel, "Density-Aware Single Image De-raining Using a Multi-stream Dense Network," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 695-704, doi: 10.1109/CVPR.2018.00079.
9. Jaesung Rim, Haeyun Lee, Jucheol Won, and Sunghyun Cho. Real-world blur dataset for learning and benchmarking deblurring algorithms. In *ECCV*, 2020.
10. S. Nah, T. H. Kim and K. M. Lee, "Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 257-265, doi: 10.1109/CVPR.2017.35.
11. Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks. *Advances in Neural Information Processing Systems*, 2014, 27: 2672-2680.
12. O. Kupyn, V. Budzan, M. Mykhailych, D. Mishkin and J. Matas, "DeblurGAN: Blind Motion Deblurring Using Conditional Adversarial Networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8183-8192, doi: 10.1109/CVPR.2018.00854.
13. K. Zhang et al., "Deblurring by Realistic Blurring," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 2734-2743, doi: 10.1109/CVPR42600.2020.00281.
14. F. Zhuang et al., "A Comprehensive Survey on Transfer Learning," in *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43-76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.
15. R. Ribani and M. Marengoni, "A Survey of Transfer Learning for Convolutional Neural Networks," *2019 32nd SIBGRAP Conference on Graphics, Patterns and Images Tutorials (SIBGRAP-T)*, 2019, pp. 47-57, doi: 10.1109/SIBGRAP-T.2019.00010.
16. Y. Tai et al., "Nonlinear Camera Response Functions and Image Deblurring: Theoretical Analysis and Practice," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 10, pp. 2498-2512, Oct. 2013, doi: 10.1109/TPAMI.2013.40.
17. J. Wu and X. Di, "Integrating Neural Networks Into the Blind Deblurring Framework to Compete With the End-to-End Learning-Based Methods," in *IEEE Transactions on Image Processing*, vol. 29, pp. 6841-6851, 2020, doi: 10.1109/TIP.2020.2994413.
18. J. Pan, D. Sun, H. Pfister and M. Yang, "Blind Image Deblurring Using Dark Channel Prior," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1628-1636, doi: 10.1109/CVPR.2016.180.
19. L. Xu, S. Zheng and J. Jia, "Unnatural L0 Sparse Representation for Natural Image Deblurring," *2013 IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1107-1114, doi: 10.1109/CVPR.2013.147.
20. T. H. Kim and K. M. Lee, "Segmentation-Free Dynamic Scene Deblurring," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2766-2773, doi: 10.1109/CVPR.2014.348.
21. J. Sun, Wenfei Cao, Zongben Xu and J. Ponce, "Learning a convolutional neural network for non-uniform motion blur removal," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 769-777, doi: 10.1109/CVPR.2015.7298677.
22. D. Gong et al., "From Motion Blur to Motion Flow: A Deep Learning Solution for Removing Heterogeneous Motion Blur," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 3806-3815, doi: 10.1109/CVPR.2017.405.
23. X. Tao, H. Gao, X. Shen, J. Wang and J. Jia, "Scale-Recurrent Network for Deep Image Deblurring," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8174-8182, doi: 10.1109/CVPR.2018.00853.
24. H. Zhang, Y. Dai, H. Li and P. Koniusz, "Deep Stacked Hierarchical Multi-Patch Network for Image Deblurring," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 5971-5979, doi: 10.1109/CVPR.2019.00613.
25. H. Sun, M. H. Ang and D. Rus, "A Convolutional Network for Joint Deraining and Dehazing from A Single Image for Autonomous Driving in Rain," *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 962-969, doi: 10.1109/IROS40897.2019.8967644.
26. M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein generative adversarial networks. *34th International Conference on Machine Learning, ICML 2017*, v 1, p 298-321, 2017
27. I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville. Improved Training of Wasserstein GANs *Advances in Neural Information Processing Systems*, v 2017-December, p 5768-5778, 2017
28. D. Huang, L. Kang, Y. F. Wang and C. Lin, "Self-Learning Based Image Decomposition With Applications to Single Image Denoising," in *IEEE Transactions on Multimedia*, vol. 16, no. 1, pp. 83-93, Jan. 2014, doi: 10.1109/TMM.2013.2284759.
29. Y. Li, R. T. Tan, X. Guo, J. Lu, and M. S. Brown. "Rain streak removal using layer priors," In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2736-2744, June 2016
30. Y.-L. Chen and C.-T. Hsu, "A generalized low-rank appearance model for spatio-temporally correlated rain streaks," in *IEEE ICCV*, 2013, pp. 1968-1975.
31. H. Sun, M. H. Ang and D. Rus, "A Convolutional Network for Joint Deraining and Dehazing from A Single Image for Autonomous Driving in Rain," *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 962-969, doi: 10.1109/IROS40897.2019.8967644.
32. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
33. J. Johnson, A. Alahi, and L. Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, 2016.
34. P. Isola, J. Zhu, T. Zhou and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial

- Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 5967-5976, doi: 10.1109/CVPR.2017.632.
35. Tae Hyun Kim, Byeongjoo Ahn, and Kyoung Mu Lee. Dynamic scene deblurring. In *ICCV*, 2013
 36. Z. Hu, S. Cho, J. Wang and M. Yang, "Deblurring Low-Light Images with Light Streaks," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3382-3389, doi: 10.1109/CVPR.2014.432.
 37. H. Zhang, V. Sindagi and V. M. Patel, "Image De-raining Using a Conditional Generative Adversarial Network," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 11, pp. 3943-3956, Nov. 2020, doi: 10.1109/TCSVT.2019.2920407.
 38. W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo and S. Yan, "Deep Joint Rain Detection and Removal from a Single Image," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1685-1694, doi: 10.1109/CVPR.2017.183.
 39. X. Fu, J. Huang, D. Zeng, Y. Huang, X. Ding and J. Paisley, "Removing Rain from Single Images via a Deep Detail Network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1715-1723, doi: 10.1109/CVPR.2017.186.
 40. X. Fu, J. Huang, X. Ding, Y. Liao and J. Paisley, "Clearing the Skies: A Deep Network Architecture for Single-Image Rain Removal," in *IEEE Transactions on Image Processing*, vol. 26, no. 6, pp. 2944-2956, June 2017, doi: 10.1109/TIP.2017.2691802.
 41. W. Wei, D. Meng, Q. Zhao, Z. Xu and Y. Wu, "Semi-Supervised Transfer Learning for Image Rain Removal," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 3872-3881, doi: 10.1109/CVPR.2019.00400.