IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 15, 2019, accepted December 1, 2019, date of publication December 10, 2019, date of current version December 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2958607

Single Image Dehazing via NIN-DehazeNet

KANGLE YUAN, JIANGUO WEI, WENHUAN LU, AND NAIXUE XIONG, (Senior Member, IEEE)
College of Intelligence and Computing, Tianjin University, Tianjin 300350, China

Corresponding author: Wenhuan Lu (wenhuan@tju.edu.cn)

This work was supported in part by National key Research and Development Program under Grant 2018YFC0806802, in part by the NSFC key Project of Tianjin under Grant 16JCZDJC35400, and in part by the National Science Foundation of China under Grant 61876131 and Grant 61573254.

ABSTRACT Single image dehazing has always been a challenging problem in the field of computer vision. Traditional image defogging methods use manual features. With the development of artificial intelligence, the defogging method based on deep learning has developed rapidly. In this paper, we propose a novel image defogging approach called NIN-DehazeNet for single image. This method estimates the transmission map by NIN-DehazeNet combining Network-in-Network with MSCNN(Single Image Dehazing via Multi-Scale Convolutional Neural Networks). In the test stage, we estimate the transmission map of the input hazy image based on the trained model, and then generate the dehazed image using the estimated atmospheric light and computed transmission map. Extensive experiments have shown that the proposed algorithm overperformance traditional methods.

INDEX TERMS Single image dehazing, manual features, deep learning, NIN-DehazeNet, Network-in-Network, multi-scale convolutional neural networks,atmospheric scattering model.

I. INTRODUCTION

Image dehazing is a challenging problem in the field of computer vision. The purpose of Image dehazing is to recover a clear image from one single noisy frame caused by haze, fog or smoke. The dehazing algorithms have thus been widely considered, as a challenging instance of (ill-posed) image restoration [1] and enhancement [2]. Similar to other problems like image denoising and super-resolution [25], [26], earlier dehazing work [27]–[30] assumed the availability of multiple images from the same scene. However, the haze removal from one single image has now gained the dominant popularity, since it is more practical for realistic settings [5]. This paper focused on the problem of single image dehazing.

With the development of the defogging field, many classic defogging algorithms have been proposed [3], [5]–[10] in recent years. Tan [3] proposed a method for defogging based on statistical rule which based on the contrast between a clear image and a foggy image. It successfully dehazed a single image by using image contrast [4]. However, this is an image enhancement method and does not restore the radiation of the object (or scene) from the perspective of the imaging mechanism. Thus, this method will cause the restored color to be over saturated. Fattal [5] assumed that the target radiation and the medium transmission have local statistical irrelevance. Then they estimated a scene by independent component analysis (ICA), and finally they obtained the fog-free image. However, this method is limited by the assumption of statistical independence. In addition, the method is based on color information and is not applicable to grayscale images and dense fog weather. Meanwhile, a dehazing algorithm which used dark channel prior(DCP) was proposed by He *et al.* [6]. It plays an important role in the field of defogging. The method is mainly based on the principle of dark channels: some pixels always have at least one color channel in most non-sky partial areas which has a very low value. In other words, the minimum value of light intensity is in this area.

All of the above are manual features [3], [5], [6]. Image dehazing technology is also gradually developed along the manual features. However, the human visual system does not need to rely on these explicit feature transformations to estimate the concentration of the fog and the depth of the scene.

DehazeNet [7] (an end-to-end system for single image haze removal) is a specially designed deep convolutional network that uses deep learning to learn the features of smog and solves the difficulties and obstacles of manual feature design. It can achieve a more positive effect of dehazing and overcome some of the shortcomings of [6]. Moreover, as the dataset increases in size, the performance of dehazing will also increase. The idea of MSCNN [8] is very similar to

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague.

that of DehazeNet [7], but has a multi-scale convolutional network.

Dehazing by a CNN is needed to estimate the global atmospheric brightness, but the key issue in dehazing is estimating the transmission map. AOD-Net [9] is not a dehazing method that estimates transmission, but its principle is similar to that of MSCNN [8]. DCPDN [10] can simultaneously study transmission maps and atmospheric illumination. The network enables end-to-end learning by embedding atmospheric scattering models directly into the network, ensuring that the proposed method strictly follows the physical drive scattering model for dehazing.

II. RELATED WORK

The formulation of a hazy image can be modeled as:

$$I(\mathrm{x}) = J(x)t(x) + A(1 - t(x)) \tag{1}$$

where $I(x)$ is observed hazy image, $J(x)$ is the scene radiance to be recovered, $A(x)$ denotes the global atmospheric light and $t(x)$ is the transmission matrix defined as:

$$t(x) = e^{-\beta d(x)} \tag{2}$$

where β is the scattering coefficient of the atmosphere and $d(x)$ is the depth map .

Early approaches for dehazing often require multiple images to deal with this problem [29], [30], [27], [33], [34], [28]. These methods [29], [30], [27], [33], [34], [28] need to assume that these images are multiple images from the same scene. However, in practical applications, many images are unique to a given scene.

Dark channel prior(DCP) [6] plays an important role in the field of defogging. Many methods about defogging are based on DCP [35]–[38]. But, these methods about dehazing computationally expensive [39]–[41].

MSCNN [8] combines convolutional networks with multi-scale features. So, compared with the ordinary convolutional network, MSCNN [8] could extract more features of image and get better effect about defogging. Therefore, in this paper, we improve the MSCNN [8] to get better effect about defogging.

Following,we will discuss the method of using deep learning to estimate the transmission map. In this paper, we design a improved network about dehazing to estimate the transmission map. We combining Network-in-Network [11] with MSCNN(Single Image Dehazing via Multi-Scale Convolutional Neural Networks) [8]. In other words, we use the mlpconv layer which was proposed in 2014 to instead of the normal linear convolution layer. With this method, a higher quality map about dehazing can be obtained.

The contributions of this work are summarized below:

1) A method about dehazing for estimating a transmission map by combining Network-in-Network with MSCNN(Single Image Dehazing via Multi-Scale Convolutional Neural Networks) [8] is proposed.
2) The method for estimating a transmission map by combining Network-in-Network with MSCNN(Single Image Dehazing via Multi-Scale Convolutional Neural Networks) [8]. So we have adopted several different feature extraction methods for MSCNN [8] and compared them with the method we proposed in this paper. Finally, we showed that the MSCNN [8] with Mlpconv layer [11] to extraction features is better than the MSCNN [8] with other convolutional layer to extraction features.
3) We analyzed some existing methods of dehazing by 2019 and compared them with the method we proposed. Finally, we proved that the algorithm about defogging we proposed can achieve better defogging effect.

Next, we will introduce the algorithm we proposed in detail.

III. NIN-DEHAZENET FOR TRANSMISSION MAPS

In this section, we propose a novel image defogging approach that estimates the transmission map by combining Network-in-Network with MSCNN(Single Image Dehazing via Multi-Scale Convolutional Neural Networks). We call this method NIN-DehazeNet.

For each scene, we propose to estimate the scene transmission map t(x) based on a multi-scale CNN [8]and Mlpconv layer [11]. The coarse structure [8] of the scene transmission map for each image is obtained from the coarse-scale network and we add the Mlpconv layer [11] to the coarse-scale network [8]. And then refined by the fine-scale network [8]. We also add the Mlpconv layer [11] to the fine-scale network [8]. Both coarse and fine scale networks are applied to the original input hazy image. And, the output of the coarse network is passed to the fine network as additional information. By Mlpconv layer [11], almost all fog-related features could be extracted, such as dark channels, hue disparity, and color attenuation.

The main steps of the proposed algorithm are shown in Figure 1.

A. NIN-DEHAZENET

A single foggy image was given, our goal is to obtain a fog-free image by estimating the transmission map. We refer to the structure of Network-in-Network and add the mlpconv layer to MSCNN(Single Image Dehazing via Multi-Scale Convolutional Neural Networks) to instead of the ordinary linear convolution layer. We call this method NIN-DehazeNet. By NIN-DehazeNet, we could get a more accurate transmission map. And then substitute it into the atmospheric scattering model to obtain a image about dehazing. The overall architecture of the model is shown in the Figure 2. The details of the procedure are described below.

1)Mlpconv layer:''Network In Network''(NIN) [11] is a novel deep network structure to enhance model discriminability for local patches within the receptive field. The conventional convolutional layer uses linear filters followed by a nonlinear activation function to scan the input. But in NIN [11], the rdinary linear convolution layer is replaced with

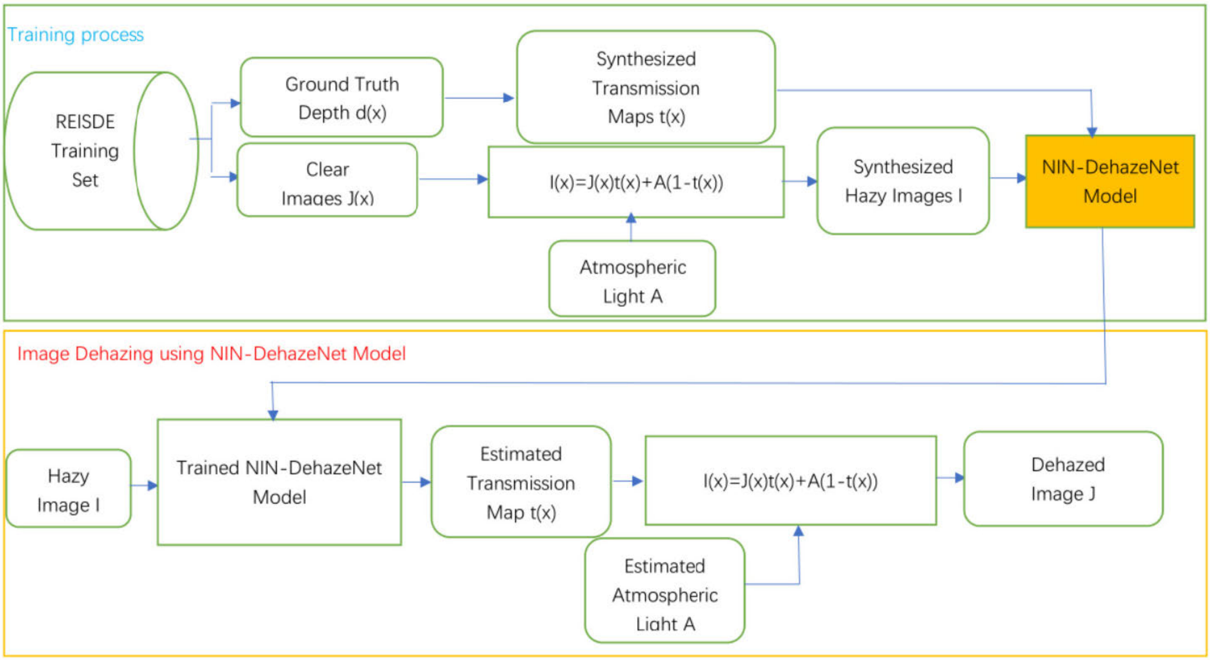


FIGURE 1. Main steps of the proposed single-image dehazing algorithm. We estimate the transmission map of the input hazy image based on the trained model, and then generate the dehazed image using the estimated atmospheric light and computed transmission map.

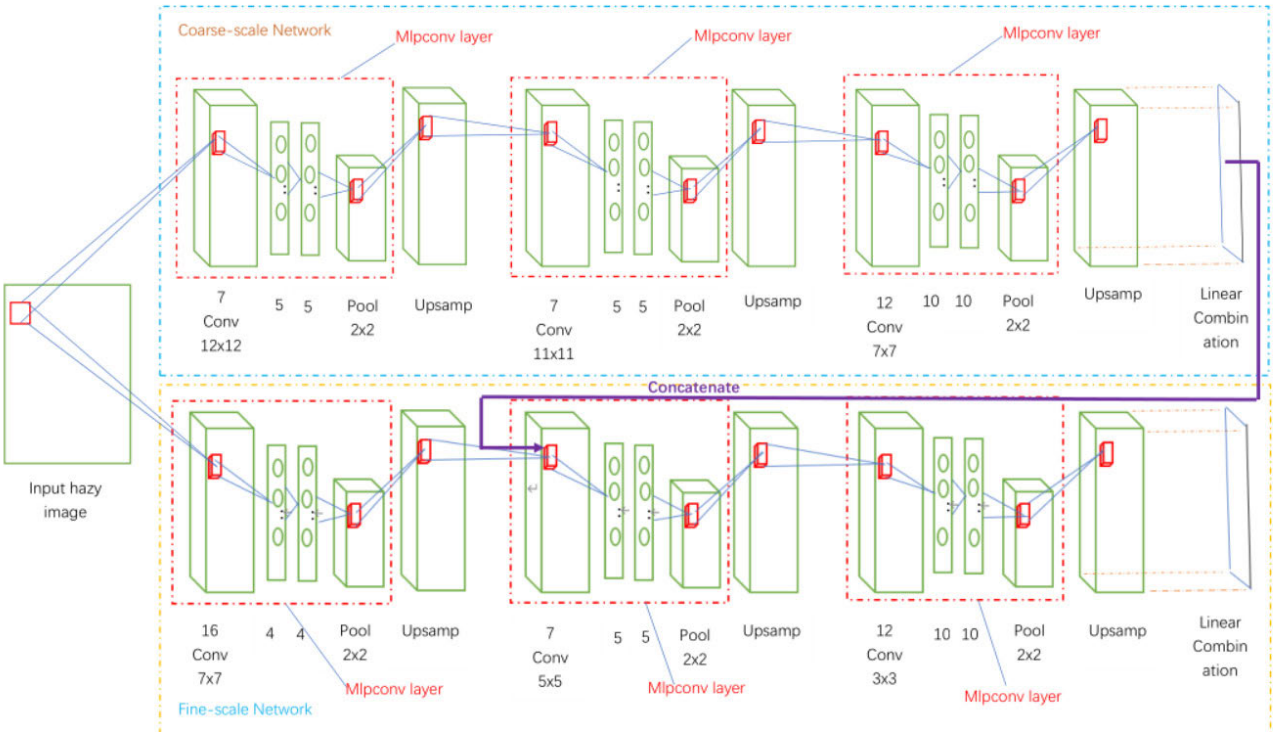


FIGURE 2. The overall architecture of the model. The network named NIN -DEHAZENET includes Coarse-scale and Fine-scale network. And mlpconv layers (the purple dashed rectangle) are put in MSCNN. Given a hazy image, the coarse- scale network (the black dashed rectangle) predicts a holistic transmission map and feeds it to the fine-scale network (the red dashed rectangle) in order to generate a refined transmission map.

a "micro network" structure which is a general nonlinear function approximator. And it was called mlpconv layer. The resulting structure of mlpconv layer is compared with CNN in Figure 3.

Both the linear convolutional layer and the mlpconv layer map the local receptive field to an output feature vector. The mlpconv maps the input local patch to the output feature vector with a multilayer perceptron(MLP) [11] consisting

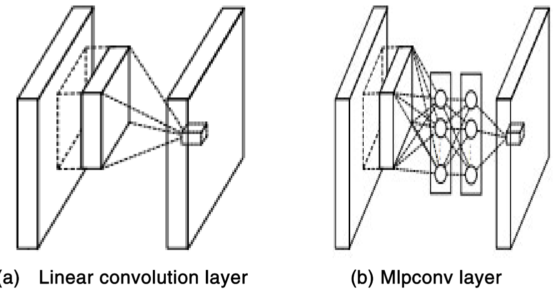


FIGURE 3. Comparison of linear convolution layer and mlpconv layer.

of multiple fully connected layers with nonlinear activation functions. The MLP is shared among all local receptive fields. The feature maps are obtained by sliding the MLP over the input in a similar manner as CNN and are then fed into the next layer.

DehazeNet [7] uses the maxout method to extracts features. The maxout unit is a simple feedforward nonlinear activation function used in a multi-layer perceptron or CNN framework. After feature extraction by maxout, almost all fog-related features (dark channels, hue disparity, and color attenuation) can be extracted.

The maxout layers in the maxout network [12] performs max pooling across multiple affine feature maps. The feature maps of maxout layers are calculated as follows:

$$f_{i,j,k} = \max_{m}(w_{k_m}^{T} x_{i,j}) \tag{3}$$

Here (i, j) is the pixel index in the feature map, $x_{i,j}$ stands for the input patch centered at location(i, j), and k is used to index the channels of the feature map.

However, maxout network [12] imposes the prior that instances of a latent concept lie within a convex set in the input space, which does not necessarily hold. It would be necessary to employ a more general function approximator when the distributions of the latent concepts are more complex. Mlpconv layer [11] differs from maxout layer in that the convex function approximator is replaced by a universal function approximator, which has greater capability in modeling various distributions of latent concepts.

The calculation performed by mlpconv layer [11] is shown as follows:

$$\begin{aligned} f_{i,j,k_1}^{1} &= \max(w_{k_1}^{1} x_{i,j} + b_{k_1}, 0) \\ &\vdots \\ f_{i,j,k_n}^{n} &= \max(w_{k_n}^{n} x_{i,j} + b_{k_n}, 0) \end{aligned} \tag{4}$$

Here n is the number of layers in the multilayer perceptron [11]. Rectified linear unit is used as the activation function in the multilayer perceptron [11]. From cross channel (cross feature map) pooling point of view, equation 4 is equivalent to cascaded cross channel parametric pooling on a normal convolution layer [11]. Each pooling layer performs weighted linear recombination on the input feature maps [11], which then go through a rectifier linear unit. The cross channel pooled feature maps are cross channel pooled again and again in the next layers. This cascaded cross channel parameteric pooling structure allows complex and learnable interactions of cross channel information. And the cross channel parametric pooling layer is also equivalent to a convolution layer with 1×1 convolution kernel.

2)Coarse-scale and Fine-scale network: This network is designed to estimate the transmission map in MSCNN [8] as that improve the prediction results. For this reason, we adopt the Coarse-scale and Fine-scale network in the subsequent network structure. The task of the coarse-scale network is to predict a holistic transmission map of the scene. Then we make refinements using a fine-scale network. The network mainly includes max-pooling, an up-sampling [31] layer and a linear combination layer. In this final layer, the feature channel from the last upsampling layer is integrated by linear combination,and the final output is obtained by the sigmoid activation function.

Max-pooling: a maximum pooling layer is used after each convolutional layer.

Up-sampling: an up-sampling layer [31] is used after each pooling layer to ensure a transmission map of the output and an input foggy pattern equal in size.

Linear combination: In this final layer, the feature channel from the last up-sampling layer is integrated by linear combination [31] and the final output is obtained by the sigmoid activation function.

B. TRAINING

For training the NIN-DehazeNet, we synthesize hazy images and the corresponding transmission maps based on REalistic Single Image DEhazing (RESIDE) training set [17].

When training the neural network, it is necessary to set the learning rate to control the speed of the parameter update. The learning rate determines the amplitude of update step of the parameters. If the amplitude is too large, the parameters may be at an excellent value. It can move back and forth on both sides; However, if the amplitude is too small, it will greatly reduce the speed of optimization. So, in this paper we have chosen a more flexible method to set learning rate, named Adam algorithm [14]. The Adam algorithm sets of advantages for two stochastic gradient descent extensions:

The adaptive gradient algorithm (AdaGrad) [15] preserves a learning rate for each parameter to improve performance on sparse gradients (the problem of natural language and computer vision).

Root mean square propagation (RMSProp) [16] adaptively preserves the learning rate based on the mean of the nearest magnitude of the weight gradient for each parameter.

The Adam algorithm takes advantage of the AdaGrad and RMSProp algorithms. Adam not only calculates the adaptive parameter learning rate based on the first-order moment mean such as the RMSProp algorithm, but also makes full use of the gradient's second-order mean variance (uncentered variance). Therefore, it can achieve excellent results very quickly.

Learning the mapping between hazy images and corresponding transmission maps is learned by minimizing the loss function between the predicted transmission map and the corresponding ground truth map. And finally, a mapping between the foggy image and the transmission map is obtained. The loss function is as follows:

$$loss = \frac{1}{N}\sum_{i=1}^{N}(y_i - y(x_i))^2 \tag{5}$$

where y_i is the reconstructed transmission map and $y(x)$ is the corresponding ground truth map. Where N is the number of hazy images in the training set. We minimize the loss using the Adam algorithm. In this paper,we first train the coarse-scale network, and then use the coarse-scale output transmission maps to train the fine-scale network [8]. The training loss is used in both coarse-scale and fine-scale networks.

C. DEHAZING WITH THE NIN-DEHAZENET

In the test stage, we estimate the transmission map of the input hazy image based on the trained model, and then generate the dehazed image using the estimated atmospheric light and computed transmission map.

Atmospheric light (A) is also estimated before using the atmospheric scattering model. In this paper, we estimate the atmospheric light by selecting the darkest pixel of 0.1% in the transmission map and then choose the brightness of the corresponding position of the foggy original image.

IV. EXPERIMENTAL RESULTS

To train the NIN-DehazeNet, we generate a dataset with synthesized hazy images and their corresponding transmission maps. We randomly sample 10500 images from the REalistic Single Image DEhazing (RESIDE) training set [17] to construct the training set. The REISDE training set contains 13, 990 synthetic hazy images, generated using 1, 399 clear images from existing indoor depth datasets NYU2 [18] and Middlebury stereo [19], [32], [33].

A. EXPERIMENTAL SETTING

In this paper, we did two sets of comparative experiments. And we all use 100 images(225∗225) in testing at two experiments. The computer configuration used in this experiment is Core i7-7700CPU@2. 80GHz, 8GB memory. We did this through tensorflow(GPU).

This experiment is mainly compared in the following two directions:

1) Because different feature extraction methods has an effect on the result, we compare the defogged images obtained when different feature extraction methods is used. In this paper,we use MSCNN,improved MSCNN which extracting features by maxout unit and NIN-DehazeNet to compare.

TABLE 1. **Average PSNR and SSIM result.**

Metrics	MSCNN	MSCNN(maxout unit)	NIN-DehazeNet
PSNR	18.2213	18.3668	**18.4827**
SSIM	0.8331	0.8559	**0.8623**

2) The optimal dehazing method obtained after the comparison above was compared with other advanced dehazing methods.

B. EVALUATION INDEX

To evaluate the quality of defogging results, we must first evaluate the indicators. This paper mainly uses peak signal to noise ratio (PSNR) [20] and structural similarity (SSIM) [21] to evaluate the images after dehazing.

PSNR is a full-reference image quality evaluation index. The unit of PSNR is dB, and the larger the value, the smaller the distortion.

SSIM (structural similarity) is also a full-reference image quality evaluation index, and it measures image similarity from three aspects: brightness, contrast, and structure. The SSIM value range is [0, 1], and the larger the value, the smaller the image distortion.

C. QUANTITATIVE EVALUATION ON DATASET

The testset used for the deep learning based de-fogging method is mainly a synthetic dehazing image with known basic facts. So we take the 100 images from the related videos and use them to synthesize foggy images as the TestSet .

1) EXPERIMENT A

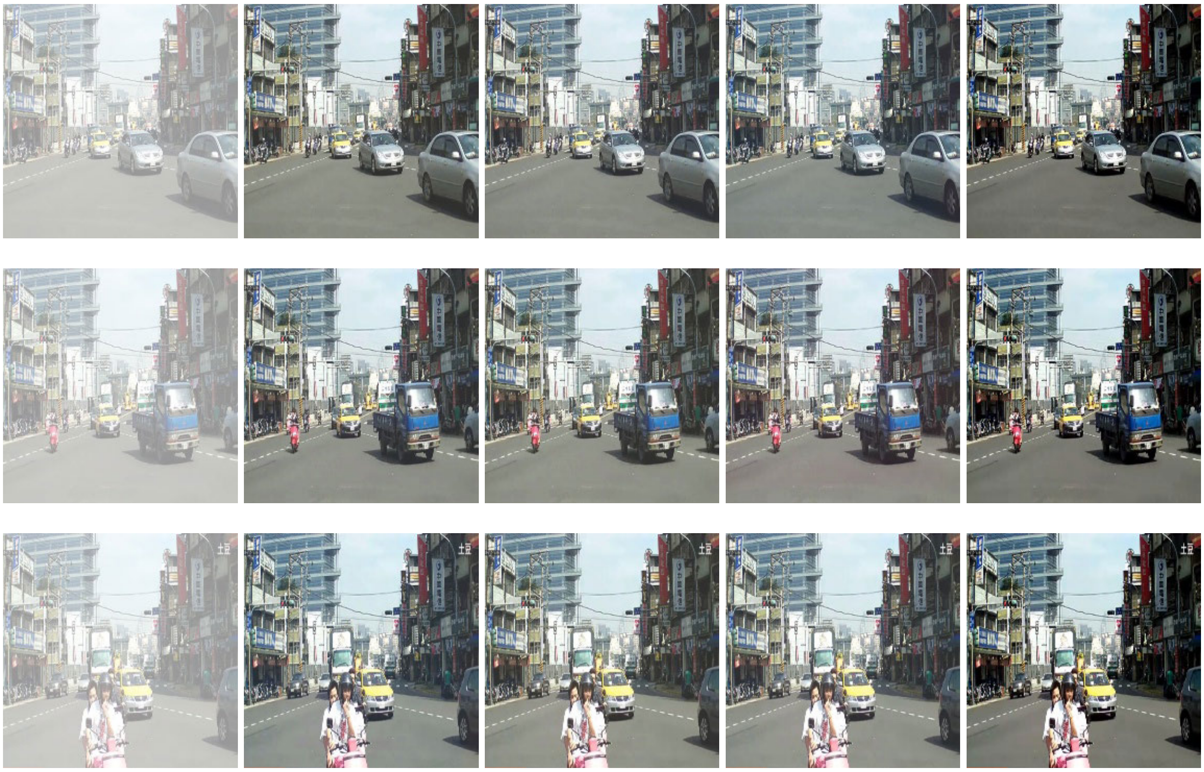
Because different feature extraction methods has an effect on the result, we compare the defogged images obtained when different feature extraction methods is used. In this paper,we use MSCNN,improved MSCNN which extracting features by maxout unit and NIN-DehazeNet to compare.

The conventional convolutional layer uses linear filters followed by a nonlinear activation function to scan the input. The maxout unit is a simple feedforward nonlinear activation function used in a multi-layer perceptron or CNN framework. After feature extraction by maxout, almost all fog-related features (dark channels, hue disparity, and color attenuation) can be extracted. However, maxout network impoes the prior that instances of a latent concept lie within a convex set in the input space [11], which does not necessarily hold.

Table 1 displays the average PSNR and SSIM results on TestSets .

According to Table 1, it can be concluded that when different feature extraction methods is used, NIN-DehazeNet could get the best defogging effect .

When the linear convolutional layer,maxout unit and the mlpconv layer map were used, the effect diagrams are shown in Figure 4. Three pictures are respectively



(a) Input image (b)MSCNN (c)MSCNN(maxout unit) (d)NIN-DehazeNet (e)Clear image

FIGURE 4. Three images were selected from 100 images, and we obtained a fog-free image with MSCNN,improved MSCNN which extracting features by maxout unit and NIN-DehazeNet.

selected from 100 images, the foggy image is shown in Figure 4(a). We note that the dehazed results by MSCNN in Figure 4(b), the dehazed results by MSCNN(maxout unit) in Figure 4(c), and dehazed results by the proposed algorithm in Figure 4(d). And fog-free original images are shown in Figure 4(e).

2) EXPERIMENT B

We compare the proposed algorithm with the state-of-the-art dehazing methods [6]–[10] using the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) metrics.

The Dark Channel Prior (DCP) algorithm, which was proposed by He *et al.* [6], plays an important role in image defogging. In this paper, we use the DCP [6] dehazing algorithm modified by the steering filter [22]. And the improved algorithm can achieve finer transmission.

DCPDN [10] dehazing algorithms proposed in CVPR in 2018. And it is one of the best algorithms for defogging at present.

The experiment mainly uses DCP [6] which was improved, Dehaze-Net [7], MSCNN [8], AOD-Net [9] and DCPDN [10] compared with the dehazing method we proposed in this paper.

In order to ensure the requirements of the qualitative and quantitative, the 100 images which were used in the above tests would still be used at this time.

The effect diagrams are shown in Figure 5. Three identical pictures are respectively selected from 100 images. We note that the dehazed results by [6] in Figure 5(a), the dehazed results by [7] in Figure 5(b), the dehazed results by [9] in Figure 5(c),the dehazed results by [8] in Figure 5(d),the dehazed results by [10] in Figure 5(e),and dehazed results by the proposed algorithm in Figure 5(f).

TABLE 2. Average PSNR and SSIM result.

Metric s	DCP	Dehaz e-Net	AOD-Net	MSCN N	DCPD N	**Ours**
PSNR	17.81 74	18.03 72	18.25 32	18.22 13	18.46 28	**18.48 27**
SSIM	0.792 2	0.821 5	0.832 6	0.833 1	0.853 2	**0.862 3**

Tables 2 display the average PSNR and SSIM results on TestSets .

According to Table 2, it can be concluded that the method we proposed has higher PSNR performance than others and it gains even greater SSIM advantages over all competitors.

D. RUNING TIME COMPARISON

We select 50 images from TestSet for all models to run, on the same machine(Intel(R) Core(TM) i7-7700 CPU@2. 80GHz and 8GB memory), with GPU acceler-ation.

The per-image average running time of all models are shown in Table 3. Despite the speed of the algorithm is not the fastest, it could be used for real-time dehazing, and it


(a)DCP (b)Dehaze-Net (c)AOD-Net (d)MSCNN (e)DCPDN (f)Ours

FIGURE 5. We use the evaluation index about dehazing to get the evaluation of the effect about dehazing.

TABLE 3. Comparison of average model running time (in seconds).

	DCP	Dehaze-Net	AOD-Net	MSCNN	DCPDN	**Ours**
Running time	0.92	0.051	0.032	0.037	0.076	0.039

has the best effect about dehazing. Therefore, considering the algorithm comprehensively, it could be applied to video defogging.

According to Table 3, it can be concluded that the method we propose is fast, and it can be applied in real scenarios of video during foggy weather [23].

Generative Adversarial Networks [24] (GAN) is a deep learning model and one of the most promising methods for unsupervised learning in complex distribution in recent years. In the future, we are going to combine NIN-DehazeNet (the method about defogging we proposed) with GAN to get a algorithm with better effect about defogging.

V. CONCLUSION

In this paper we propose NIN-DehazNet, which estimates the transmission map by combining Network-in-Network with MSCNN(Single Image Dehazing via Multi-Scale Convolutional Neural Networks). And we compare it with some methods which could get good effect about dehazing. The experiments show that the method we propose perform best and it has the best effect about dehazing than other dehazing algorithms . The PSNR increased as well as the SSIM improved. The experiments also show that our method is fast, and therefore it can be applied in real scenarios of video during foggy weather. The work can facilitate the automatic monitoring in Smart City applications.

REFERENCES

[1] D. Guo, P. Liu, Y. She, H. Yang, and D. Liu, “Ultrasonic imaging contrast enhancement using modified dehaze image model,” *Electron. Lett.*, vol. 49, no. 19, p. 1209, 2013.
[2] J. Li, Z. Hong, Y. Ding, and M. Sun, “Single image dehazing using the change of detail prior,” *Neurocomputing*, vol. 156, no. 26, pp. 1–11, 2015.
[3] R. T. Tan, “Visibility in bad weather from a single image,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.
[4] J. P. Oakley and B. L. Satherley, “Improving image quality in poor visibility conditions using a physical model for contrast degradation,” *IEEE Trans. Image Process.*, vol. 7, no. 2, pp. 167–179, Feb. 1998.
[5] R. Fattal, “Single image dehazing,” *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, 2008.
[6] K. He, J. Sun, and X. Tang, “Single image haze removal using dark channel prior,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
[7] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, “DehazeNet: An end-to-end system for single image haze removal,” *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016.
[8] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, “Single image dehazing via multi-scale convolutional neural networks,” in *Proc. 14th Eur. Conf. Comput. Vis. (ECCV)*, in Lecture Notes in Computer Science, vol. 9906. Amsterdam, The Netherlands: Springer, Oct. 2016.
[9] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “Aod-net: All-in-one dehazing network,” in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2017, pp. 4770–4778.
[10] Z. He and V. M. Patel, “Densely connected pyramid dehazing network,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3194–3203.
[11] M. Lin, Q. Chen, and S. Yan, “Network in network,” 2013, *arXiv:1312.4400*.
[12] I. J. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, and Y. Bengio, “Maxout networks,” in *Proc. 30th Int. Conf. Mach. Learn. (ICML)*, Atlanta, GA, USA, Jun. 2013.

[13] R. Herzog, E. Eisemann, K. Myszkowski, and H.-P. Seidel, "Spatio-temporal upsampling on the GPU," in *Proc. Symp. Interact. 3D Graph. (SI3D)*, Washington, DC, USA, Feb. 2010, pp. 91–98.

[14] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015.

[15] F. E. Curtis and X. Que, "An adaptive gradient sampling algorithm for non-smooth optimization," *Optim. Method. Softw.*, vol. 28, no. 6, pp. 1302–1324, 2013.

[16] S. Roy, N. Das, M. Kundu, and M. Nasipuri, "Handwritten isolated Bangla compound character recognition: A new benchmark using a novel deep learning approach," *Pattern Recogn. Lett.*, vol. 90, no. 4, pp. 15–21, 2017.

[17] Y. Zhang, D. Li, and G. Sharma, "HazeRD: An outdoor scene dataset and benchmark for single image dehazing," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2018, pp. 3205–3209.

[18] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Proc. Comput. Vis. (ECCV)*, 2012, pp. 746–760.

[19] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 2003, p. 1.

[20] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[21] H. R. Sheikh, A. C. Bovik, and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2117–2128, Dec. 2005.

[22] M. Barbaro, P. Y. Burgi, A. Mortara, P. Nussbaum, and F. Heitger, "A 100 × 100 pixel silicon retina for gradient extraction with steering filter capabilities and temporal output coding," *IEEE J. Solid-State Circuits*, vol. 37, no. 2, pp. 160–172, Feb. 2002.

[23] W. Ren, J. Zhang, X. Xu, L. Ma, X. Cao, G. Meng, and W. Liu, "Deep video dehazing with semantic segmentation," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1895–1908, Apr. 2019.

[24] J. Li, S. Liu, H. He, and L. Li, "A novel framework for gear safety factor prediction," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 1998–2007. 2019.

[25] M. Tico, "Multi-frame image denoising and stabilization," In *Proc. 16th Eur. Signal Process. Conf.*, 2008, pp. 1–4.

[26] X. Li, Y. Hu, X. Gao, D. Tao, and B. Ning, "A multi-frame image super-resolution method," *Signal Process.*, vol. 90, no. 2, pp. 405–414, 2010.

[27] S. G. Narasimhan and S. K. Nayar, "Contrast restoration of weather degraded images," *IEEE Trans. Pattern Anal. Mach. Learn.*, vol. 25, no. 6, pp. 713–724, Jun. 2003.

[28] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, "Instant dehazing of images using polarization," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, 2001, p. 1.

[29] T. Treibitz and Y. Y. Schechner, "Polarization: Beneficial for visibility enhancement?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 525–532.

[30] J. Kopf, B. Neubert, B. Chen, M. Cohen, D. Cohen-Or, O. Deussen, M. Uyttendaele, and D. Lischinski, "Deep photo: Model-based photograph enhancement and viewing," *ACM Trans. Graph.*, vol. 27, no. 5, 2008, Art. no. 116.

[31] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf, "Learning to deblur," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 7, pp. 1439–1451, Sep. 2016.

[32] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.*, vol. 47, nos. 1–3, pp. 7–42, 2002.

[33] H. Hirschmüller and D. Scharstein, "Evaluation of cost functions for stereo matching," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Minneapolis, MN, USA, Jun. 2007.

[34] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," in *Proc. CVPR*, 2000, pp. 1598–1605.

[35] L. Kratz and K. Nishino, "Factorizing scene albedo and depth from a single foggy image," in *Proc. ICCV*, 2009, pp. 1701–1708.

[36] J. P. Tarel, N. Hautiere, L. Caraffa, A. Cord, H. Halmaoui, and D. Gruyer, "Vision enhancement in homogeneous and heterogeneous fog," *Intell. Transp. Syst. Mag.*, vol. 4, no. 2, pp. 6–20, 2012.

[37] K. Nishino, L. Kratz, and S. Lombardi, "Bayesian defogging," *Int. J. Comput. Vis.*, vol. 98, no. 3, pp. 263–278, Jul. 2012.

[38] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *Proc. ICCV*, 2013, pp. 617–624.

[39] J. P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in *Proc. ICCV*, 2009, pp. 2201–2208.

[40] K. B. Gibson, D. T. Vo, and T. Q. Nguyen, "An investigation of dehazing effects on image and video coding," *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 662–673, Feb. 2012.

[41] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 397–1409, Jun. 2013.

KANGLE YUAN was born in 1994. He received the bachelor's degree in vehicle engineering from the Shandong University of Technology, in 2017. He is currently pursuing the master's degree in software engineering with Tianjin University. His research interests include speech signal processing and computer graphics processing associate with it.

JIANGUO WEI received the master's degree from Jilin University and Tianjin University, and the Ph.D. degree from the University of North Korea and the University of Science and Technology, in 2007. He is currently with the National Demonstration Software College, Tianjin University. He is also a Professor, the Associate Dean, and the Dean of the School of Computer Science, Qinghai University for Nationalities (job-to-the-job support) and a Distinguished Visiting Professor with the Hokuriku University of Science and Technology, Japan. He then worked as a Postdoctoral Researcher with the French National Research Center CNRS/Telecom and Japan JAIST. In 2010, he was an Associate Researcher and a Professor with Tianjin University. He is also a Distinguished Professor of the University of Science and Technology of Hokuriku, the Associate Dean of the School of Software, Tianjin University, and the Dean of the School of Computer Science, Qinghai University for Nationalities (job). In recent years, he has devoted himself to research in computer basic education, software engineering, artificial intelligence, and big data. He is also the Director of the ACM China Computer Education Committee and the Secretary General of Computer Basic Education Branch of Tianjin Computer Society. At present, he is responsible for the National 973 Sub-project 1 project, the National Fund Project, and the Tianjin Key Fund. At the same time, he participated in the science and technology support project of the Ministry of Science and Technology and the national fund key projects. He has published more than 60 academic articles in the past five years. He is also a member of ACM and ISCA, a Senior Member of CCF, and a CCF Human–Computer Interaction Committee Member.

WENHUAN LU received the Ph.D. degree from the Japan Advanced Institute of Science and Technology (JAIST), Japan. She is currently an Associate Professor with the School of Computer Software, Tianjin University, China. She has also worked as a Postdoctoral Researcher with the JAIST. Afterwards, she joined the School of Computer Software, Tianjin University. She is working on knowledge modeling and knowledge representation with semantic technology applied to knowledge-based system. She is also the Principle Investigator for several national grants in China, including the National High-Technology Research and Development Program of China (863 Program), the National Natural Science Foundation of China, and the Scientific Research Foundation from Ministry of Education of China. She has also (co)authored more than 20 academic articles. She received the Best Paper Award in IJCSS, in 2011.

NAIXUE XIONG (S'05–M'07–SM'12) received the Ph.D. degree in sensor system engineering from Wuhan University, in 2007, and the Ph.D. degree in dependable sensor networks from the Japan Advanced Institute of Science and Technology, in 2008.

Before, he attended Tianjin University, he worked with Northeastern State University, Georgia State University, Wentworth Technology Institution, and Colorado Technical University (a Full Professor about five years) about ten years. He is currently a Professor with the College of Intelligence and Computing, Tianjin University, China, in 2017. He has published over 300 international journal articles and over 100 international conference papers. Some of his works were published in the IEEE JSAC, the IEEE or *ACM transactions*, ACM Sigcomm workshop, the IEEE INFOCOM, the ICDCS, and the IPDPS. His research interests include cloud computing, security and dependability, parallel and distributed computing, networks, and optimization theory.

Dr. Xiong is a Senior member of the IEEE Computer Society, from 2012. He has received the Best Paper Award in the 10th IEEE International Conference on High Performance Computing and Communications (HPCC-08) and the Best student Paper Award in the 28th North American Fuzzy Information Processing Society Annual Conference (NAFIPS2009). He is the Chair of Trusted Cloud Computing Task Force, the IEEE Computational Intelligence Society (CIS), http://www.cs.gsu.edu/~cscnxx/index-TF.html, and the Industry System Applications Technical Committee, http://ieee-cis.org/technical/isatc/. He has been the General Chair, the Program Chair, the Publicity Chair, the Program Committee member and Organizing Committee member of over 100 international conferences and a Reviewer of about 100 international journals, including the IEEE JSAC, the IEEE SMC (Park: A/B/C), the IEEE TRANSACTIONS ON COMMUNICATIONS, the IEEE TRANSACTIONS ON MOBILE COMPUTING, the IEEE TRANSACTION ON PARALLEL AND DISTRIBUTED SYSTEMS. He is serving as an Editor-in-Chief, an Associate Editor or a Editor member for over 10 international journals (including Associate Editor for the IEEE TRANSACTION ON SYSTEMS, MAN AND CYBERNETICS: SYSTEMS, an Associate Editor for *Information Science*, an Editor-in-Chief for *Journal of Internet Technology* (JIT), and *Journal of Parallel and Cloud Computing* (PCC)), and a Guest Editor for over ten international journals, including *Sensor* Journal, WINET and MONET.