Does Haze Removal Help CNN-based Image Classification?

Yanting Pei[1,2], Yaping Huang[1,*], Qi Zou[1], Yuhang Lu[2], and Song Wang[2,3,*]

[1] Beijing Key Laboratory of Traffic Data Analysis and Mining, Beijing Jiaotong University, Beijing, China
[2] Department of Computer Science and Engineering, University of South Carolina, Columbia, SC, USA
[3] School of Computer Science and Technology, Tianjin University, Tianjin, China
{15112073, yphuang, qzou}@bjtu.edu.cn, yuhang@email.sc.edu,
songwang@cec.sc.edu

Abstract. Hazy images are common in real scenarios and many dehazing methods have been developed to automatically remove the haze from images. Typically, the goal of image dehazing is to produce clearer images from which human vision can better identify the object and structural details present in the images. When the ground-truth haze-free image is available for a hazy image, quantitative evaluation of image dehazing is usually based on objective metrics, such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM). However, in many applications, large-scale images are collected not for visual examination by human. Instead, they are used for many high-level vision tasks, such as automatic classification, recognition and categorization. One fundamental problem here is whether various dehazing methods can produce clearer images that can help improve the performance of the high-level tasks. In this paper, we empirically study this problem in the important task of image classification by using both synthetic and real hazy image datasets. From the experimental results, we find that the existing image-dehazing methods cannot improve much the image-classification performance and sometimes even reduce the image-classification performance.

Keywords: Hazy images · Haze removal · Image classification · Dehazing · Classification accuracy.

1 Introduction

Haze is a very common atmospheric phenomenon where fog, dust, smoke and other particles obscure the clarity of the scene and in practice, many images collected outdoors are contaminated by different levels of haze, even on a sunny day and in computer vision society, such images are usually called *hazy images*, as shown in Fig. 1(a). With intensity blurs and lower contrast, it is usually more

---

* Co-corresponding authors.

difficult to identify object and structural details from hazy images, especially when the level of haze is strong. To address this issue, many *image dehazing* methods [9, 25, 26, 20, 33, 2, 3, 21, 15] have been developed to remove the haze and try to recover the original clear version of an image. Those dehazing methods mainly rely on various image prior, such as dark channel prior [9] and color attenuation prior [33]. As shown in Fig. 1, the images after the dehazing are usually more visually pleasing – it can be easier for the human vision to identify the objects and structures in the image. Meanwhile, many objective metrics, such as Peak Signal-to-Noise Ratio (PSNR) [11] and Structural Similarity (SSIM) [30], have been proposed to quantitatively evaluate the performance of image dehzaing when the ground-truth haze-free image is available for a hazy image.

(a)   (b)   (c)   (d)

Fig. 1. An illustration of image dehazing. (a) A hazy image. (b), (c) and (d) are the images after applying different dehazing methods to the image (a).

However, nowadays large-scale image data are collected not just for visual examination. In many cases, they are collected for high-level vision tasks, such as automatic image classification, recognition and categorization. *One fundamental problem is whether the performance of these high-level vision tasks can be significantly improved if we preprocess all hazy images by applying an image-dehazing method.* On one hand, images after the dehazing are visually clearer with more identifiable details. From this perspective, we might expect the performance improvement of the above vision tasks with image dehazing. On the other hand, most image dehazing methods just process the input images without introducing new information to the images. From this perspective, we may not expect any performance improvement of these vision tasks by using image dehazing since many high-level vision tasks are handled by extracting image information for training classifiers. In this paper, we empirically study this problem in the important task of image classification.

By classifying an image based on its semantic content, *image classification* is an important problem in computer vision and has wide applications in autonomous driving, surveillance and robotics. This problem has been studied for a long time and many well known image databases, such as Caltech-256 [8], PASCAL VOCs [7] and ImageNet [5], have been constructed for evaluating the performance of image classification. Recently, the accuracy of image classification has been significantly boosted by using deep neural networks. In this paper, we will conduct our empirical study by taking Convolutional Neural Network

(CNN), one of the most widely used deep neural networks, as the image classifier and then evaluate the image-classification accuracy with and without the preprocessing of image dehazing.

More specifically, in this paper we pick eight state-of-the-art image dehazing methods and examine whether they can help improve the image-classification accuracy. To guarantee the comprehensiveness of empirical study, we use both synthetic data of hazy images and real hazy images for experiments and use AlexNet [14], VGGNet [22] and ResNet [10] for CNN implementation. Note that the goal of this paper is not the development of a new image-dehazing method or a new image-classification method. Instead, we study whether the preprocessing of image dehazing can help improve the accuracy of hazy image classification. We expect this study can provide new insights on how to improve the performance of hazy image classification.

2   Related Work

Hazy images and their analysis have been studied for many years. Many of the existing researches were focused on developing reliable models and algorithms to remove haze and restore the original clear image underlying an input hazy image. Many models and algorithms have been developed for outdoor image haze removal. For example, in [9], dark channel prior was used to remove haze from a single image. In [20], an image dehazing method was proposed with a boundary constraint and contextual regularization. In [33], color attenuation prior was used for removing haze from a single image. In [3], an end-to-end method was proposed for removing haze from a single image. In [21], multi-scale convolutional neural networks were used for haze removal. In [15], a haze-removal method was proposed by directly generating the underlying clean image through a light-weight CNN and it can be embedded into other deep models easily. Besides, researchers also investigated haze removal from the images taken at nighttime hazy scenes. For example, in [16], a method was developed to remove the nighttime haze with glow and multiple light colors. In [32], a fast haze removal method was proposed for nighttime images using the maximum reflectance prior.

Image classification has attracted extensive attention in the community of computer vision. In the early stage, hand-designed features [31] were mainly used for image classification. In recent years, significant progress has been made on image classification, partly due to the creation of large-scale hand-labeled datasets such as ImageNet [5], and the development of deep convolutional neural networks (CNN) [14]. Current state-of-the-art image classification research is focused on training feedforward convolutional neural networks using "very deep" structure [22, 23, 10]. VGGNet [22], Inception [23] and residual learning [10] have been proposed to train very deep neural networks, resulting in excellent image-classification performances on clear natural images. In [18], a cross-convolutional-layer pooling method was proposed for image classification. In [28], CNN is combined with recurrent neural networks (RNN) for improving the performance of image classification. In [6], three important visual recogni-

tion tasks, image classification, weakly supervised point-wise object localization and semantic segmentation, were studied in an integrative way. In [27], a convolutional neural network using attention mechanism was developed for image classification.

Although these CNN-based methods have achieved excellent performance on image classification, most of them were only applied to the classification of clear natural images. Very few of existing works explored the classification of degradation images. In [1], strong classification performance was achieved on corrupted MNIST digits by applying image denoising as an image preprocessing step. In [24], a model was proposed to recognize faces in the presence of noise and occlusion. In [29], classification of very low resolution images was studied by using CNN, with applications to face identification, digit recognition and font recognition. In [12], a preprocessing step of image denoising is shown to be able to improve the performance of image classification under a supervised training framework. In [4], image denoising and classification were tackled by training a unified single model, resulting in performance improvement on both tasks. Image haze studied in this paper is a special kind of image degradations and, to our best knowledge, there is no systematic study on hazy image classification and whether image dehazing can help hazy image classification.

3   Proposed Method

In this section, we elaborate on the hazy image data, image-dehazing methods, image-classification framework and evaluation metrics used in the empirical study. In the following, we first discuss the construction of both synthetic and real hazy image datasets. We then introduce the eight state-of-the-art image-dehazing methods used in our study. After that, we briefly introduce the CNN-based framework used for image classification. Finally, we discuss the evaluation metrics used in our empirical study.

3.1   Hazy-Image Datasets

For this empirical study, we need a large set of hazy images for both image-classifier training and testing. Current large-scale image datasets that are publicly available, such as Caltech-256, PASCAL VOCs and ImageNet, mainly consist of clear images without degradations. In this paper, we use two strategies to get the hazy images. First, we synthesize a large set of hazy images by adding haze to clear images using available physical models. Second, we collect a set of real hazy images from the Internet.

We synthesize hazy images by the following equation [13], where the atmospheric scattering model is used to describe the hazy image generation process:

$$I(x,y) = t(x,y) \cdot J(x,y) + [1 - t(x,y)] \cdot A, \qquad (1)$$

where (x,y) is the pixel coordinate, I is the synthetic hazy image, and J is the original clear image. A is the global atmospheric light. The scene transmission

$t(x,y)$ is distance-dependent and defined as

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

where β is the atmospheric scattering coefficient and $d(x,y)$ is the normalized distance of the scene at pixel (x,y). We compute the depth map $d(x,y)$ of an image by using the algorithm proposed in [17]. An example of such synthetic hazy image, as well as its original clear image and depth map, are shown in Fig. 2. In this paper, we take all the images in Caltech-256 to construct synthetic hazy images and the class label of each synthetic image follow the label of the corresponding original clear image. This way, we can use the synthetic images for image classification.
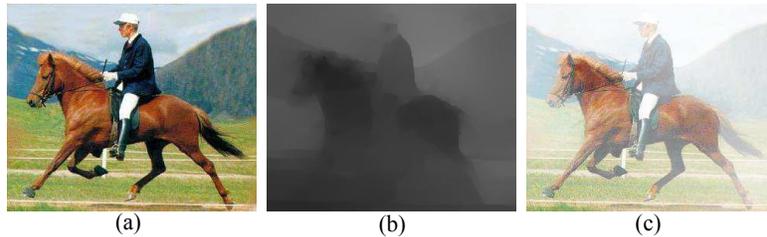

(a)   (b)   (c)

Fig. 2. An illustration of hazy image synthesis. (a) Clear image. (b) Depth map of (a). (c) Synthetic hazy image.

While we can construct synthetic hazy images by following well-acknowledged physical models, real haze models can be much more complicated and a study on synthetic hazy image datasets may not completely reflect what we may encounter on real hazy images. To address this issue, we collect a new dataset of hazy images by collecting images from the Internet. This new dataset contains 4,610 images from 20 classes and we named it as *Haze-20*. These 20 image classes are bird (231), boat (236), bridge (233), building (251), bus (222), car (256), chair (213), cow (227), dog (244), horse (237), people (279), plane (235), sheep (204), sign (221), street-lamp (216), tower (230), traffic-light (206), train (207), tree (239) and truck (223), and in the parenthesis is the number of images collected for each class. The number of images per class varies from 204 to 279. Some examples in Haze-20 are shown in Fig. 3.

In this study, we will try the case of training the image-classifier using clear images and testing on hazy images. For synthetic hazy images, we have their original clear images, which can be used for training. For real images in Haze-20, we do not have their underlying clear images. To address this issue, we collect a new *HazeClear-20* image dataset from the Internet, which consists of haze-free images that fall in the same 20 classes as in Haze-20. HazeClear-20 consists of 3,000 images, with 150 images per class.

Fig. 3. Sample hazy images in our new Haze-20 dataset.

3.2 Dehazing Methods

In this paper we try eight state-of-the-art image-dehazing methods: Dark-Channel Prior (**DCP**) [9], Fast Visibility Restoration (**FVR**) [25], Improved Visibility(**IV**) [26], Boundary Constraint and Contextual Regularization (**BCCR**) [20], Color Attenuation Prior (**CAP**) [33], Non-local Image Dehazing (**NLD**) [2], DehazeNet (**DNet**) [3], and **MSCNN** [21]. We examine each of them to see whether it can help improve the performance of hazy image classification.

- **DCP** removes haze using dark channel prior, which is based on a key observation – most local patches of outdoor haze-free images contain some pixels whose intensity is very low in at least one color channel.
- **FVR** is a fast haze-removal algorithm based on the median filter. Its main advantage is its fast speed since its complexity is just a linear function of the input-image size.
- **IV** enhances the contrast of an input image so that the image visibility is improved. It computes the data cost and smoothness cost for every pixel by using Markov Random Fields.
- **BCCR** is an efficient regularization method for removing haze. In particular, the inherent boundary constraint on the transmission function combined with a weighted L_1-norm based contextual regularization, is modeled into an optimization formulation to recover the unknown scene transmission.
- **CAP** removes haze using color attenuation prior that is based on the difference between the saturation and the brightness of the pixels in the hazy image. By creating a linear model, the scene depth of the hazy image is computed with color attenuation prior, where the parameters are learned by a supervised method.
- **NLD** is a haze-removal algorithm based on a non-local prior, by assuming that colors of a haze-free image are well approximated by a few hundred of distinct colors in the form of tight clusters in RGB space. In a hazy image, these tight color clusters change due to haze and form lines in RGB space that pass through the airlight coordinate.
- **DNet** is an end-to-end haze-removal method based on CNN. The layers of CNN architecture are specially designed to embody the established priors

in image dehazing. DNet conceptually consists of four sequential operations
– feature extraction, multi-scale mapping, local extremum and non-linear
regression, which are constructed by three convolution layers, a max-pooling,
a Maxout unit and a bilinear ReLU activation function, respectively.
– **MSCNN** uses a multi-scale deep neural network for image dehazing by
learning the mapping between hazy images and their corresponding trans-
mission maps. It consists of a coarse-scale net which predicts a holistic trans-
mission map based on the entire image, and a fine-scale net which refines
results locally. The network consists of four operations: convolution, max-
pooling, up-sampling and linear combination.

3.3　Image Classification Model

In this paper, we implement CNN-based model for image classification by using
AlexNet [14], VGGNet-16 [22] and ResNet-50 [10] on Caffe. The AlexNet [14]
has 8 weight layers (5 convolutional layers and 3 fully-connected layers). The
VGGNet-16 [22] has 16 weight layers (13 convolutional layers and 3 fully-connected
layers). The ResNet-50 [10] has 50 weight layers (49 convolutional layers and 1
fully-connected layer). For those three networks, the last fully-connected layer
has N channels (N is the number of classes).

3.4　Evaluation Metrics

We will quantitatively evaluate the performance of image dehazing and the per-
formance of image classification. Other than visual examination, Peak Signal-to-
Noise Ratio (PSNR) [11] and Structural Similarity (SSIM) [30] are widely used
for evaluating the performance of image dehazing when the ground-truth haze-
free image is available for each hazy image. For image classification, classification
accuracy is the most widely used performance evaluation metric.

Note that, both PSNR and SSIM are objective metrics based on image statis-
tics. Previous research has shown that they may not always be consistent with the
image-dehazing quality perceived by human vision, which is quite subjective. In
this paper, what we concern about is the performance of image classification after
incorporating image dehazing as preprocessing. Therefore, we will study whether
PSNR and SSIM metrics show certain correlation to the image classification per-
formance. In this paper, we simply use the classification accuracy $Accuracy = \frac{R}{N}$
to objectively measure the image-classification performance, where N is the to-
tal number of testing images and R is the total number of testing images that
are correctly classified by using the trained CNN-based models.

4　Experiments

4.1　Datasets and Experiment Setup

In this section, we evaluate various image-dehazing methods on the hazy images
synthesized from Caltech-256 and our newly collected Haze-20 datasets.

We synthesize hazy images using all the images in Caltech-256 dataset, which has been widely used for evaluating image classification algorithms. It contains 30,607 images from 257 classes, including 256 object classes and a clutter class. In our experiment, we select six different hazy levels for generating synthetic images. Specifically, we set the parameter $\beta = 0, 1, 2, 3, 4, 5$ respectively in Eq.(2) for hazy image synthesis where $\beta = 0$ corresponds to original images in Caltech-256. In Caltech-256, we select 60 images randomly from each class as training images, and the rest are used for testing. Among the training images, 20% per class are used as a validation set. We follow this to split the synthetic hazy image data: an image is in training set if it is synthesized from an image in the training set and in testing set otherwise. This way, we have a training set of 60×257 $= 15,420$ images (60 per class) and a testing set of $30,607 - 15,420 = 15,187$ images for each hazy level.

For the collected real hazy images in Haze-20, we select 100 images randomly from each class as training images, and the rest are used for testing. Among the training images, 20% per class are used as a validation set. So, we have a training set of $100 \times 20 = 2,000$ images and a testing set of $4,610 - 2,000 = 2,610$ images. For HazeClear-20 dataset, we also select 100 images randomly from each class as training images, and the rest are used for testing. Among the training images, 20% per class are used as a validation set. So, we have a training set of $100 \times 20 = 2,000$ images and a testing set of $50 \times 20 = 1,000$ images.

While the proposed CNN model can use AlexNet, VGGNet, ResNet or another network structures, for simplicity, we use AlexNet, VGGNet-16, ResNet-50 on Caffe in this paper. The CNN architectures are pre-trained on ImageNet dataset that consists of 1,000 classes with 1.2 million training images. We then use the collected images to fine-tune the pre-trained model for image classification, in which we change the number of channels in the last fully connected layer from 1,000 to N, where N is the number of classes in our datasets. To more comprehensively explore the effect of haze-removal to image classification, we study different combinations of the training and testing data, including training and testing on images without applying image dehazing, training and testing on images after dehazing, and training on clear images but testing on hazy images.

4.2   Quantitative Comparisons on Synthetic and Real Hazy Images

To verify whether haze-removal preprocessing can improve the performance of hazy image classification, we test on the synthetic and real hazy images with and without haze removal for quantitative evaluation. The classification results are shown in Fig. 4, where (a-e) are the classification accuracies on testing synthetic hazy images with $\beta = 1, 2, 3, 4, 5$, respectively using different dehazing methods. For these five curve figures, the horizontal axis lists different dehazing methods, where "Clear" indicates the use of the testing images in the original Caltech-256 datasets and this assumes a perfect image dehazing in the ideal case. The case of "Haze" indicates the testing on the hazy images without any dehazing. (f) is the classification accuracy on the testing images in Haze-20 using different dehazing methods, where "Clear" indicates the use of testing images in HazeClear-20 and

"Haze" indicates the use of testing images in Haze-20 without any dehazing. *AlexNet_1*, *VGGNet_1* and *ResNet_1* represent the case of training and testing on the same kinds of images, e.g., training on the training images in Haze-20 after DCP dehazing, then testing on testing images in Haze-20 after DCP dehazing, by using AlexNet, VGGNet and ResNet, respectively. *AlexNet_2*, *VGGNet_2* and *ResNet_2* represent the case of training on clear images, i.e., for (a-e), we train on training images in original Caltech-256, and for (f), we train on training images in HazeClear-20, by using AlexNet, VGGNet and ResNet, respectively.

Fig. 4. The classification accuracy on different hazy images. (a-e) Classification accuracies on testing synthetic hazy images with $\beta = 1, 2, 3, 4, 5$, respectively. (f) Classification accuracy on the testing images in Haze-20.

We can see that when we train CNN models on clear images and test them on hazy images with and without haze removal (e.g., *AlexNet_2*, *VGGNet_2* and *ResNet_2*), the classification performance drop significantly. From Fig. 4(e), image classification accuracy drop from 71.7% to 21.7% when images have a haze level of $\beta = 5$ by using AlexNet. Along the same curve shown in Fig. 4(e), we can see that by applying a dehazing method on the testing images, the classification accuracy can move up to 42.5% (using MSCNN dehazing). But it is still much lower than 71.7%, the accuracy on classifying original clear images. These experiments indicate that haze significantly affects the accuracy of CNN-based image classification when training on original clear images. However, if we directly train the classifiers on the hazy image of the same level, the classification

accuracy moves up to 51.9%, as shown in the red curve in Fig. 4(e), where no dehazing is involved in training and testing images. Another choice is to apply the same dehazing methods to both training and testing images: From results shown in all the six subfigures in Fig. 4, we can see that the resulting accuracy is similar to the case where no dehazing is applied to training and testing images. This indicates that the dehazing conducted in this study does not help image classification. We believe this is due to the fact that the dehazing does not introduce new information to the image.

There are also many non-CNN-based image classification methods. While it is difficult to include all of them into our empirical study, we try the one based on sparse coding [31] and the results are shown in Fig. 5, where $\beta = 1, 2, 3, 4, 5$ represent haze levels of synthetic hazy images in Caltech-256 dataset and *Haze-20* represents Haze-20 dataset. For this specific non-CNN-based image classification method, we can get the similar conclusion that the tried dehazing does not help image classification, as shown in Fig. 5. Comparing Figs. 4 and 5, we can see that the classification accuracy of this non-CNN-based method is much lower than the state-of-the-art CNN-based methods. Therefore, we focus on CNN-based image classification in this paper.
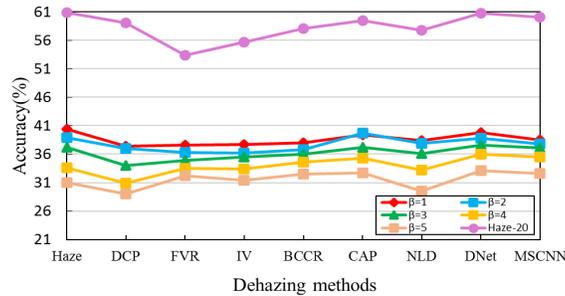

Fig. 5. Classification accuracy (%) on synthetic and real-world hazy images by using a non-CNN-based image classification method. Here the same kinds of images are used for training, i.e., building the basis for sparse coding, and testing, just like the case corresponding to the solid curves (*AlexNet_1*, *VGGNet_1* and *ResNet_1*) in Fig. 4.

4.3   Training on Mixed-Level Hazy Images

For more comprehensive analysis of dehazing methods, we conduct experiments of training on hazy images with mixed haze levels. For synthetic dataset, we try two cases. In Case 1, we mix all six levels of hazy images by selecting 10 images per class from each level of hazy images as training set and among the training images, two images per class per haze level are taken as validation set. We then test on the testing images of the involved haze levels – actually all six levels for this case – respectively. Results are shown in Fig. 6(a), (b)
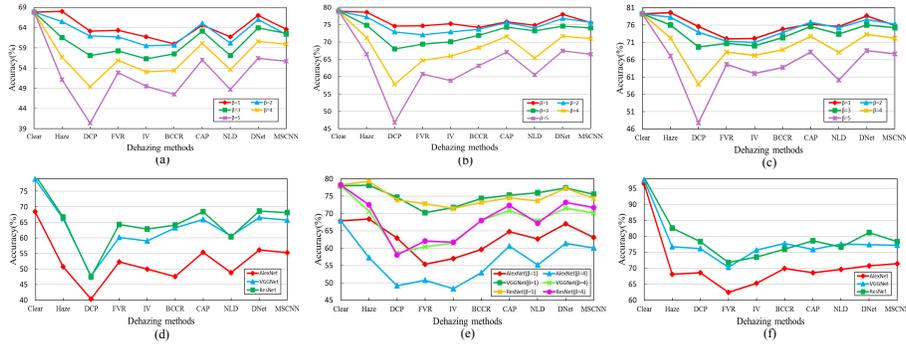
Fig. 6. Classification accuracy when training on mixed-level hazy images. (a, b, c) Mix all six levels of synthetic images. (d) Mix two levels $\beta = 0$ and $\beta = 5$. (e) Mix two levels $\beta = 1$ and $\beta = 4$. (f) Mix Haze-20 and HazeClear-20.

and (c) when using AlexNet, VGGNet and ResNet respectively. In Case 2, we randomly choose images from two different haze levels and mix them. In this case, 30 images per class per level are taken as training images and among the training images, 6 images per class per level are used as validation images. This way we have 60 images per class for training. Similarly, we then test on the testing images of the involved two haze levels, respectively. Results are shown in Fig. 6(d) and (e) for four different kinds of level combinations, respectively. For real hazy images, we mix clear images in HazeClear-20 and hazy images in Haze-20 by picking 50 images per class for training and then test on the testing images in Haze-20 and HazeClear-20 respectively. Results are shown in Fig. 6(f). Similarly, combining all the results, the use of dehazing does not clearly improve the image classification accuracy, over the case of directly training and testing on hazy images.

4.4 Performance Evaluation of Dehazing Methods

In this section, we study whether there is a correlation between the dehazing metrics PSNR/SSIM and the image classification performance. On the synthetic images, we can compute the metrics PSNR and SSIM on all the dehazing results, which are shown in Fig. 7. In this figure, the PSNR and SSIM values are averaged over the respective testing images. We pick the red curves (*AlexNet_1*) from Fig. 4(a-e) and for each haze level in $\beta = 1, 2, 3, 4, 5$, we rank all the dehazing methods based on the classification accuracy. We then rank these methods based on average PSNR and SSIM at the same haze level. Finally we calculate the rank correlation between image classification and PSNR/SSIM at each haze level. Results are shown in Table 1. Negative values indicate negative correlation, positive values indicate positive correlation and the greater the absolute value, the higher the correlation. We can see that their correlations are actually low, especially when $\beta = 3$.
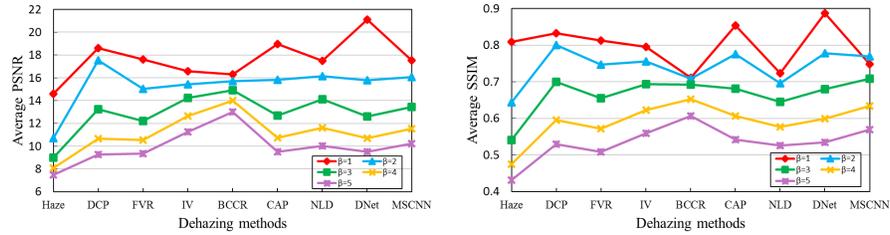
Fig. 7. Average PSNR and SSIM values on synthetic image dataset at different haze levels.

Table 1. The rank correlation between image-classification accuracy and PSNR/SSIM at each haze level.

Correlation	$\beta = 1$	$\beta = 2$	$\beta = 3$	$\beta = 4$	$\beta = 5$
(Accuracy, PSNR)	-0.3095	0.3571	0.0952	-0.2143	0.1905
(Accuracy, SSIM)	-0.2381	-0.5238	-0.0714	0.6905	0.6190

4.5   Subjective Evaluation

In this section, we conduct an experiment for subjective evaluation of the image dehazing. By observing the dehazed images, we randomly select 10 images per class with $\beta = 3$ and subjectively divide them into 5 with better dehazing effect and 5 with worse dehazing effect. This way, we have 2,570 images in total (set M) and 1,285 images each with better dehazing (set A) and worse dehazing (set B). Classification accuracy (%) using VGGNet is shown in Fig. 8 and we can see that there is no significant accuracy difference for these three sets. This indicates that the classification accuracy is not consistent with the human subjective evaluation of the image dehazing quality.
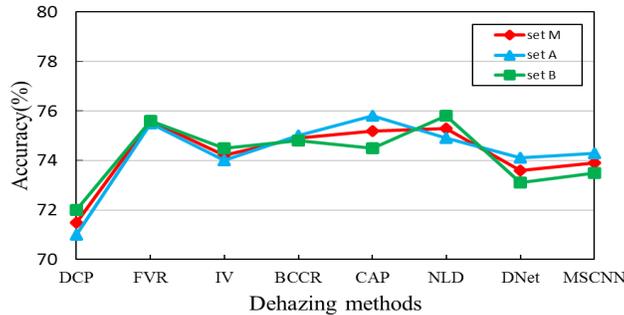

Fig. 8. Classification accuracy of different sets of dehazed images subjectively selected by human.

4.6   Feature Reconstruction

The CNN networks used for image classification consists of multiple layers to extract deep image features. One interesting question is whether certain layers in the trained CNN actually perform image dehazing implicitly. We picked a reconstruction method [19] to reconstruct the image according to feature maps of all the layers in AlexNet. The reconstruction results are shown in Fig. 9, from which we can see that, for the first several layers, the reconstructed images do not show any dehazing effect. For the last several layers, the reconstructed images have been distorted, let alone dehazing. One possibility of this is that many existing image dehazing methods aim to please human vision system, which may not be good to CNN-based image classification. Meanwhile, many existing image dehazing methods introduce information loss, such as color distortion, and may increase the difficulty of image classification.
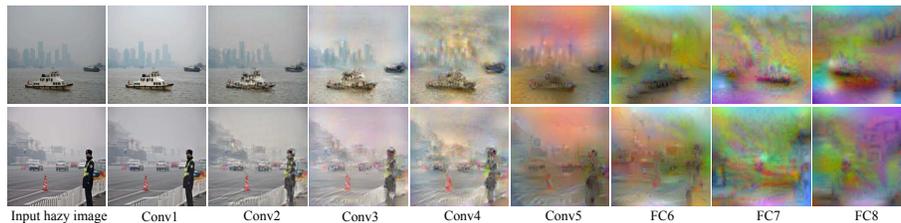

Fig. 9. Sample feature reconstruction results for two images, shown in two rows respectively. The leftmost column shows the input hazy images and the following columns are the images reconstructed from different layers in AlexNet.

4.7   Feature Visualization

In order to further analyze different dehazing methods, we extract and visualize the features at hidden layers using VGGNet. For an input image with size $H \times W$, the activations of a convolution layer is formulated as an order-3 tensor with $H \times W \times D$ elements, where D is the number of channels. The term "activations" is a feature map of all the channels in a convolution layer. The activations in haze-removal images with different dehazing methods are displayed in Fig. 10. From top to bottom are haze-removal images, and the activations at $pool_1$, $pool_3$ and $pool_5$ layers, respectively. We can see that different dehazing methods actually have different activations, such as the activations of $pool_5$ layer of NLD and DNet.

5   Conclusions

In this paper, we conducted an empirical study to explore the effect of image dehazing to the performance of CNN-based image classification on synthetic and
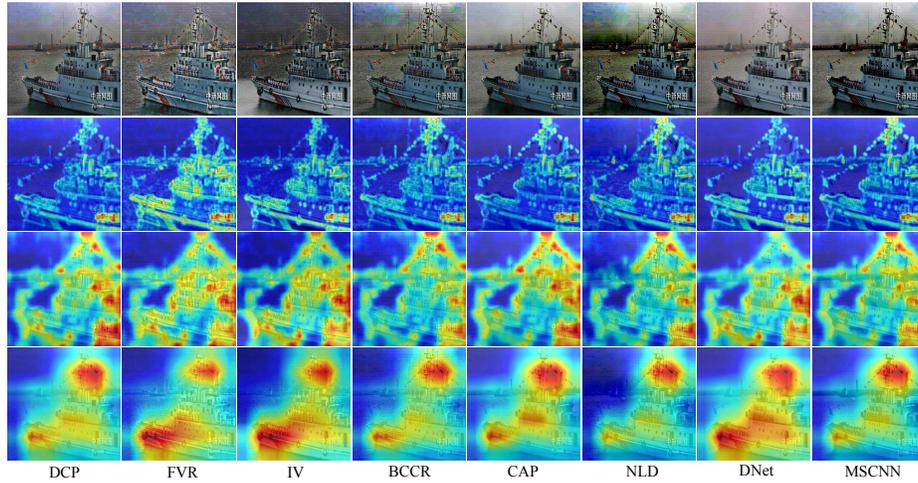
| DCP | FVR | IV | BCCR | CAP | NLD | DNet | MSCNN |

Fig. 10. Activations of hidden layers of VGGNet on image classification. From top to bottom are the haze-removal images, and the activations at $pool_1$, $pool_3$ and $pool_5$ layers, respectively.

real hazy images. We used physical haze models to synthesize a large number of hazy images with different haze levels for training and testing. We also collected a new dataset of real hazy images from the Internet and it contains 4,610 images from 20 classes. We picked eight well-known dehazing methods for our empirical study. Experimental results on both synthetic and real hazy datasets show that the existing dehazing algorithms do not bring much benefit to improve the CNN-based image-classification accuracy, when compared to the case of directly training and testing on hazy images. Besides, we analyzed the current dehazing evaluation measures based on pixel-wise errors and local structural similarities and showed that there is not much correlation between these dehazing metrics and the image-classification accuracy when the images are preprocessed by the exsiting dehazing methods. While we believe this is due to the fact that image dehazing does not introduce new information to help image classification, we do not exclude the possibility that the existing image-dehazing methods are not sufficiently good in recovering the original clear image and better image-dehazing methods developed in the future may help improve image classification. We hope this study can draw more interests from the community to work on the important problem of haze image classification, which plays a critical role in applications such as autonomous driving, surveillance and robotics.

References

1. Agostinelli, F., Anderson, M.R., Lee, H.: Adaptive multi-column deep neural networks with application to robust image denoising. In: Advances in Neural Information Processing Systems. pp. 1493–1501 (2013)
2. Berman, D., Treibitz, T., Avidan, S., et al.: Non-local image dehazing. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1674–1682 (2016)
3. Cai, B., Xu, X., Jia, K., Qing, C., Tao, D.: Dehazenet: An end-to-end system for single image haze removal. IEEE Transactions on Image Processing **25**(11), 5187–5198 (2016)
4. Chen, G., Li, Y., Srihari, S.N.: Joint visual denoising and classification using deep learning. In: IEEE International Conference on Image Processing. pp. 3673–3677 (2016)
5. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 248–255 (2009)
6. Durand, T., Mordan, T., Thome, N., Cord, M.: Wildcat: Weakly supervised learning of deep convnets for image classification, pointwise localization and segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition (2017)
7. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International Journal of Computer Vision **88**(2), 303–338 (2010)
8. Griffin, G., Holub, A., Perona, P.: Caltech-256 object category dataset (2007)
9. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence **33**(12), 2341–2353 (2011)
10. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778 (2016)
11. Huynh-Thu, Q., Ghanbari, M.: Scope of validity of psnr in image/video quality assessment. Electronics letters **44**(13), 800–801 (2008)
12. Jalalvand, A., De Neve, W., Van de Walle, R., Martens, J.P.: Towards using reservoir computing networks for noise-robust image recognition. In: International Joint Conference on Neural Networks. pp. 1666–1672 (2016)
13. Koschmieder, H.: Theorie der horizontalen sichtweite. Beitrage zur Physik der freien Atmosphare pp. 33–53 (1924)
14. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. pp. 1097–1105 (2012)
15. Li, B., Peng, X., Wang, Z., Xu, J., Feng, D.: Aod-net: All-in-one dehazing network. In: IEEE International Conference on Computer Vision. pp. 4770–4778 (2017)
16. Li, Y., Tan, R.T., Brown, M.S.: Nighttime haze removal with glow and multiple light colors. In: IEEE International Conference on Computer Vision. pp. 226–234 (2015)
17. Liu, F., Shen, C., Lin, G.: Deep convolutional neural fields for depth estimation from a single image. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 5162–5170 (2015)
18. Liu, L., Shen, C., van den Hengel, A.: The treasure beneath convolutional layers: Cross-convolutional-layer pooling for image classification. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 4749–4757 (2015)

19. Mahendran, A., Vedaldi, A.: Understanding deep image representations by inverting them. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 5188–5196 (2015)
20. Meng, G., Wang, Y., Duan, J., Xiang, S., Pan, C.: Efficient image dehazing with boundary constraint and contextual regularization. In: IEEE International Conference on Computer Vision. pp. 617–624 (2013)
21. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., Yang, M.H.: Single image dehazing via multi-scale convolutional neural networks. In: European Conference on Computer Vision. pp. 154–169 (2016)
22. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
23. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., et al.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–9 (2015)
24. Tang, Y., Salakhutdinov, R., Hinton, G.: Robust boltzmann machines for recognition and denoising. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 2264–2271 (2012)
25. Tarel, J.P., Hautiere, N.: Fast visibility restoration from a single color or gray level image. In: IEEE International Conference on Computer Vision. pp. 2201–2208 (2009)
26. Tarel, J.P., Hautiere, N., Cord, A., Gruyer, D., Halmaoui, H.: Improved visibility of road scene images under heterogeneous fog. In: IEEE Intelligent Vehicles Symposium. pp. 478–485 (2010)
27. Wang, F., Jiang, M., Qian, C., Yang, S., Li, C., Zhang, H., Wang, X., Tang, X.: Residual attention network for image classification. arXiv preprint arXiv:1704.06904 (2017)
28. Wang, J., Yang, Y., Mao, J., Huang, Z., Huang, C., Xu, W.: Cnn-rnn: A unified framework for multi-label image classification. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 2285–2294 (2016)
29. Wang, Z., Chang, S., Yang, Y., Liu, D., Huang, T.S.: Studying very low resolution recognition using deep networks. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 4792–4800 (2016)
30. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**(4), 600–612 (2004)
31. Yang, J., Yu, K., Gong, Y., Huang, T.: Linear spatial pyramid matching using sparse coding for image classification. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1794–1801 (2009)
32. Zhang, J., Cao, Y., Fang, S., Kang, Y., Chen, C.W.: Fast haze removal for nighttime image using maximum reflectance prior. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 7418–7426 (2017)
33. Zhu, Q., Mai, J., Shao, L.: A fast single image haze removal algorithm using color attenuation prior. IEEE Transactions on Image Processing **24**(11), 3522–3533 (2015)