

Enhanced densely dehazing network for single image haze removal under railway scenes

Ruhao Zhao

State Key Laboratory of Rail Traffic Control and Safety, Beijing, China

Xiaoping Ma

*State Key Laboratory of Traffic Control and Safety,
School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China*

He Zhang

Rutgers, Piscataway, New Jersey, USA

Honghui Dong

State Key Laboratory of Rail Traffic Control and Safety, Beijing, China, and

Yong Qin and Limin Jia

Beijing Jiaotong University, Beijing, China

Abstract

Purpose – This paper aims to propose an enhanced densely dehazing network to suit railway scenes' features and improve the visual quality degraded by haze and fog.

Design/methodology/approach – It is an end-to-end network based on DenseNet. The authors design enhanced dense blocks and fuse them in a pyramid pooling module for visual data's local and global features. Multiple ablation studies have been conducted to show the effects of each module proposed in this paper.

Findings – The authors have compared dehazed results on real hazy images and railway hazy images of state-of-the-art dehazing networks with the dehazed results in data quality. Finally, an object-detection test is taken to judge the edge information preservation after haze removal. All results demonstrate that the proposed dehazing network performs better under railway scenes in detail.

Originality/value – This study provides a new method for image enhancing in the railway monitoring system.

Keywords Deep learning, Haze removal, Railway intelligent transportation system, Railway monitoring system

Paper type Research paper



1. Introduction

Railway has taken an important role in the economic and social development of modern countries. Boosting railway mileages and high-level requirements of speed, transport capacity and service quality have led railway security maintenance into a crucial research field. Railway intelligent transportation system (RITS) (Li *et al.*, 2003) is broadly applied nowadays in railway daily operation. The system is integrated by electronic techniques, communication, artificial intelligence and information techniques for keep railway operation safety.

Railway monitoring system (Sacchi and Regazzoni, 2000) has been taken an important place in RITS, based on developments of visual equipment and network transmission bandwidth. The system could capture visual data by cameras, which are always used along railway tracks or on UAVs. It is designed to serve for railway operation department for 24 h to monitoring dramatic changes of facilities and environment around the railway and analyze visual data to help railway operation and dispatch department on railway planning, dispatching, interlocking, etc. It is helpful to avoid railway operation accidents, provide alarms and show real-time situations of the railway operation site.

Haze (Nayar and Narasimhan, 2002) is one kind of severe weather phenomena, which is caused by polluted atmospheric aerosol. It would cause light absorbing and scattering during the way between the original illumination source and receivers. For human visual tasks, haze would degrade visibility badly, which interrupts human judgments of the surrounding environment. For computer visual tasks, the quality of visual data captured by camera-lens in hazy days would be degraded in hue, saturation and value. Hazy visual data would lose important pixels information and is hard to be directly processed in high-level visual tasks, such as image segmentation and detection.

In recent years, haze has been constant around industrial cities, which have dense railway networks. It declines the railway monitoring system's efficiency and degraded railway drivers' visibility hardly. Railway operation apartment has to make decisions such as railway outage and speed reduction in hazy days to decline railway operation accidents. Haze has threatened railway operation safety seriously. Therefore, haze removal has been more and more important for railway monitoring systems (Cao *et al.*, 2020; Liu *et al.*, 2020; Wu *et al.*, 2020).

Because of development of computer vision and artificial intelligence techniques in recent years, railway monitoring system has taken a more intelligent role in daily railway operation with the help of deep learning, high-definition cameras and graphics processing unit (GPU) computing power. Visual data captured by the Railway monitoring system has the most various information in RITS. Researchers and practitioners have got remarkable achievements of visual objects classification, detection and segmentation techniques in automatic driving and traffic operation. However, visual data captured in foggy or hazy days usually suffers bad data quality, which affects the efficiency and precision rate of visual tasks such as image detection and recognition in railway monitoring systems badly.

Image dehazing algorithm has been a crucial challenge to enhance hazy and foggy images' qualities and restore blurred pixel information in computer vision. Multi-image haze removal and single image haze removal are two main directions in this field. Multi-image haze removal (Jiang and Lu, 2018; Li *et al.*, 2015) needs images from multiple sources to estimate climate parameters. However, equipment limitations and running time of multi-image haze removal methods are not suitable under railway scenes. Most cameras are monocular and deployed in distances of over 150 m. Multi-image haze removal methods are hard to collect data set and take experiments in the railway monitoring system. Single image haze removal (Ancuti and Ancuti., 2013; Fattal, 2008; Tan, 2008; He *et al.*, 2011) is an

ill-posed problem that estimating climate parameters for dehazing from single images is difficult. However, single image haze removal methods have achieved great results in recent works. In addition, their technical characters fit monocular railway scenes better (Plate 1).

To improve visual data's quality in hazy days under railway scenes, an enhanced densely dehazing network is proposed for single image haze removal in this paper. Because of complex monitoring scenes and railway visual data's features, the paper makes the following contributions:

First, a novel end-to-end densely dehazing network is proposed. Enhanced dense blocks (EDBs) have taken in DenseNet (Huang *et al.*, 2017) structure. EDBs can capture feature tensors of haze images more efficiently than original dense blocks. The network is an end-to-end structure without human guidance and prior information.

Second, adding a pyramid pooling module (PPM) (He *et al.*, 2015; Han *et al.*, 2017) tends to fuse localize and global image features well. Multi-scale average-pooling results of hazy image feature tensors concatenate with original feature tensors. Fusing features are helpful to avoid dehazing results with a non-uniform style.

Finally, an edge-preserving loss function is proposed for restoring more edge information of hazy images. Image dehazing methods under railway scenes need to enhance image quality and prepare data for high-level visual tasks. This loss function could increase the accuracy of foreign objects detection in the railway monitoring system.

The structure of this paper is following:

In Section 2, the research background and related works are introduced. Railway monitoring system, railway hazy images' features and single image haze removal methods have been discussed. In Section 3, an enhanced densely dehazing network and its modules are proposed. In Section 4, extensive experiments are taken on synthetic data sets and railway hazy images data set. Ablation tests, image dehazing tests and image detection tests have demonstrated our proposed network's improvements for railway hazy images dehazing. The conclusion would be discussed in Section 5.

2. Background and related works

Railway monitoring systems now is broadly applied in daily railway operations. Because of computer vision techniques such as image recognition and image detection have got great development in recent years, visual data is getting much more helpful to keep railway operation safety.

These high-level vision models' training needs high quality image data sets to match the precision ratio demands, but visual data captured by monitoring cameras in railway monitoring systems on hazy days would be degraded seriously. A typical haze image under the railway scene is shown in Plate 2. The objects, which are far from the camera lens, are blurred and difficult to be detected and recognized by human or computer vision algorithms.



Plate 1.
Haze removal sample
under railway scene



Plate 2.
Haze image under
railway scene

The data flow of visual data in railway monitoring system needs to be optimized like in [Figure 1](#).

The visual data captured in railway monitoring system should be first addressed by the low-level vision tasks, such as classification (judging whether the input data needs to be dehazed), data enhancement (expanding size of data sets to assure neural networks' efficiency) and image enhancement algorithms (image dehazing, deraining and deblurring). In this paper, a novel single image haze removal method is proposed for the railway monitoring system.

For a good description of hazy removal progress, [Nayar and Narasimhan \(2002\)](#) proposed the atmosphere scattering model, which can explain the image degradation caused by haze and fog in math. The model is formulated as:

$$I(x) = J(x)t(x) + A(x)(1 - t(x))$$

where I is the hazy image, J is the true scene radiance, A is the global atmospheric light, t is the transmission map and x is one pixel. The transmission map can be expressed as $t(x) = e^{-\beta d(x)}$, where β is the attenuation coefficient of the atmosphere and d is the scene depth. In a haze removal task, the I is given to estimate the J .

Most haze removal algorithms are based on the atmosphere scattering model. Estimating global atmospheric light, attenuation coefficient and image depth into transmission map is an ill-posed problem. During the past two decades, single image haze removal methods are broadly divided into two main groups, namely, prior-based methods and learning-based methods. Prior-based methods need prior character information to estimate climate parameters for image dehazing, such as dark-channel prior, color-lines prior and haze-lines prior methods. As the convolutional neural networks (CNNs) have got great development in the computer vision tasks, the learning methods are able to estimate atmosphere light and transmission maps directly without priors. It makes the dehazing methods more general to fit different scenes and decrease the difficulty of algorithms. Some important works are introduced as follows:

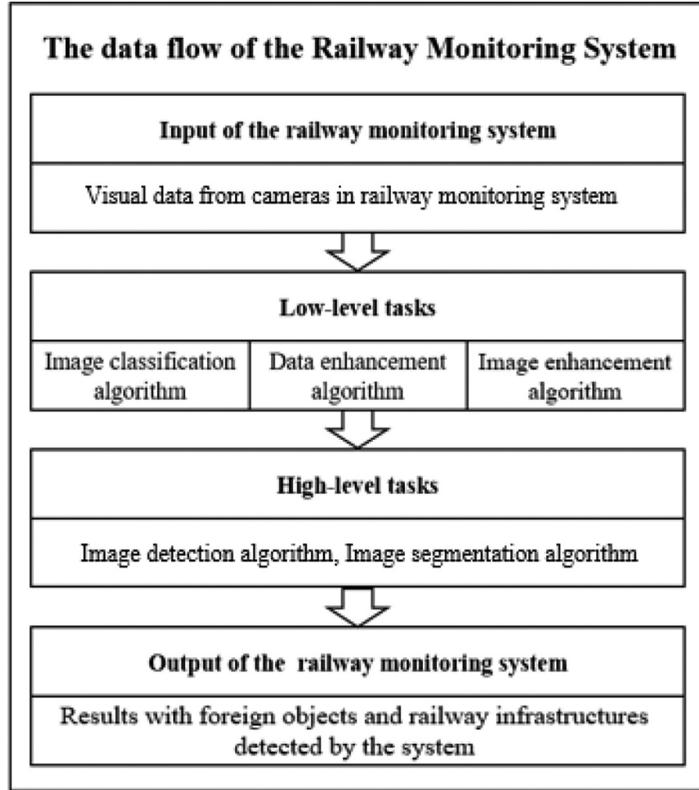


Figure 1.
Data flow of railway
monitoring system

- *Prior-based methods*: Researchers in this field are trying to build physical models for describing hazy images and concluding image features. Fattal (2008) proposed a physical-based method based on assumption that transmission maps and pixel colors are uncorrelated in the local region to estimate the albedo of haze images' background. Tan (2008) proposed a patch-based contrast-maximization method based on Markov random fields to enhance hazy images' color contrast. He et al. (2011) proposed a dark channel prior model, which depends on the observation that RGB outdoor images have at least one low intensity channel called dark channel, to estimate the transmission maps for image dehazing. Berman et al. (2017) proposed a haze-line method based the color-line method and depth estimation.
- *Learning-based methods*: Learning-based single image haze removal methods are data-driven to estimate transmission maps without prior information. Cai et al. (2016) has firstly introduced an end-to-end CNN network with a novel BReLU unit for image dehazing. Ren et al. (2016) proposed a deep neural network with the multi-scale structure to estimate hazy images' transmission maps. Li et al. (2017a) proposed a image dehazing network named all-in-one, which estimated the transmission map and the atmosphere light into one variable K. More recently, Ren et al. (2018) proposed a gate-mixed method, which is an end to end network mixing the contrast, white-balance and gamma corrections features together.

Zhang and Patel (2018) proposed a single image dehazing network, which added a multi-level PPM to estimate transmission maps.

Haze removal methods are helpful to deal with the degradation problem in hazy images. However, haze images under railway scenes are different from samples in popular haze images data sets. The challenges can be confronted as follows:

- *Isolated background scenes.* The hazy images' backgrounds are isolated under railway scenes. Most visual data's upper background is the sky and lower background is railway infrastructures. Two parts are quite different from the main color and pixel distribution. Image dehazing methods must overcome local-concentration results. Uniform image style is important for visual data quality. In addition, hazy images under railway scenes are hard to be collected on the internet without open-source data sets. It is hard to train a data-driven railway image dehazing network based on *the* above methods.
- *Depth information.* The image depth under railway scenes is hard to be estimated (Li *et al.*, 2017b; Ummenhofer *et al.*, 2017). Most kinds of depth images camera, such as TOF, stereo cameras and Kinect are limited in distance between 5 m and 30 m. However, cameras in railway monitoring systems have huge image depth over the range. Lidar is the most popular solution for detecting depth of field in long distance, such as auto-driving and robots. While the expensive cost makes the employment plan hard to be taken in the railway monitoring system. All above means haze removal methods based on atmosphere scattering model are not very suitable for hazy images under railway scenes.
- *Precision rate of image segmentation and detection algorithms in haze removal results.* The images under railway scenes always contain obvious edge information. The center regions of hazy images are blurred seriously by aerobics, which are always treated as the most far points along railway tracks. The edge information needs to be restored after the haze removal in *the* railway monitoring system, which serves for high-level visual tasks (Eitel *et al.*, 2015). Image dehazing method under railway scenes should be able to improve images' quality and restore edge information at the same time.

Because of the observations above, recent single image haze removal methods based on the atmosphere scattering model to estimate climate parameters and transmission maps are hard to be applied in the railway monitoring system. To tackle the problems above, a single image haze removal method without transmission maps is necessary. CNNs for image style transfer, image de-blurring and image de-raining have archived developments in directly building non-linear links between inputs and targets. Proposed image dehazing method should complete the transformation between hazy images and clear images, restore edge information and keep the whole dehazed image plane in a uniform style matching human perceptual-satisfied demands.

3. Enhanced densely dehazing network

The proposed enhanced densely dehazing network architecture is illustrated in Figure 2. The network consists of the following three modules, namely, EDBs, PPM and edge-preserving loss function. For building the direct transformation form hazy images and clear images, the network structure is end-to-end without prior information and using an auto encoder and decoder.

The network backbone is based on DenseNet (Huang *et al.*, 2017), which is introduced by Huang. DenseNet could increase network layers by concatenating features and require fewer parameters than former convolutional networks. To ensure maximum information flow between layers, each layer in DenseNet has direct access to gradients from loss functions and original input. These advantages mean that DenseNet is easier to train and has smaller parameter sets. In addition, image feature maps could be kept in the data flow in feed forward progress. Therefore, the DenseNet is taken as auto-encoder's backbone in this work.

3.1 Enhanced dense blocks

A densely encoder is proposed to extract more features from input hazy images. Original dense blocks just consist of bottleneck layers with Conv layers, Conv (1×1) and Conv (3×3), in kernel sizes of one and three. The Conv layers mean a batchnorm (Ioffe and Szegedy, 2015) layer, a ReLU layer and a convolutional layer. To build direct non-linear transformation from hazy images to clear images, the network must have more efficient feature extraction layers. Conv (1×1) layer based on the optimization method of ResNet (He *et al.*, 2016) is increased into the dense block. In addition, only the first batchnorm layer is retained. EDBs can be described in Figure 3. The transition blocks are the same to the original DenseNet module. To boost network training, three former dense blocks' parameters are taken from pre-trained densenet-101 model. EDBs can refined network dictionaries based on them. The feature sizes are transformed in the range of 512 and 16. The whole progress contains five EDBs and five responding refined transition blocks are prepared for the decoder.

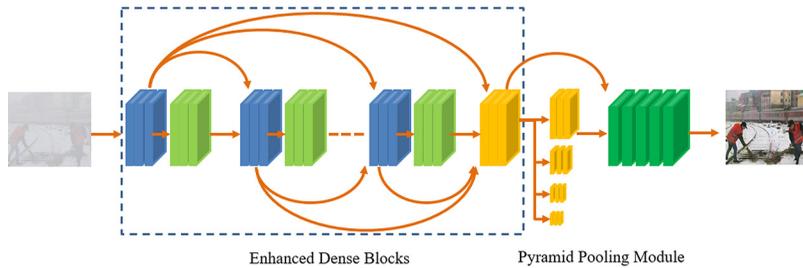


Figure 2.
The architecture of
Enhanced Densely
Dehazing Network

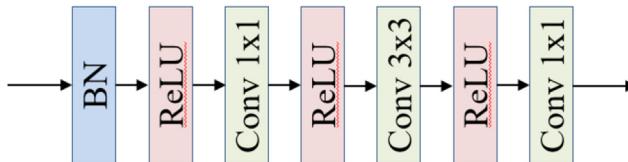


Figure 3.
The structure of
enhanced dense
blocks

3.2 Pyramid pooling module

To fuse localize and global image features, a PPM (He *et al.*, 2015; Han *et al.*, 2017) is taken into our decoder. In the encoder, the network is focused on features extraction as the layers increase. However, for a better result for image dehazing, the global features of image planes need to be considered. They are important for restoring a uniform image style. Without different scales of global image features, image dehazing results could avoid localize haze concentration. To address this issue efficiently, a multi-scale PPM is adopted to make sure that features from different scales concatenate with encoder results. This idea is inspired in image classification and segmentation tasks for efficient use of global context information. The encoder results are tackled by average-pooling layers to 1/4, 1/8, 1/16 and 1/32 sizes of inputs. These features would be up-sampled to the original encoder results' size and concatenated back with the encoder results before the final image translation.

3.3 Edge-preserving loss function

It is important to restore more edge information of hazy images under railway scenes for high-level visual tasks. Inspired of previous methods, the Euclidean loss (L2 loss) (Huang *et al.*, 2014) leads to blur the final result and the results lose the details, leading to the halo artifacts in the images. However, L2 loss is helpful to keep structure consistency during images translation. To overcome this problem, image dehazing methods in railway monitoring systems need other loss functions to restore edge information. The edge information can be captured from the following two observations. The images' gradients can characterize the image intensities and when the gradients change sharply, the lines of edges can be captured. As the low-levels of CNN's structure has been found that the features in these levels are the simple features, such as edges and contours, the first few layers can be treated as the edge detector in the deep learning network. The pre-trained VGG-16 model's output of layer relu1_2 can capture the edge information clearly. The perceptual loss based on the pre-trained VGG-16 model in low-level vision tasks is proposed by Johnson *et al.* (2016).

Based on these observations and inspired by the gradient loss used in depth estimation and image segmentation, the edge-preserving loss is proposed and contains three different parts, namely, two-directional gradient loss, perceptual loss and smooth L1 loss:

$$L_e = aL_g + bL_f + cL_s$$

where L_e indicates the edge-preserving loss, L_g indicates the gradients loss, L_f indicates the edge feature loss and L_s indicates the smooth L1 loss. a , b and c are parameters trained by the network:

$$L_g = \sum_{w,h} \left\| (G_x(I))_{w,h} - (G_x(J))_{w,h} \right\|_2 + \left\| (G_y(I))_{w,h} - (G_y(J))_{w,h} \right\|_2$$

where G_x and G_y are the gradients in horizontal direction and vertical direction and the w and h indicate the width and height of the output feature map. I and J are the input and output of the network:

$$L_f = \sum_{c,w,h} \left\| (V(I))_{c,w,h} - (V(J))_{c,w,h} \right\|_2$$

where V represents the edge detector, the layers before relu1_2 from the pre-trained VGG-16. c , w and h are the dimensions of the corresponding low-level feature in V :

$$L_s = \begin{cases} 0.5(I - J')^2, & \text{if } |I - J'| < 1 \\ |I - J'| - 0.5, & \text{otherwise} \end{cases}$$

Smooth L1 loss has replaced L2 loss in Faster-RCNN (Ren *et al.*, 2017), for it is less sensitive to outliers than L2 loss. It means the training with a smooth L1 loss would have more robustness.

4. Experiments

In this section, the experimental details and evaluation results on synthetic hazy data sets, real-world hazy data sets and railway hazy data sets have been introduced. Enough ablation tests have been taken to demonstrate the improvements of our proposed enhanced network architecture. The dehazing performance on the synthetic data set and real data set is evaluated in terms of peak signal to noise ratio (PSNR) and structure similarity (SSIM). Image dehazing results of the proposed method are compared with other state-of-the-art dehazing methods' results. The image detection experiment tests the dehazed outputs in railway data set by the Faster R-CNN model for edge-preserving loss function.

4.1 Data sets

Now learning-based dehazing methods are mostly data-driven. In theory, it is extremely hard to make hazy images data set that clear and hazy images are in the same illumination and environment conditions in large scale. Researchers have to make synthetic hazy images data set. Similar to the existed learning-based haze removal methods, the synthetic hazy data set is based on the NYU-depth 2 data set (Silberman *et al.*, 2012) by the method proposed by Li *et al.* (2017a). Each pair of clear images and their depth matrices generate four corresponding hazy images. During the synthesis progress, atmosphere light conditions are random in the range of $A \in [0.4, 1]$ and the scattering coefficients are random in the range of $\beta \in [0.4, 1.6]$. A random set of 1,200 images are selected from the NYU-depth 2 data set to generate the training and valuation data set. In addition, another 300 images are made into synthetic test data set in a similar way. Training data set and test data set have been ensured with no duplicate images.

Outdoor hazy images data set in NTIRE challenge are chosen for outside hazy images test. Hazy images in this data set are made by the fogging machine to simulate haze climate close to real world. However, the data set's scale is small with only 25 indoor images and 35 outdoor images. It would be a good test data set for image dehazing methods without image depth information. In addition, the high-resolution clear and hazy images in pair are easy for image quality assessment in PSNR and SSIM.

Finally, image dehazing test and image detection test under railway scenes are based on railway hazy images collected on the internet. The dehazing algorithm proposed in this paper has been optimized for railway hazy images' features. While there is no open-source railway hazy data set, railway hazy images would be only used for the test in this paper.

4.2 Training details

During training, ADAM (Kingma and Ba, 2014) is chosen as the optimization algorithm with a learning rate of 5×10^{-3} and the batch size of 8. All the training samples are resized to 512×512 and clear images are pair with hazy images. The model is trained for 100 epochs. The a , b and c in L_e are selected as $a = 0.5$, $b = 0.8$ and $c = 1$. All training and experiments are done on a PC with an Intel i7-CPU and an NVIDIA TITAN X GPU. The network parameters

are trained on the deep learning platform, Pytorch and for 400 epochs on the synthetic hazy data set.

4.3 Evaluation standard

Inspired by other image dehazing methods, PSNR and SSIM index are usually chosen as the evaluation standard. PSNR and SSIM are defined as follow:

$$PSNR(J', I) = 10 \times \log_{10} \frac{255^2}{\sqrt{|J' - I|^2}}$$

where J and J' are the ground truth image and corresponding predicted haze-free image:

$$SSIM = \frac{(2\mu_J\mu_{J'} + C_1)(2\sigma_{JJ'} + C_2)}{(\mu_J^2 + \mu_{J'}^2 + C_1)(\sigma_J^2 + \sigma_{J'}^2 + C_2)}$$

where μ_J and $\mu_{J'}$ are the average value of the ground truth image J and corresponding predicted haze-free image J' . σ_J and $\sigma_{J'}$ are the standard deviations of J and J' .

4.4 Ablation study

To prove each module's improvement in the proposed dehazing network, a set of ablation tests is designed like the following experiments:

- Image dehazing network with EDBs and with original dense blocks.
- Image dehazing with and without PPM.
- Image dehazing network trained by L2 loss and by our edge-preserving loss. All test models are trained in the same conditions for fair comparisons.

4.4.1 Effect of enhanced dense blocks. In our proposed network, EDBs are designed to take place of original dense blocks for capturing image features. In this experiment, the image dehazing network would be trained in two different structures, with original dense blocks and with our EDBs. The PSNR and SSIM test results show proposed EDBs could improve the image quality of dehazing results on NYU-Depth 2 data set. The PSNR and SSIM ratios are averages of 300 images (Table 1).

4.4.2 Effect of pyramid pooling module. To demonstrate the improvement by PPM, the network has been trained in two settings as before, namely, without PPM and with PPM. The dehazing results in the NTIRE indoor data set have clearly displayed differences between two networks in Figure 4. The pixels in image c around the blue toy have been in a more uniform style than pixels in image b.

Table 1.
Average PSNR and SSIM results for ablation study for effect of enhanced dense blocks

	Original dense blocks	Enhanced dense blocks
PSNR	19.51	19.72
SSIM	0.8212	0.8346

4.4.3 *Effect of loss functions.* In the third ablation study, the effects of using edge-preserving loss and L2 loss are tested. Image dehazing network is trained with two loss functions separately. Image dehazing results in Figure 5 could show that tire lines in image c are more clear than those in image b. It proves using edge-preserving loss function to train dehazing networks could restore more edge information of hazy images.

4.5 *Image dehazing experiments*

For verification of the proposed enhanced densely dehazing network, there are three experiments taken to make comparisons with state-of-the-art single image haze removal methods in this part. DehazeNet (Cai *et al.*, 2016), MSCNN (Ren *et al.*, 2016) and AODNet (Li *et al.*, 2017a) are taken in experiments. All of these methods are data-driven and rely on transmission maps estimation. The experiments are as follows:

- Test on NYU-depth 2 synthetic hazy image data set.
- Test on NTIRE outdoor hazy image data set.
- Test on railway hazy image data set.

4.5.1 *Test on NYU-depth 2 synthetic hazy image data set.* To evaluate the image dehazing quality of the proposed method, the first experiment is taken on NYU-depth 2 synthetic hazy image data set. The proposed method could complete image dehazing directly without transmission maps estimation. Hazy images in the NYU data set are under indoor scenes and provide image depth information. In PSNR and SSIM test of 300 synthetic hazy image,



Figure 4.
Image dehazing with/
without pyramid
pooling module.

Notes: (a) Input image; (b)Result without PPM; (c)Result with PPM

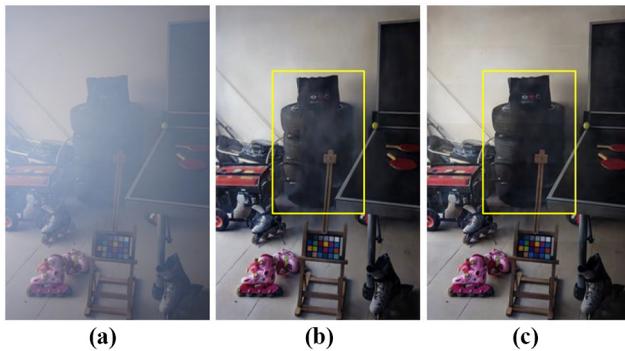


Figure 5.
Image dehazing with
L2 loss and edge-
preserving loss.

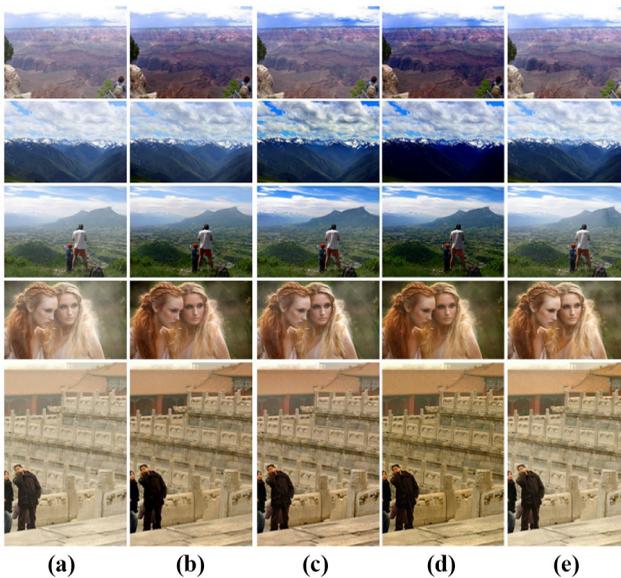
Notes: (a) Input image; (b) result with L2 loss; (c) result with edge-preserving loss

proposed method get the best score in PSNR test and AODNet gets the best score in SSIM test. In addition, the other algorithms are running on matconvnet that cost a longer time. While dehazing results of these models on traditional real-world test images are different, results of MSCNN and AODNet have higher image contrast and image distortion. DehazeNet has halo artifacts around the human body. Proposed method achieve better visual perceptual results (Table 2 and Figure 6).

4.5.2 *Test on NTIRE outdoor hazy image data set.* For hazy data in NTIRE outdoor hazy image data set is made by fogging machine, there is no image depth information. The proposed method have been trained on the NTIRE outdoor hazy image training data set. As the other methods rely on the transmission map estimation model, the network parameters are provided by their authors (Table 3 and Figure 7).

	DehazeNet	MSCNN	AODNet	Proposed
PSNR	18.85	19.21	19.68	19.72
SSIM	0.7812	0.8283	0.8401	0.8346

Table 2.
Average PSNR and SSIM results on NYU-depth 2 synthetic hazy image data set



Notes: (a) Input image; (b) DehazeNet; (c) MSCNN; (d) AODNet; (e) Proposed method

Figure 6.
Dehazing results evaluated on real-world images.

	DehazeNet	MSCNN	AODNet	Proposed
PSNR	16.53	17.56	15.03	24.598
SSIM	0.6312	0.6495	0.5385	0.777

Table 3.
Average PSNR and SSIM results on NYU-depth 2 synthetic hazy image data set

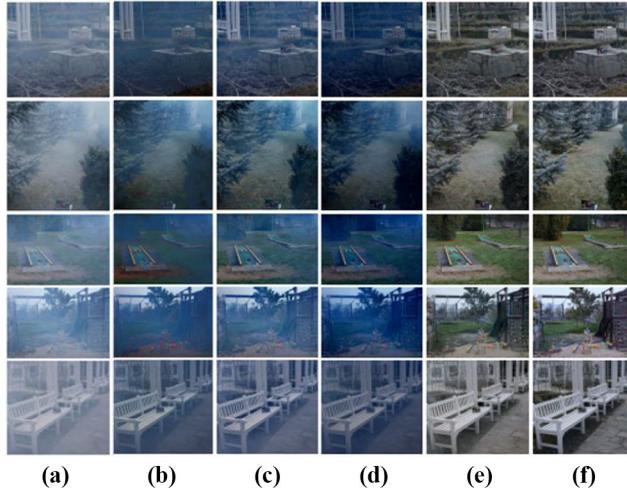


Figure 7.
Dehazing results
evaluated on NTIRE
outdoor hazy images.

Notes: (a) Input image; (b) DehazeNet; (c) MSCNN;
(d) AODNet; (e) Proposed method; (f) Ground truth

This test is trying to simulate railway hazy images data set. Without precise depth estimation, methods based on transmission maps tend to under dehaze results and darker image planes. Results of proposed methods are closer to the ground-truth and have less image distortions.

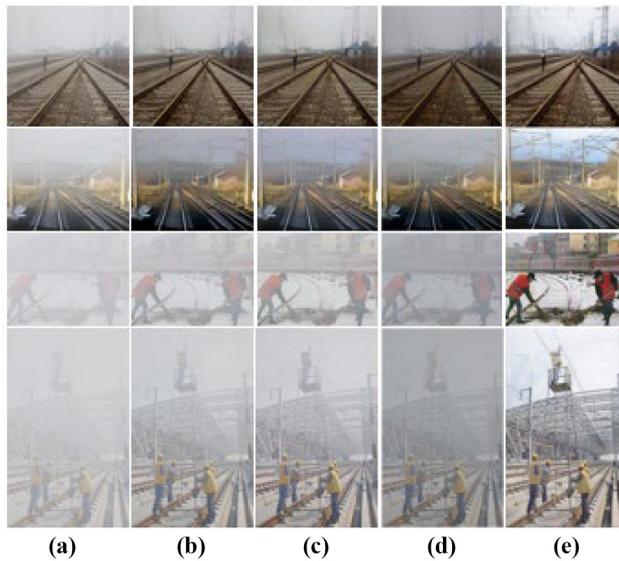


Figure 8.
Dehazing results
evaluated on railway
hazy images dataset.

Notes: (a) Input image; (b) DehazeNet; (c) MSCNN;
(d) AODNet; (e) Proposed method

4.5.3 *Test on railway hazy image data set.* Finally, above image dehazing methods are tested on real-world railway hazy images. As the prior experiment, methods based on transmission maps estimation all archive under-dehaze results. Our proposed methods' results are much more clear and match human perceptual demands (Figure 8).

4.6 *Image detection experiments*

Railway monitoring system now is going to take more high-level visual tasks. Foreign objects detection by visual surveillance has been a crucial problem now. To evaluate the edge information preserving ability in our proposed image dehazing method, dehazed results are taken in an image detection experiment based on Faster RCNN (Ren *et al.*, 2017).

From observation of image detection results in Figure 9, a number of detected objects does not change in the first pair of hazy and dehazing images, but the precision ratio has increased from 0.967 to 0.995. In the other pair, a number of detected objects and precision ratio are both enhanced in dehazing results. These demonstrate proposed image dehazing method is helpful to enhance the accuracy and efficiency of the image detection algorithm under railway scenes.



Notes: (a) Detection results of railway hazy images;
(b) Detection results of dehazing railway images

Figure 9.
Foreign objects
detection results
evaluated on railway
hazy images and
dehazing results.

5. Conclusion

In this paper, an enhanced densely dehazing network for hazy railway monitoring images is proposed without transmission maps estimation. Different from the existing methods, which are based on the atmosphere scattering model for transmission maps estimation and image dehazing, proposed image dehazing method could complete non-linear translation from hazy images and clear images. The network is efficient in synthetic, real-world and railway hazy image data sets.

The proposed method has been proved in the following contributions. First, the novel network architecture could get image dehazing results in an end-to-end way. The EDBs could capture features more efficiently than original dense blocks.

Second, the PPM is taken to avoid dehazing artifacts to a certain extent. It could fuse hazy image local features and global features and especially works well in open background railway images.

Finally, the edge-preserving loss function could help the network to restore edge information from hazy images. It has been tested by an image detection algorithm. Railway monitoring system could depend on our dehazing results for complex and precise visual tasks.

It should be noted that this work provides a new method for image enhancing in the railway monitoring system. However, the accuracy and running time still cannot match railway operation demands. For further research, railway image dehazing method needs to match real-time and robustness requirements in railway monitoring systems for practical application.

References

- Ancuti, C.O. and Ancuti, C. (2013), "Single image dehazing by multi-scale fusion", *IEEE Transactions on Image Processing*, Vol. 22 No. 8, pp. 3271-3282.
- Berman, D., Treibitz, T. and Avidan, S. (2017), "Air-light estimation using haze-lines", *2017 IEEE International Conference on Computational Photography (ICCP)*, IEEE, pp. 1-9.
- Cai, B., Xu, X., Jia, K., Qing, C. and Tao, D. (2016), "DehazeNet: an end-to-end system for single image haze removal", *IEEE Transactions on Image Processing*, Vol. 25 No. 11, pp. 5187-5198.
- Cao, Z., Qin, Y., Jia, L., Xie, Z., Liu, Q., Ma, X. and Yu, C. (2020), "Haze removal of railway monitoring images using multi-scale residual network", *IEEE Transactions on Intelligent Transportation Systems*, pp. 1-14.
- Eitel, A., Springenberg, J.T., Spinello, L., Riedmiller, M. and Burgard, W. (2015), "Multimodal deep learning for robust rgb-d object recognition", *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, pp. 681-687.
- Fattal, R. (2008), "Single image dehazing", *Proceeding of ACM SIGGRAPH Papers-SIGGRAPH*, pp. 72:1-72:9.
- Han, D., Kim, J. and Kim, J. (2017), "Deep pyramidal residual networks", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5927-5935.
- He, K., Sun, J. and Tang, X. (2011), "Single image haze removal using dark channel prior", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33 No. 12, pp. 2341-2353.
- He, K., Zhang, X., Ren, S. and Sun, J. (2016), "Deep residual learning for image recognition", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778.
- He, K., Zhang, X., Ren, S. and Sun, J. (2015), "Spatial pyramid pooling in deep convolutional networks for visual recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 37 No. 9, pp. 1904-1916.

-
- Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017), "Densely connected convolutional networks", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2261-2269.
- Huang, S.C., Chen, B.H. and Wang, W.J. (2014), "Visibility restoration of single hazy images captured in real-world weather conditions", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 24 No. 10, pp. 1814-1824.
- Ioffe, S. and Szegedy, C. (2015), "Batch normalization: accelerating deep network training by reducing internal covariate shift", *Proceeding of 32nd International Conference on Machine Learning (ICML)*, Vol. 1, pp. 448-456.
- Jiang, H. and Lu, N. (2018), "Multi-scale residual convolutional neural network for haze removal of remote sensing images", *Remote Sensing*, Vol. 10 No. 6, pp. 53-69.
- Johnson, J., Alahi, A. and Fei-Fei, L. (2016), "Perceptual losses for real-time style transfer and super-resolution", *European Conference on Computer Vision*, Springer, pp. 694-711.
- Kingma, D.P. and Ba, J.L. (2014), "Adam: a method for stochastic optimization", arXiv:1412.6980, available at: <https://arxiv.org/abs/1412.6980>.
- Li, B., Peng, X., Wang, Z., Xu, J. and Feng, D. (2017a), "AOD-net: All-in-one dehazing network", *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4780-4788.
- Li, J., Klein, R. and Yao, A. (2017b), "A two-streamed network for estimating fine-scaled depth maps from single rgb images", *The IEEE International Conference on Computer Vision (ICCV)*.
- Li, P., Jia, L.M. and Nie, A.X. (2003), "Study on railway intelligent transportation system architecture", *Intelligent Transportation Systems IEEE*.
- Li, Z., Tan, P., Tan, R.T., Zou, D., Zhiying Zhou, S. and Cheong, L.F. (2015), "Simultaneous video defogging and stereo reconstruction", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4988-4997.
- Liu, Q., Qin, Y., Xie, Z., Cao, Z. and Jia, L. (2020), "An efficient residual-based method for railway image dehazing", *Sensors*, Vol. 20 No. 21, p. 6204.
- Nayar, S.K. and Narasimhan, S.G. (2002), "Vision in bad weather", *Proceedings of the Seventh IEEE International Conference on Computer Vision IEEE*.
- Ren, S., He, K., Girshick, R. and Sun, J. (2017), "Faster R-CNN: towards real-time object detection with region proposal networks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39 No. 6, pp. 1137-1149.
- Ren, W., Ma, L., Zhang, J., Pan, J., Cao, X., Liu, W. and Yang, M.H. (2018), "Gated fusion network for single image dehazing", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3253-3261.
- Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X. and Yang, M.H. (2016), "Single image dehazing via multi-scale convolutional neural networks", *Proceeding of European Conference on Computer Vision (ECCV)*, pp. 154-169.
- Sacchi, C. and Regazzoni, C.S. (2000), "A distributed surveillance system for detection of abandoned objects in unmanned railway environments", *IEEE Transactions on Vehicular Technology*, Vol. 49 No. 5, pp. 2013-2026.
- Silberman, N., Hoiem, D., Kohli, P. and Fergus, R. (2012), "Indoor segmentation and support inference from RGBD images", *Proceeding of European conference on computer vision*, pp. 746-760.
- Tan, R.T. (2008), "Visibility in bad weather from a single image", *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, pp. 1-8.
- Ummenhofer, B., Zhou, H., Uhrig, J., Mayer, N., Ilg, E., Dosovitskiy, A. and Brox, T. (2017), "Demon: Depth and motion network for learning monocular stereo", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5038-5047.

-
- Wu, Y., Qin, Y., Wang, Z., Ma, X. and Cao, Z. (2020), "Densely pyramidal residual network for UAV-based railway images dehazing", *Neurocomputing*, Vol. 371, pp. 124-136.
- Zhang, H. and Patel, V.M. (2018), "Densely connected pyramid dehazing network", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3194-3203.

Further reading

- Glorot, X., Bordes, A. and Bengio, Y. (2011), "Deep sparse rectifier neural networks", *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 315-323.
- Jégou, S., Drozdal, M., Vazquez, D., Romero, A. and Bengio, Y. (2017), "The one hundred layers tiramisù: Fully convolutional densenets for semantic segmentation", *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, IEEE, pp. 1175-1183.
- Kratz, L. and Nishino, K. (2009), "Factorizing scene albedo and depth from a single foggy image", *IEEE 12th International Conference on Computer Vision*, IEEE, pp. 1701-1708.
- Ren, W., Pan, J., Zhang, H., Cao, X. and Yang, M.H. (2020), "Single image dehazing via multi-scale convolutional neural networks with holistic edges", *International Journal of Computer Vision*, Vol. 128 No. 1, pp. 240-259.
- Sulami, M., Glatzer, I., Fattal, R. and Werman, M. (2014), "Automatic recovery of the atmospheric light in hazy images", *2014 IEEE International Conference on Computational Photography (ICCP)*, IEEE, pp. 1-11.
- Zeiler, M.D. and Fergus, R. (2014), "Visualizing and understanding convolutional networks", *European Conference on Computer Vision*, Springer, pp. 818-833.

Corresponding author

Limin Jia can be contacted at: lmjia@bjtu.edu.cn