

Let You See in Haze and Sandstorm: Two-in-One Low-Visibility Enhancement Network

Yuan Gao^{ID}, Wenyu Xu^{ID}, and Yuxu Lu^{ID}

Abstract—Scattering particles (such as water vapor and dust) in the atmosphere excessively absorb and scatter light, which causes the image collected by imaging sensors with low visibility, low contrast, color shifting, etc. This limitation hinders the development of vision-driven artificial intelligence applications, such as the intelligent traffic monitoring system (ITMS). To improve the imaging quality in hazy and sandstorm conditions, this article proposes a novel two-in-one low-visibility enhancement network (termed TOENet). Specifically, we take into account the changes in the histogram distributions of the red, green, and blue channels of the hazy and sandstorm images and present a multilayer perceptron (MLP)-based channel correlation extraction module (CCEM) to learn and reconstruct the correlation between the channels of low-visibility images. The multiscale channel attention features generated by CCEM will be embedded into the backbone (i.e., the encoder-decoder network) of TOENet to extract additional latent feature information from low-visibility images. TOENet can achieve both dehazing and sandstorm enhancement through a single deep network model without incurring extra computational costs. Extensive experimental results demonstrate that the proposed TOENet outperforms several state-of-the-art methods on challenging datasets. Furthermore, the running time and object detection experiments indicate the practical benefits of our method for high-level visual task under low-visibility weathers in ITMS. The source code is available at <https://github.com/YuanGao-YG/TOENet>.

Index Terms—Channel correlation, dehazing, intelligent traffic monitoring system (ITMS), low-visibility, sandstorm enhancement.

I. INTRODUCTION

HIGH-QUALITY visual imaging is a key prerequisite for the development of intelligent traffic monitoring system (ITMS) [1], [2], [3]. But in harsh imaging situations, such as haze and sandstorm, the collected videos and images have poor contrast, low visibility, and color shifting, which hinders the implementation of vision-driven intelligent systems [4]. To alleviate these problems, researchers proposed

Manuscript received 13 June 2023; accepted 21 July 2023. Date of publication 14 August 2023; date of current version 30 August 2023. This work was supported by the Fundamental Research Funds for the Central Universities under Grant 225212002. The Associate Editor coordinating the review process was Dr. Yu Yang. (*Yuan Gao and Wenyu Xu are co-first authors.*) (*Corresponding author: Yuxu Lu.*)

Yuan Gao is with the School of Navigation, Wuhan University of Technology, Wuhan 430063, China (e-mail: yuangao@whut.edu.cn).

Wenyu Xu is with Wuhan Baosight Software Company Ltd., Wuhan 430080, China (e-mail: xuwenyu_971652@baosight.com).

Yuxu Lu is with the Department of Logistics and Maritime Studies, The Hong Kong Polytechnic University, Hong Kong, China (e-mail: louisyuxulu@gmail.com).

Digital Object Identifier 10.1109/TIM.2023.3304668

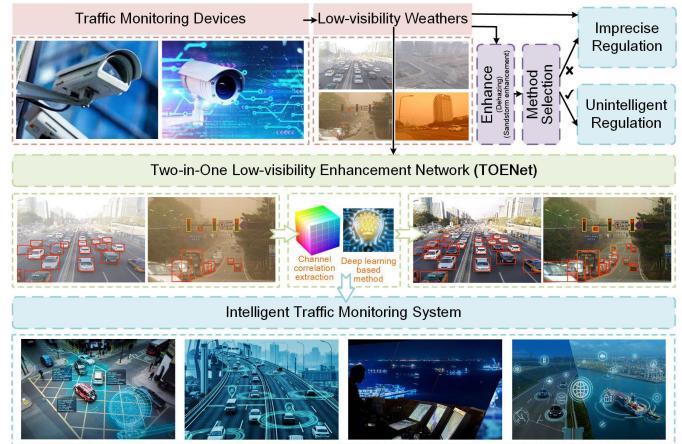


Fig. 1. Workflow of our TOENet-based low-visibility image enhancement for the vision-driven ITMS.

traditional (including enhancement- and restoration-based) and learning-based methods to extract valuable information from low-visibility images and achieved great progress. However, in the real world, the weather is often changeable, and low-visibility weather conditions such as haze and sandstorm are similar and can be challenging to distinguish. This variability poses a problem when attempting to enhance images captured under such conditions. Fig. 1 shows that embedding the wrong enhancement method in traffic monitoring devices will result in poor image quality and imprecise regulation. If the regulatory authority arranges duty personnel to select different enhancement methods in real-time according to the weather conditions, even if the enhancement method is selected correctly for the low-visibility weather, the level of intelligence is limited. Thus, a more sophisticated solution is needed to address these challenges. Therefore, it is necessary to design an efficient two-in-one general-purpose dehazing and sandstorm enhancement low-visibility enhancer.

The enhancement-based methods are composed of histogram equalization (HE)-based [5], [6], Retinex-based [7], [8], homomorphic filtering [9], gamma correction [10], etc. Shi et al. [11] propose a normalized gamma-transformation-based contrast-limited adaptive HE with color correction in laboratory color space to reduce the impact of the sandstorm. Fu et al. [8] propose a Retinex-based framework to decompose the reflected and illuminated components of low-visibility images and optimize the underlying detail and brightness information. Despite the fact that enhancement-based methods aim to improve image contrast and color, it is challenging

to identify the particular causes of image degradation, which may result in the loss of information. The restoration-based methods examine the image deterioration process from the standpoint of physical optics, assess the degradation model, and then use the degradation model to restore the prospective clear image as much as possible. The dark channel prior (DCP) [12] and its improved methods have been widely used in dehazing and sandstorm enhancement. Liu et al. [13] propose an intensity projection strategy that estimates transmittance more precisely and extracts valuable feature information from haze and sandstorm more efficiently. However, when the imaging scene is too complicated, the restoration-based methods are difficult to accurately estimate the imaging process of the damaged image, which causes the enhanced image with inaccurate color expression.

In recent years, deep-learning-based methods have attracted much attention in the field of image restoration due to their powerful nonlinear mapping capabilities. Cai et al. [14] use a deep convolutional neural network to estimate the transmission map of the atmospheric scattering model and construct the mapping relationship between hazy images and intermediate parameters. Ren et al. [15] extract the global and local features of the transmittance map through convolution kernels of different receptive fields to finely estimate the transmittance. AODNet [16] uses the atmospheric scattering model and generates latent dehazed images with lightweight densely connected blocks [17]. Guo et al. [18] design a self-paced semi-curricular attention network to comply with image dehazing. Li et al. [19] construct an all-in-one image restoration to achieve degraded image enhancement in various scenarios. Qu et al. [20] propose a learned parameter-sharing network to achieve general image enhancement. Due to the low dependence of the generative adversarial network (GAN) on the training dataset, some researchers [21], [22] can realistically simulate hazy images and generate natural-looking dehazed images from complex imaging environments. Chen et al. [23] propose a physics and depth guided end-to-end GAN image dehazing method. Meantime, the rapid development of transformer provides a new solution for image restoration, Yang et al. [24] construct a multiscale transformer fusion module to capture the long-range dependencies of image information in space. Song et al. [25] improve the transformer network structure so that it can meet the high requirements of image restoration tasks for texture structure, color, etc.

As discussed above, traditional methods in image dehazing and sandstorm image enhancement primarily focus on improving image contrast and color, which may not fully address the challenges posed by complex degradation caused by sandstorms or other types of haze. As a result, these methods may lead to information loss. In addition, they heavily rely on prior knowledge and assumptions, which may not accurately estimate the imaging process of damaged images in complex and realistic hazy or sandstorm weather conditions, resulting in undesirable enhancement results. The performance of deep-learning-based dehazing research has made significant advancements. However, the lack of interpretability of many networks results in heavy reliance on the depth of the network and the training dataset, which can lead to

limited generalization capability. Furthermore, real-world hazy scenarios are often complex, causing many methods to perform poorly in actual scenarios, despite good performance on synthetic datasets. In addition, the deep-learning-based sandstorm image enhancement methods are few and underperform in real scenes, mainly due to the scarcity of paired sandstorm datasets and the limited analysis of sandstorm image characteristics, resulting in suboptimal performance in actual sandstorm scenarios. Therefore, this article deeply analyzes the characteristics of the two types of low-visibility images and uses the image RGB distribution features for image restoration instead of treating the neural network as a black box. Our method can improve the network's enhancement performance in real-world scenarios using limited datasets and the powerful feature mapping capabilities of neural network. Notably, the correct application of RGB distribution features enables us to simultaneously enhance two types of low-visibility images in a single network, making it a more practical solution for use in ITMS.

In this work, we present a two-in-one low-visibility enhancement network (termed TOENet) to improve the imaging quality under both hazy and sandstorm weather conditions. Specifically, the proposed multilayer perceptron (MLP)-based channel correlation extraction module (CCEM) achieves effective learning and reconstruction of correlations between the red, green, and blue channels. Then we suggest the encoder-decoder framework as the backbone of the network, aided by the generated channel attention features, which may extract abundant and valuable feature information from hazy and sandstorm images with a single network model. The main contributions are summarized as follows.

- 1) We propose a two-in-one low-visibility enhancement network to improve the imaging performance in hazy and sandstorm weather conditions. Specifically, the proposed TOENet takes the encoder-decoder as the backbone, aided by channel attention features to extract potential feature information.
- 2) We suggest an MLP-based CCEM to learn and reconstruct the pixel value distribution of the RGB channels of the low-visibility images.
- 3) Without loss of generalization, we conduct extensive experiments to verify the effectiveness of TOENet in dehazing and sandstorm enhancement with competitive methods.

The rest of this article is organized as follows. Physical imaging models for hazy and sandstorm images are presented in Section II. Section III introduces our TOENet. Numerous experiments have been implemented to evaluate the performance of TOENet in Section IV. Conclusions are given in Section V.

II. PROBLEM FORMULATION

Scattering particles in the air scatter atmospheric light to varying degrees. During hazy and sandstorm weather conditions, water droplet particles and sand particles suspended in the air will refract light, which will reduce the imaging quality of imaging devices. The atmospheric scattering model divides the information received by the imaging device into

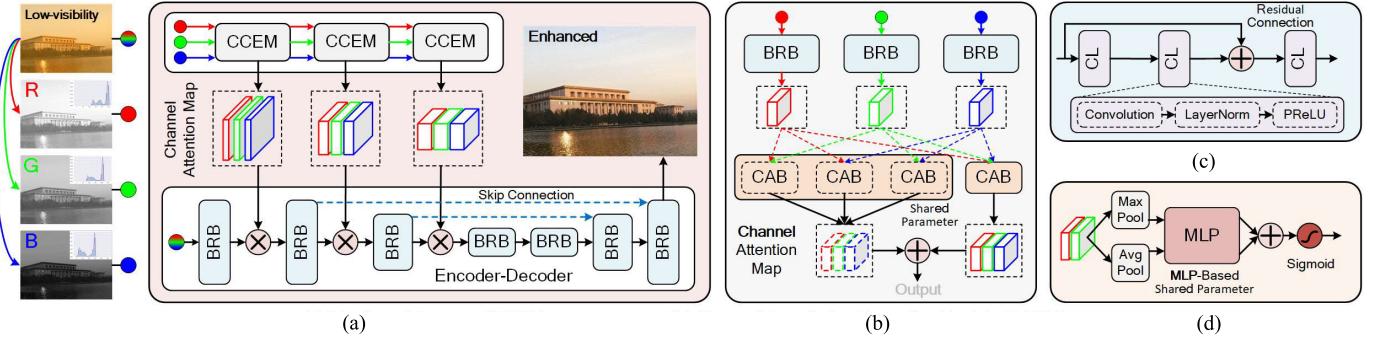


Fig. 2. Flowchart of the two-in-one low-visibility enhancement network (termed TOENet). {8, 16, 32} in (a) are the number of channels of the three scales in the encoder-decoder. (b) Pipeline of the proposed CCEM, which will generate more RGB reconstruction features. (c) and (d) Detailed network structures of the suggested BRB and CAB, respectively.

two parts: the attenuation part of the imaging source and the scattering part of atmospheric light in the environment. Therefore, the process of atmospheric degradation scattering model for the image degradation in hazy weather can be expressed as follows:

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

where x represents the corresponding position of each pixel in the image, J represents a clean (no haze) image, A represents the value of atmospheric light in the environment, t represents the transmittance, and I represents the hazy image. The transmittance t is used to measure the degree of scattering of light by suspended particles in the haze, i.e.,

$$t(x) = e^{-\beta(\lambda) \text{depth}(x)} \quad (2)$$

where λ represents the wavelength, and β represents the atmospheric scattering coefficient, which indicates the different scattering degrees of the particles in the hazy to certain wavelengths of light. β is related to the wavelength; the larger value means a stronger ability to scatter light. depth represents the distance between the imaging device and the target.

The particle radius in the dust storm is much larger than the particle size in the haze. Si et al. [26] propose a novel dust scattering model, and the mathematical expression can be given as follows:

$$I(x) = (J(x) - A + 1)t(x) + A. \quad (3)$$

Simulating more realistic hazy and sandstorm images is one of the important means to improve deep network learning and map low-visibility images to high-quality clear images. The hazy and sandstorm images synthesized by (1) and (3) are used for training the proposed network. And the real-world restoration results of low-visibility images tested by this trained model still achieve satisfactory performance.

III. TOENET: TWO-IN-ONE LOW-VISIBILITY ENHANCEMENT NETWORK

Hazy and sandstorm imaging environments have different atmospheric light values, resulting in obvious differences in the RGB histogram distribution of the collected images. How to learn the correlation among channels of low-quality images is the key to flexibly reconstructing the pixel distribution of hazy and sandstorm images. Therefore, this article proposes a

two-in-one low-visibility enhancement network based on RGB histogram distribution. TOENet uses the encoder-decoder as the backbone and embeds the channel attention map generated by the proposed CCEM, which achieves satisfactory performance on dehazing and sandstorm enhancement tasks with a single network.

A. Basic Modules

Residual learning has demonstrated its efficient performance in different fields of computer vision. So we propose a basic residual block (BRB, $\mathcal{R}(\cdot)$) as the basic unit of the encoder-decoder structure. As shown in Fig. 2(c), a residual block consists of three convolutional layers (CLs, $\mathbf{CL}(\cdot)$). And a CL sequentially contains convolutional operation ($\mathcal{C}(\cdot)$), layer normalization ($\mathcal{N}(\cdot)$) [27], and parametric rectified linear unit (PReLU, $\mathcal{P}(\cdot)$) [28], i.e.,

$$\mathbf{CL}(x_{cl}) = \mathcal{P}(\mathcal{N}(\mathcal{C}(x_{cl}))) \quad (4)$$

where x_{cl} is the input of the CL. LayerNorm can better balance the extraction of single-layer channel features and the association of multilayer channel features in image restoration tasks [29]. As shown in Fig. 2(d), to fully extract the channel-to-channel correlation during learning, we suggest a channel attention block (CAB, $\mathcal{A}(\cdot)$), which can be defined as follows:

$$\mathcal{A}(x_a) = \sigma(\mathbf{MLP}(\text{Avg}(x_a)) + \mathbf{MLP}(\text{Max}(x_a))) \quad (5)$$

where x_a is the input of the CAB, and Avg and Max represent the average pooling and max pooling, respectively. σ is a Sigmoid nonlinear activation. Different from [30], we optimize **MLP** for image restoration tasks and embed LayerNorm to further improve the sensitivity to channel information. As the basic learning modules, BRB and CAB will be suggested to construct a two-in-one enhancer with stable enhancement performance and low computational cost.

B. Channel Correlation Extraction Module

We analyzed the histogram distribution of the three RGB channels of the hazy and sandstorm images. As shown in Fig. 3, compared with the original image, the correlation between the three RGB channels of the synthetic low-visibility images has obvious changes. Therefore, as shown in Fig. 2(b), we propose a CCEM to reconstruct the RGB three-channel

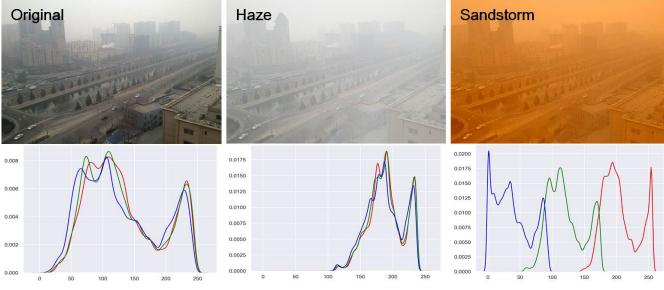


Fig. 3. RGB three-channel histogram distribution of synthetic hazy and sandstorm images.

feature information of the degraded image, which constrains the multiscale features generated by the encoder with channel attention. CCEM consists of two basic modules, i.e., BRB and CAB. After the initial three-channel feature information is mapped by BRB, we establish the correlation between RG, RB, GB, and RGB channels through CAB, i.e.,

$$\begin{cases} A_{rg} = \mathcal{A}_{rgb}(\text{cat}[\mathcal{R}_r(F_r), \mathcal{R}_g(F_g)]) \\ A_{rb} = \mathcal{A}_{rgb}(\text{cat}[\mathcal{R}_r(F_r), \mathcal{R}_b(F_b)]) \\ A_{gb} = \mathcal{A}_{rgb}(\text{cat}[\mathcal{R}_g(F_g), \mathcal{R}_b(F_b)]) \\ A_{rgb} = \mathcal{A}_{RGB}(\text{cat}[\mathcal{R}_r(F_r), \mathcal{R}_g(F_g), \mathcal{R}_b(F_b)]) \end{cases} \quad (6)$$

where F_r , F_g , and F_b represent the feature of the three channels of RGB corresponding to the degraded image. $\text{cat}[\cdot]$ is the concatenation operation of feature maps. The two-channel shared learning \mathcal{A}_{rgb} and three-channel global learning \mathcal{A}_{RGB} will provide richer correlations to reconstruct channel features. A_{rg} , A_{rb} , A_{gb} , and A_{rgb} are the channel correlation attention feature maps. In the end, we concatenate the attention feature maps between the three sets of channels and reinforce the input of the encoder, which can be given as follows:

$$F_{rgb}^{\text{out}} = \mathcal{E}\left(F_{rgb}^{\text{in}} \otimes (\text{cat}[A_{rg}, A_{rb}, A_{gb}] + A_{rgb})\right) \quad (7)$$

where F_{rgb}^{in} represents the feature of the degraded image, \otimes denotes pixel multiplication, \mathcal{E} denotes encoding operation, and F_{rgb}^{out} is the output of a single-scale encoder. CCEM can enhance the sensitivity of the enhancer to channel information and balance image dehazing and sandstorm enhancement.

C. Encoder–Decoder Backbone Network

In this article, we suggest the encoder–decoder network as the backbone and take the BRB as the basic block. As shown in Fig. 2, we combine the generated channel attention maps at three scales with the features, which are multiplied separately to strengthen the encoder’s attention to inter-channel features. We verify in Section IV-C that CCEM can significantly improve the feature mining capability of the encoder. Skip connections are also applied to reduce the phenomenon of network gradient disappearance and speed up network training. The number of channels of the three scales is set to 8, 16, and 32, respectively.

D. Loss Function

To ensure the effect of training, for a given synthetic low-visibility image x_{low} , we use loss \mathcal{L}_{ℓ_2} to minimize the

ℓ_2 -norm between ground truth y and the recovered clean image $\hat{y} = f(x_{\text{low}})$ through TOENet. Mathematically

$$\mathcal{L}_{\ell_2} = \frac{1}{T} \sum_{i=1}^T \|f(x_i) - y_i\|^2 \quad (8)$$

where T is the number of pixels, and i is the index of the pixel. To balance the sensitivity to color during dehazing and sandstorm enhancement learning and inference, we suggest introducing a color loss $\mathcal{L}_{\text{color}}$, which can be given as follows:

$$\mathcal{L}_{\text{color}} = 1 - \frac{1}{T} \sum_{i=1}^T \frac{\langle f(x_i), y_i \rangle}{\|f(x_i)\|_2 \times \|y_i\|_2}. \quad (9)$$

The color loss will be beneficial for increasing the sensitivity of the deep network to the RGB three-channel pixel values, thereby generating clearer images with more natural colors [41]. Therefore, the overall loss function $\mathcal{L}_{\text{total}}$ can be given as follows:

$$\mathcal{L}_{\text{total}} = \gamma_1 \mathcal{L}_{\ell_2} + \gamma_2 \mathcal{L}_{\text{color}} \quad (10)$$

where γ_1 and γ_2 denote the weight value for each loss term, respectively. Extensive experimental results show that $\gamma_1 = 0.8$ and $\gamma_2 = 0.2$ achieve the best quantitative and qualitative performance.

IV. EXPERIMENTS AND DISCUSSION

In this section, we conduct massive experiments to demonstrate the low-visibility enhancement performance of our TOENet. We first introduce the datasets and implementation details. To clearly demonstrate the superiority of TOENet, we conduct quantitative and qualitative comparisons with several state-of-the-art methods on synthetic and real-world hazy and sandstorm images. To evaluate the significance of each module, we conduct the ablation study. Furthermore, to verify the practical application of TOENet, we also conduct the object detection experiment.

A. Datasets and Implementation Details

To examine the performance of the proposed method, we compare TOENet with several state-of-art methods. For dehazing methods, we select DCP [12], CAP [31], HL [32], ROP+ [13], MSCNN [15], AODNet [16], GCANet [33], FFANet [34], and TSDNet [35]. For sandstorm enhancement methods, we select DCP [12], HL [32], LDCP [36], Fusion [37], Retinex [8], ROP+ [13], SDIE [38], Jong [39], and Ako [40]. To quantitatively evaluate the enhancement performance of different methods, we select the reference evaluation metrics peak signal-to-noise ratio (PSNR) [42], structural similarity index measure (SSIM) [43], feature similarity index measure (FSIM) [44], visual-saliency-induced index (VSI) [45], lightness-order-error (LOE) [46], learned perceptual image patch similarity (LPIPS) [47], and the no-reference evaluation metrics natural image quality evaluator (NIQE) [48] and perception-based image quality evaluator (PIQE) [49] as the objective evaluation metrics. For the test dataset, we set 1000 images (excluding the training images) from OTS [50] to synthesize hazy and sandstorm images. We also select some real-world low-visibility images from RTTS [50] and

TABLE I
PERFORMANCE EVALUATION OF OUR AND EXISTING METHODS ON THE 1000 SYNTHETIC HAZY IMAGES.
THE BEST RESULTS ARE IN **BOLD**, AND THE SECOND BEST ARE WITH UNDERLINE

Methods	Dehazing							
	PSNR \uparrow	SSIM \uparrow	FSIM \uparrow	VSI \uparrow	LOE \downarrow	LPIPS \downarrow	NIQE \downarrow	PIQE \downarrow
DCP [12]	15.232 \pm 2.934	0.755 \pm 0.087	0.927 \pm 0.027	0.966 \pm 0.015	327.717 \pm 144.482	0.098 \pm 0.042	3.212 \pm 0.792	8.300 \pm 4.111
CAP [31]	20.337 \pm 3.099	0.864 \pm 0.075	0.958 \pm 0.020	0.985 \pm 0.007	246.873 \pm 135.901	0.052 \pm 0.034	3.410 \pm 0.917	10.613 \pm 6.619
HL [32]	17.983 \pm 3.191	0.813 \pm 0.088	0.932 \pm 0.039	0.966 \pm 0.021	441.060 \pm 187.836	0.080 \pm 0.040	3.230 \pm 0.797	8.708 \pm 4.096
ROP+ [13]	20.296 \pm 3.308	0.866 \pm 0.065	0.958 \pm 0.017	0.981 \pm 0.009	301.012 \pm 149.970	0.051 \pm 0.032	3.067\pm0.802	8.698 \pm 4.355
MSCNN [15]	18.271 \pm 3.833	0.840 \pm 0.098	0.958 \pm 0.023	0.983 \pm 0.009	290.535 \pm 141.392	0.067 \pm 0.048	3.409 \pm 0.917	10.636 \pm 6.683
AODNet [16]	19.564 \pm 2.542	0.859 \pm 0.073	0.922 \pm 0.026	0.980 \pm 0.009	294.345 \pm 142.462	0.061 \pm 0.037	3.366 \pm 0.911	10.872 \pm 5.617
GCANet [33]	21.111 \pm 4.310	0.878 \pm 0.077	0.959 \pm 0.028	0.984 \pm 0.013	329.507 \pm 196.266	0.046 \pm 0.030	3.708 \pm 0.769	8.529 \pm 4.797
FFANet [34]	21.741 \pm 5.351	0.896 \pm 0.075	0.969 \pm 0.023	0.990 \pm 0.008	180.599 \pm 105.599	0.043 \pm 0.036	3.415 \pm 0.944	8.888 \pm 5.680
TSDNet [35]	23.777 \pm 3.777	0.920 \pm 0.053	0.977 \pm 0.013	0.990 \pm 0.006	231.868 \pm 117.771	0.039 \pm 0.020	3.184 \pm 0.808	7.911\pm4.166
TOENet	24.273\pm3.122	<u>0.934\pm0.057</u>	<u>0.979\pm0.018</u>	<u>0.991\pm0.006</u>	176.705\pm112.802	0.032\pm0.022	3.171 \pm 0.837	8.108 \pm 4.302

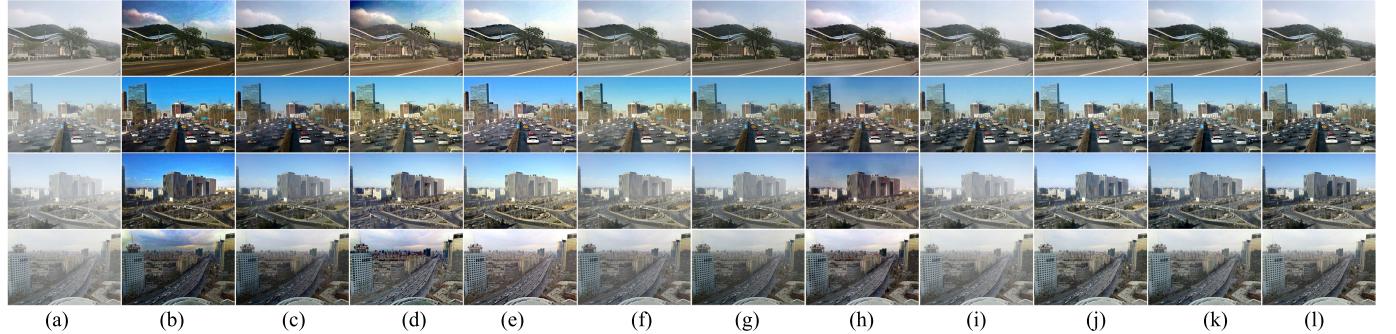


Fig. 4. Visual comparison between the results of the proposed and existing (b) DCP [12], (c) CAP [31], (d) HL [32], (e) ROP+ [13], (f) MSCNN [15], (g) AODNet [16], (h) GCANet [33], (i) FFANet [34], (j) TSDNet [35], (k) proposed TOENet, and (l) Ground Truth on (a) four synthetic hazy images.

the Internet. The TOENet is trained for 100 epochs with 2000 images. The network parameters are updated by the Adam optimizer. The initial learning rate is 0.001 and is cut to one-tenth after every 30 epochs. The TOENet model is trained and tested in the Python 3.7 environment using the PyTorch software package with an Intel¹ Core² i9-13900K CPU @5.80 GHz and Nvidia GeForce RTX 3080 GPU.

B. Analysis on Synthetic and Real-World Low-Visibility Images

1) *Synthetic Haze*: We first conduct the experiment on 1000 synthetic hazy images. As shown in Table I, on the six reference evaluation metrics, the proposed TOENet ranks first place. Meantime, the no-reference metrics results of TOENet are still competitive. To judge the results enhanced by different methods more intuitively, we also show some visual comparisons in Fig. 4, MSCNN and FFANet cannot successfully remove haze thoroughly. The saturation of DCP is too high, which leads to serious color distortion. Although HL performs better in color restoration compared with DCP, the color is still unnatural in some areas of the sky. In the meantime, ROP+ suffers from the overexposure problem. Relatively speaking, CAP, AODNet, GCANet, and TSDNet get satisfactory results. However, they still should be improved in detail reservation and color restoration. Compared with the methods mentioned above, the proposed TOENet is most similar to ground truth in terms of color, structure, contrast, etc.

We then show the detail restoration ability of TOENet on the transportation-related images. As the local magnification shown in Fig. 5, it can be clearly found that the TOENet achieves better performance on both color restoration and detail preservation compared with the state-of-art methods. To further demonstrate the superiority of TOENet over the SOTA method (TSDNet), we conducted a detailed comparison as shown in Fig. 6. The results indicate that while the SOTA method achieves similar enhancement effects to TOENet in some images, a closer examination of the detail comparison reveals that our method produces more intricate details and colors that are closer to the ground truth. And the image definition after TOENet enhancement is higher.

2) *Synthetic Sandstorm*: To demonstrate the versatility of the TOENet, we also conduct experiments on sandstorm images. Similar to the dehazing experiment, we first compare the enhancement performance of the state-of-art methods on 1000 synthetic sandstorm images. As shown in Table II, the TOENet still ranks first place in terms of reference metrics results. In addition, the no-reference metrics performance is also satisfactory. And as shown in Fig. 7, the traditional dehazing-based methods such as DCP, MSCNN, and LDCP tend to restore the image from the prominent structures but have shortcomings in suppressing the impact of dust. This is because the imaging theory of hazy and sandstorm images is different [26]. Once the assumed prior knowledge is accurate, the enhancement results will be unsatisfactory. In addition, SDIE and Jong cannot remove the dust thoroughly. And Ako has the insufficient recovery of color, which is more obvious when the concentrations of dust storms increase. Compared

¹Registered trademark.

²Trademarked.

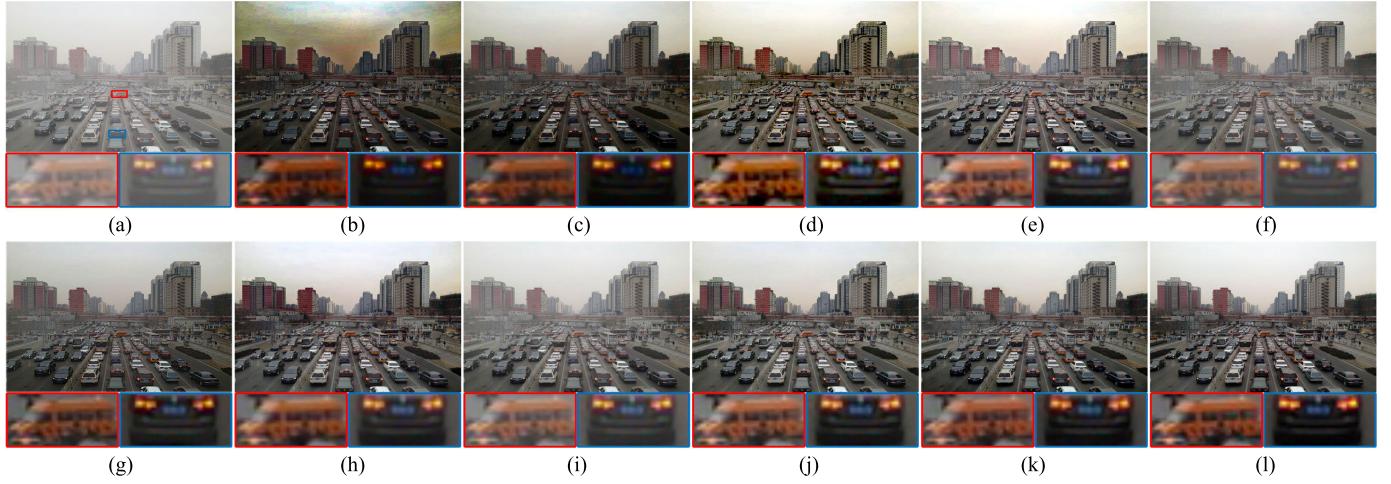


Fig. 5. Visual comparison of different enhancement methods for one typical hazy image. From the top left to the bottom right are (a) hazy image, restored images, enhanced by (b) DCP [12], (c) CAP [31], (d) HL [32], (e) ROP+ [13], (f) MSCNN [15], (g) AODNet [16], (h) GCANet [33], (i) FFANet [34], (j) TSDNet [35], (k) proposed TOENet, and (l) Ground Truth, respectively.



Fig. 6. Detailed visual comparison between the results of the proposed method and the SOTA methods in Figs. 4 and 7. The first three columns are the dehazing results: (a) TSDNet [35], (b) proposed TOENet, and (c) Ground Truth. The last three columns are the sandstorm enhancement results: (d) Ako [40], (e) proposed TOENet, and (f) Ground Truth.

with the methods mentioned above, the proposed TOENet can enhance the image with the best color restoration and detail preservation.

As the local magnification shown in Fig. 8, TOENet also achieves satisfactory performance. In the meantime, as shown in Fig. 6, TOENet preserves details better compared with the SOTA method (Ako).

3) Real-World: To better demonstrate the superiority of TOENet in practical application, we also show the results on real-world 300 hazy and 300 sandstorm images. As shown in Table III, according to the average metric results, the proposed TOENet achieves superior performance. The non-referenced metrics indicate that our method has better image quality compared with the state-of-the-art methods.

TABLE II
PERFORMANCE EVALUATION OF OUR AND EXISTING METHODS ON THE 1000 SYNTHETIC SANDSTORM IMAGES.
THE BEST RESULTS ARE IN **BOLD**, AND THE SECOND BEST ARE WITH UNDERLINE

Methods	Sandstorm enhancement							
	PSNR \uparrow	SSIM \uparrow	FSIM \uparrow	VSI \uparrow	LOE \downarrow	LPIPS \downarrow	NIQE \downarrow	PIQE \downarrow
DCP [12]	10.049 \pm 1.788	0.532 \pm 0.096	0.869 \pm 0.047	0.901 \pm 0.025	700.504 \pm 467.531	0.355 \pm 0.086	3.461 \pm 0.881	11.065 \pm 7.279
HL [32]	12.798 \pm 2.313	0.695 \pm 0.096	0.949 \pm 0.030	0.931 \pm 0.021	529.572 \pm 301.347	0.284 \pm 0.088	3.313 \pm 0.906	9.743 \pm 5.996
LDCP [36]	11.172 \pm 1.997	0.641 \pm 0.115	0.897 \pm 0.049	0.926 \pm 0.026	357.635 \pm 286.739	0.306 \pm 0.102	3.756 \pm 1.026	11.485 \pm 7.777
Fusion [37]	20.057 \pm 3.207	0.807 \pm 0.101	0.930 \pm 0.041	0.967 \pm 0.020	282.387 \pm 140.615	0.202 \pm 0.073	3.333 \pm 0.814	9.842 \pm 4.338
Retinex [8]	17.904 \pm 2.694	0.780 \pm 0.085	0.906 \pm 0.032	0.960 \pm 0.016	301.182 \pm 142.404	0.213 \pm 0.044	3.556 \pm 0.599	21.185 \pm 9.066
ROP+ [13]	18.895 \pm 2.886	0.824 \pm 0.077	<u>0.960\pm0.019</u>	0.966 \pm 0.015	297.929 \pm 137.193	<u>0.142\pm0.081</u>	<u>3.061\pm0.780</u>	<u>8.717\pm4.492</u>
SDIE [38]	17.589 \pm 2.338	<u>0.830\pm0.062</u>	0.927 \pm 0.037	0.960 \pm 0.017	399.731 \pm 280.257	0.162 \pm 0.049	3.770 \pm 1.241	10.110 \pm 6.213
Jong [39]	18.258 \pm 2.993	0.790 \pm 0.078	0.947 \pm 0.021	0.965 \pm 0.014	289.451 \pm 158.966	0.158 \pm 0.053	3.390 \pm 0.875	11.331 \pm 7.112
Ako [40]	17.559 \pm 2.338	0.759 \pm 0.089	0.941 \pm 0.022	0.965 \pm 0.018	345.411 \pm 189.842	0.143 \pm 0.999	2.877\pm0.698	8.214\pm7.161
TOENet	20.600\pm4.400	0.841\pm0.103	0.965\pm0.022	0.968\pm0.021	276.448\pm198.597	0.132\pm0.102	3.296 \pm 0.847	9.578 \pm 6.025

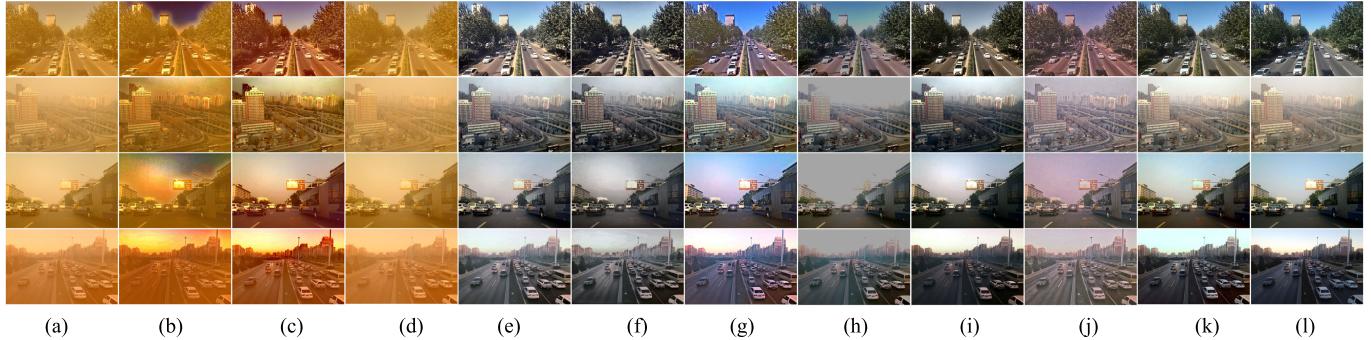


Fig. 7. Visual comparison between the results of the proposed and existing (b) DCP [12], (c) HL [32], (d) LDCP [36], (e) Fusion [37], (f) Retinex [8], (g) ROP+ [13], (h) SDIE [38], (i) Jong [39], (j) Ako [40], (k) proposed TOENet, and (l) Ground Truth on (a) four synthetic sandstorm images.

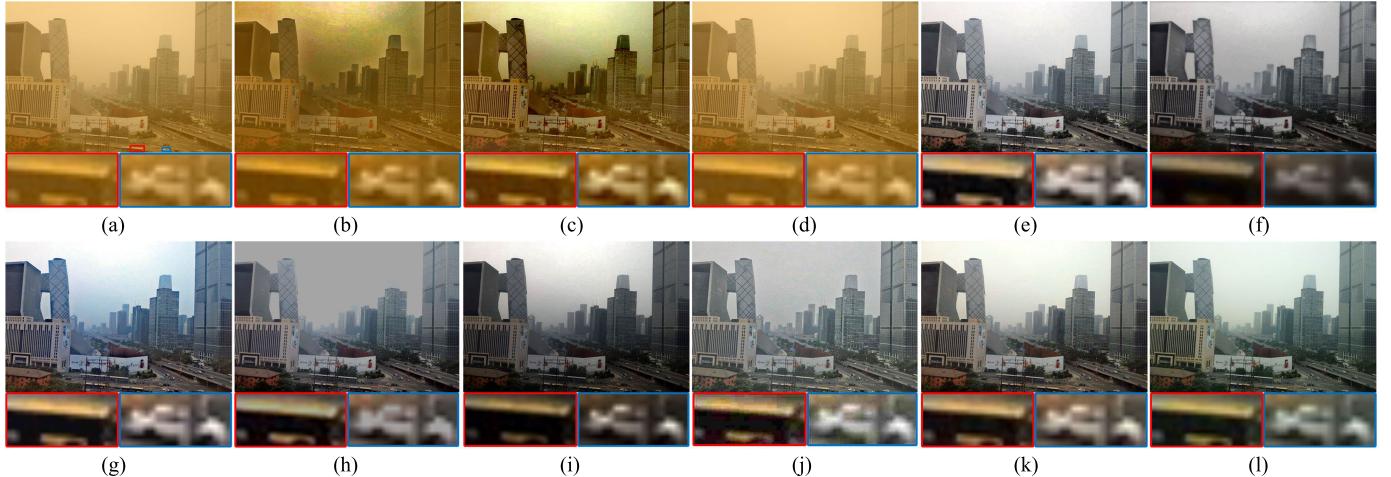


Fig. 8. Visual comparison of different enhancement methods for one typical sandstorm image. From the top left to the bottom right are (a) sandstorm image, restored images, enhanced by (b) DCP [12], (c) HL [32], (d) LDCP [36], (e) Fusion [37], (f) Retinex [8], (g) ROP+ [13], (h) SDIE [38], (i) Jong [39], (j) Ako [40], (k) proposed TOENet, and (l) Ground Truth, respectively.

Meanwhile, as shown in Fig. 9, even in complex real-world environments, the TOENet can still work effectively.

As discussed above, TOENet outperforms better than other methods, whether on synthetic or real-world images. This superior performance can be attributed to the astute analysis of the distribution characteristics of RGB channels in hazy and sandstorm images and the utilization of CCEM module to extract pertinent channel information for achieving universal image enhancement.

C. Ablation Study

1) *Channel Correlation Extraction Module*: We also conduct experiments to verify the rationality of elaborately designed parts in the channel correction extraction module. As shown in Table IV, when the CCEM is lost in the network, the PSNR and SSIM results are the worst. When the network uses CCEM to constrain the multiscale features generated by the encoder, the metrics results have an inspiring improvement. What is more, when the CAB is used

TABLE III

PERFORMANCE EVALUATION OF OUR AND EXISTING METHODS ON THE 300 REAL-WORLD HAZY AND SANDSTORM IMAGES.
THE BEST RESULTS ARE IN **BOLD**, AND THE SECOND BEST ARE WITH UNDERLINE

Methods	Dehazing		Methods	Sandstorm enhancement	
	NIQE ↓	PIQE ↓		NIQE ↓	PIQE ↓
DCP [12]	3.996±1.208	8.316±3.499	DCP [12]	3.638±0.495	21.621±7.495
CAP [31]	4.330±1.744	10.075±4.882	HL [32]	3.394±0.403	17.693±6.227
HL [32]	3.904±1.175	9.499±2.957	LDCP [36]	3.876±0.648	20.516±9.415
ROP+ [13]	3.812±1.140	7.589±2.433	Fusion [37]	3.290±0.366	16.538±6.509
MSCNN [15]	4.056±1.304	8.633±3.786	Retinex [8]	3.773±0.385	31.042±7.600
AODNet [16]	4.136±1.208	9.830±3.289	ROP+ [13]	3.328±0.392	16.530±6.102
GCANet [33]	4.091±1.112	7.742±3.713	SDIE [38]	3.381±0.408	17.826±6.547
FFANet [34]	4.166±1.299	8.638±4.600	Jong [39]	4.046±0.468	21.154±10.235
TSDNet [35]	3.912±1.190	7.115±2.646	Ako [40]	3.345±0.385	7.789±4.557
TOENet	3.784±1.116	7.460±2.139	TOENet	3.279±0.289	15.718±5.713



Fig. 9. Visual results on real-world low-visibility images. From top to bottom are (a) low-visibility images, enhanced by (b) DCP [12], (c) HL [32], (d) LDCP [36], (e) ROP+ [13], and (f) proposed TOENet, respectively.

TABLE IV
ABLATION STUDY ON CCEM. THE BEST RESULTS ARE IN **BOLD**

Number	Methods					Dehazing		Sandstorm enhancement	
	CCEM	CAB	R	G	B	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
(a)	✗	✗	✗	✗	✗	21.113±3.126	0.892±0.039	18.256±2.998	0.855±0.052
(b)	✓	✗	✓	✓	✓	21.834±3.145	0.901±0.040	19.635±3.145	0.866±0.048
(c)	✓	✓	✓	✗	✗	22.277±2.796	0.907±0.041	20.651±2.974	0.877±0.047
(d)	✓	✓	✗	✓	✗	23.223±2.660	0.911±0.033	20.146±2.598	0.872±0.045
(e)	✓	✓	✗	✗	✓	23.292±3.003	0.916±0.042	20.286±2.972	0.876±0.049
(f)	✓	✓	✓	✓	✗	23.344±3.223	0.922±0.042	20.998±3.264	0.894±0.063
(g)	✓	✓	✓	✗	✓	23.411±3.330	0.926±0.041	20.733±3.360	0.881±0.084
(h)	✓	✓	✗	✓	✓	23.397±2.666	0.920±0.040	20.644±3.304	0.883±0.050
(i)	✓	✓	✓	✓	✓	24.042±2.684	0.932±0.031	21.700±3.543	0.906±0.012

to restrain the extraction channel feature, the enhancement performance also has an improvement. We also present the visual effects of each ablation experiment in Fig. 10. It is worth mentioning that a major innovation of this work is the analysis of the patterns of hazy and sandstorm images in the RGB channels (Fig. 3) and the design of a rational module for high-quality image restoration and general image enhancement. This conclusion is drawn based on the analysis of a large number of low-visibility images rather than treating deep learning as a black box for image

enhancement. As shown in Fig. 10, solely using the traditional encoder-decoder structure to enhance both types of images simultaneously does not yield satisfactory results. Furthermore, each component of CCEM indeed has an impact on the enhancement effect. The network achieves the best result when the RGB information is used together, which demonstrates that channel correction plays a significant role in the proposed TOENet.

2) *Three Training Strategies*: An attractive point of TOENet is that it can enhance different low-visibility images in the

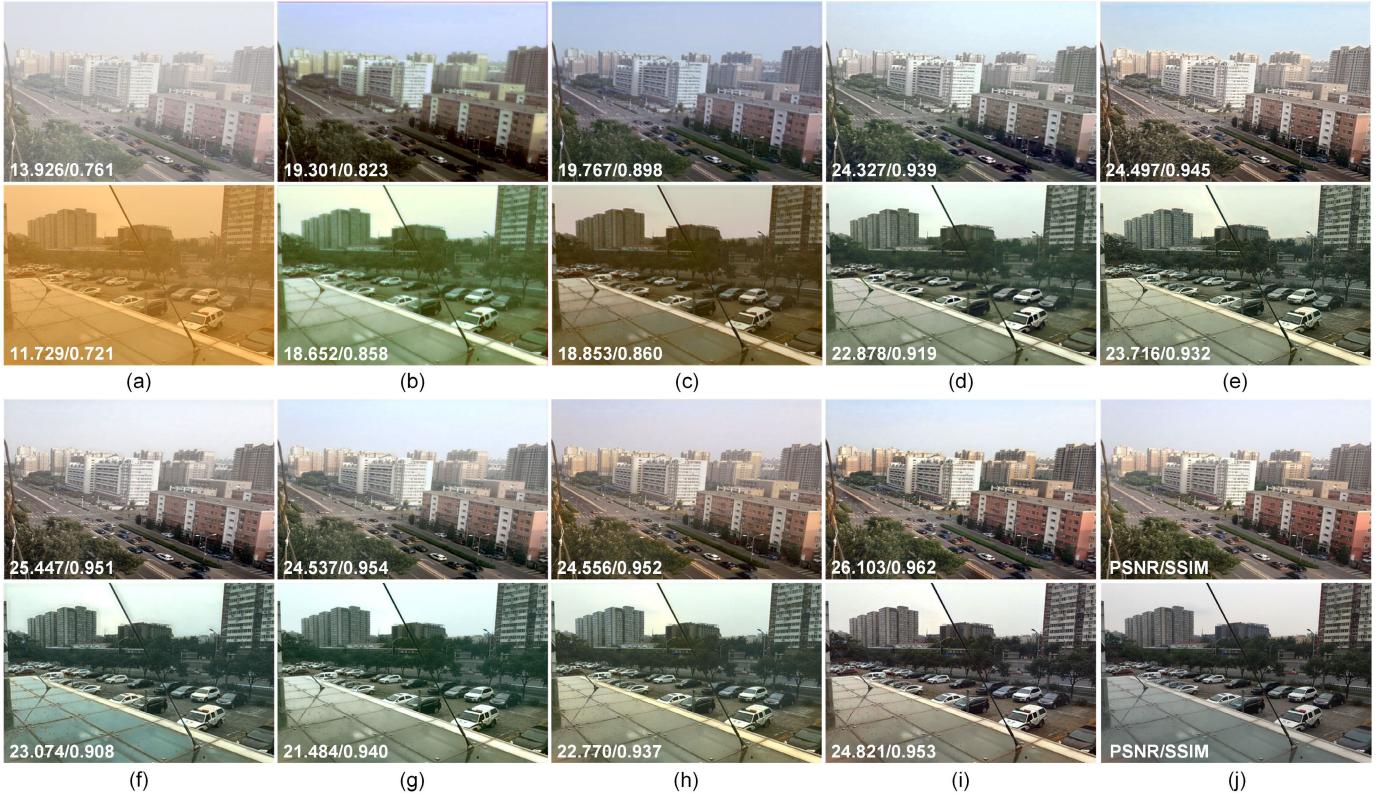


Fig. 10. Visual comparisons of the ablation study on CCEM. From the top left to the bottom right are (a) only encoder-decoder structure, (b) aided by CCEM without CAB, (c) aided by CCEM with CAB of R channel information, (d) aided by CCEM with CAB of G channel information, (e) aided by CCEM with CAB of B channel information, (f) aided by CCEM with CAB of RG channel information, (g) aided by CCEM with CAB of RB channel information, (h) aided by CCEM with CAB of GB channel information, (i) proposed TOENet, and (j) Ground Truth, respectively.

TABLE V

ABLATION STUDY ON THREE TRAINING STRATEGIES

Training Strategy	Dehazing		Sandstorm enhancement	
	Hazy	Sandstorm	PSNR ↑	SSIM ↑
✓ X	24.501±3.665	0.941±0.042	—	—
X ✓	—	—	22.364±3.323	0.912±0.057
✓ ✓	24.042±2.684	0.932±0.031	21.700±3.543	0.906±0.012

TABLE VII

COMPARISON OF RUNNING TIME (UNIT: SECOND PER IMAGE) BETWEEN THE TOENET AND THE STATE-OF-ART METHODS ON 2K IMAGE

Methods	Hazy	Sandstorm	Platform	Size (KB)	Time (Sec.)	FLOPs (G)
DCP [12]	✓	✓	Matlab	—	5.462	—
Fusion [37]	X	✓	Matlab	—	1.713	—
Retinex [8]	X	✓	Matlab	—	1.923	—
CAP [31]	✓	X	Matlab	—	1.971	—
HL [32]	✓	✓	Matlab	—	9.232	—
LDCP [36]	✓	✓	Matlab	—	5.213	—
ROP+ [13]	✓	✓	Matlab	—	0.659	—
SDIE [38]	X	✓	Python	—	0.618	—
Jong [39]	X	✓	Python	—	1.311	—
Ako [40]	X	✓	Matlab	—	4.679	—
MSCNN [15]	✓	X	Matlab	—	3.576	—
AODNet [16]	✓	X	Pytorch	9	0.067	3.218
GCANet [33]	✓	X	Pytorch	2758	1.614	522.142
FFANet [34]	✓	X	Pytorch	25999	5.008	8086.870
TSDNet [35]	✓	X	Pytorch	14275	0.014	675.142
TOENet	✓	✓	Pytorch	2557	0.006	52.347

TABLE VI
ABLATION STUDY ON THE LOSS FUNCTION.
THE BEST RESULTS ARE IN **BOLD**

\mathcal{L}_{ℓ_2}	$\mathcal{L}_{\text{color}}$	Dehazing		Sandstorm enhancement	
		PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
0.1	0.9	20.230±1.502	0.805±0.052	17.269±2.174	0.770±0.033
0.2	0.8	23.074±2.373	0.882±0.027	20.222±2.669	0.830±0.008
0.3	0.7	23.652±3.276	0.911±0.038	20.358±2.841	0.867±0.019
0.4	0.6	23.520±2.442	0.895±0.035	20.207±3.216	0.860±0.016
0.5	0.5	23.255±2.300	0.889±0.038	20.844±2.962	0.866±0.019
0.6	0.4	23.717±3.266	0.918±0.035	21.203±2.979	0.891±0.016
0.7	0.3	23.977±2.961	0.920±0.025	21.235±3.096	0.905±0.006
0.9	0.1	23.414±2.801	0.915±0.030	20.319±2.715	0.890±0.011
0.8	0.2	24.042±2.684	0.932±0.031	21.700±3.543	0.906±0.012

two-in-one network. To further explore the enhancement performance of TOENet to different visual enhancement tasks, we train the network under different strategies. As shown in Table V, TOENet also has a satisfactory performance under different training strategies, which demonstrates that TOENet finds the connection between different low-visibility images. Therefore, this work reveals that it may be possible to handle all low-visibility images in a method using the

connection between them, which may be more effective in the real world, especially when the low-visibility types cannot be accurately judged in complex environments. And this article leverages the correlation between hazy and sandstorm images, enabling them to be processed within a single network with remarkable results. Without the correlation, the network would be unable to simultaneously enhance both types of images. While the overall performance of the trained universal image

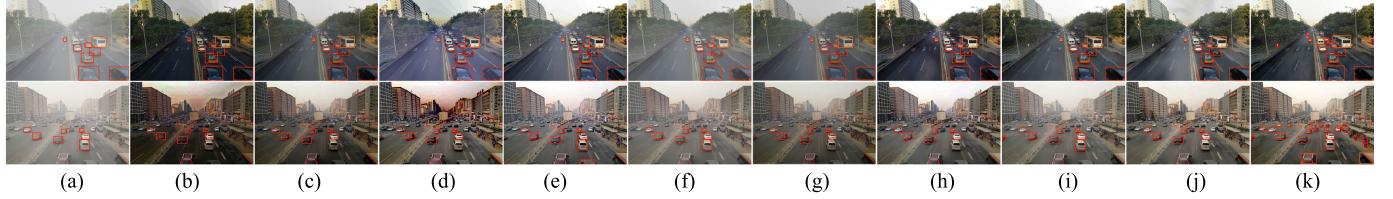


Fig. 11. Results of the object detection experiment on hazy traffic monitoring images. From left to right are: (a) Hazy, (b) DCP [12], (c) CAP [31], (d) HL [32], (e) ROP+ [13], (f) MSCNN [15], (g) AODNet [16], (h) GCANet [33], (i) FFANet [34], (j) TSDNet [35], and (k) proposed TOENet, respectively.

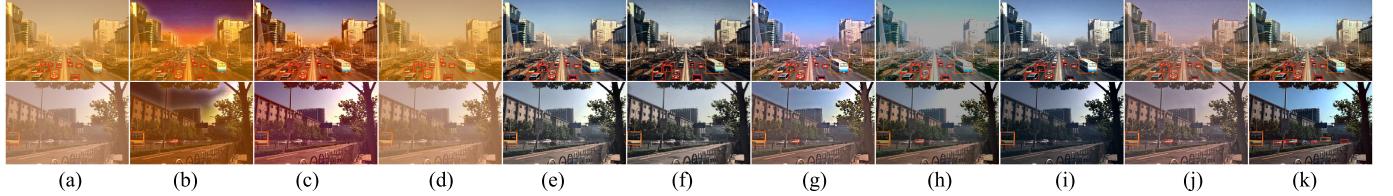


Fig. 12. Results of the object detection experiment on sandstorm traffic monitoring images. From left to right are: (a) Sandstorm, (b) DCP [12], (c) HL [32], (d) LDCP [36], (e) Fusion [37], (f) Retinex [8], (g) ROP+ [13], (h) SDIE [38], (i) Jong [39], (j) Ako [40], and (k) proposed TOENet, respectively.

TABLE VIII
DETECTION AVERAGE AP OF YOLOX ON THE TRANSPORTATION-RELATED SYNTHETIC HAZY IMAGES FROM COCO2017
AND THE IMAGES ENHANCED BY VARIOUS METHODS. THE BEST RESULTS ARE IN **BOLD**

Methods	Dehazing								
	Aeroplane	Bicycle	Boat	Bus	Car	Motorbike	Person	Train	AP
Hazy	0.756	0.874	0.500	0.920	0.900	0.844	0.828	0.760	0.797
DCP [12]	0.873	0.920	0.651	0.966	0.874	0.887	0.852	0.895	0.865
CAP [31]	0.767	0.888	0.562	0.921	0.877	0.834	0.824	0.809	0.810
HL [32]	0.806	0.874	0.621	0.935	0.828	0.882	0.811	0.846	0.826
ROP+ [13]	0.789	0.875	0.593	0.939	0.823	0.886	0.807	0.845	0.820
MSCNN [15]	0.760	0.882	0.569	0.936	0.871	0.832	0.820	0.817	0.811
AODNet [16]	0.885	0.921	0.647	0.951	0.912	0.886	0.870	0.869	0.868
GCANet [33]	0.873	0.918	0.662	0.961	0.887	0.911	0.873	0.895	0.872
FFANet [34]	0.660	0.800	0.382	0.824	0.805	0.692	0.704	0.645	0.689
TSDNet [35]	0.862	0.882	0.592	0.940	0.871	0.840	0.817	0.836	0.830
TOENet	0.946	0.926	0.734	0.964	0.914	0.941	0.893	0.913	0.906
Ground Truth	0.987	0.965	0.846	0.985	0.927	0.952	0.939	0.949	0.944

enhancement model shows a slight decrease compared with the models trained on individual tasks, it still surpasses the existing state-of-the-art methods in terms of enhancement capability. Thus, this enhancement strategy is deemed acceptable, trading off a marginal performance loss for a more intelligent approach to image enhancement, which can be more practical in real-world ITMS applications.

3) *Empirical Values in the Loss Function:* To verify the effectiveness of the empiric values in the loss function, we conduct the ablation experiment on the design of the loss function. Specifically, we set the weight of each loss differently in the training period and show the quantitative result in Table VI. The result demonstrates that the network's current weights are more effective in supervising the network and achieving more satisfactory results.

D. Running Time Comparisons

To prove the advantage of TOENet in computational efficiency, we conduct comparisons of the running time and FLOPs on both dehazing and sandstorm enhancement methods. As shown in Table VII, the size of TOENet is only 2557 KB. And it ranks first in running time among the state-of-art methods, with only taking 0.006 s on GPU to process a 2K (2560×1440) image. In addition, the FLOPs of

TOENet is still competitive, which is suitable for edge devices in ITMS.

E. Object Detection Accuracy Comparisons

To further demonstrate the benefits of our TOENet in practical application, we apply YOLOX [51] to detect objects on both low-visibility images and the enhanced by the enhancement methods. As shown in Figs. 11 and 12, whether in hazy or sandstorm scenes, the detection network cannot guarantee accurate object detection due to poor image quality. After the enhancement, the YOLOX can detect more objects. What is more, the enhanced image of TOENet performs better compared with the state-of-the-art methods. In addition, we also quantitatively compare our TOENet with various methods for the improvement for the high-level vision task. Specially, we choose the transportation-related images from the COCO2017 [52] dataset (ground truth) to synthesize low-visibility images. We display the comparison in Tables VIII and IX in terms of AP. The results show that the detection results enhanced by our TOENet achieve the closest AP results to the clear image. It demonstrates that TOENet is more effective for the higher level visual task under low-visibility environments compared with previous works, which is more beneficial to the ITMS.

TABLE IX

DETECTION AVERAGE AP OF YOLOX ON THE TRANSPORTATION-RELATED SYNTHETIC SANDSTORM IMAGES FROM COCO2017 AND THE IMAGES ENHANCED BY VARIOUS METHODS. THE BEST RESULTS ARE IN **BOLD**

Methods	Sandstorm enhancement								AP
	Aeroplane	Bicycle	Boat	Bus	Car	Motorbike	Person	Train	
Sandstorm	0.761	0.877	0.491	0.865	0.874	0.858	0.810	0.745	0.785
DCP [12]	0.880	0.930	0.650	0.939	0.890	0.918	0.868	0.860	0.867
HL [32]	0.799	0.919	0.630	0.932	0.879	0.908	0.850	0.858	0.847
LDCP [36]	0.533	0.752	0.318	0.746	0.733	0.675	0.629	0.584	0.621
Fusion [37]	0.779	0.896	0.598	0.900	0.825	0.878	0.829	0.847	0.819
Retinex [8]	0.602	0.881	0.537	0.929	0.851	0.880	0.854	0.848	0.800
ROP+ [13]	0.827	0.885	0.602	0.890	0.815	0.872	0.815	0.843	0.818
SDIE [38]	0.778	0.881	0.577	0.884	0.801	0.855	0.793	0.846	0.802
Jong [39]	0.793	0.910	0.598	0.910	0.842	0.899	0.833	0.866	0.832
Ako [40]	0.821	0.893	0.575	0.875	0.813	0.874	0.809	0.797	0.807
TOENet	0.936	0.933	0.745	0.946	0.912	0.928	0.896	0.907	0.900
Ground Truth	0.987	0.965	0.846	0.985	0.927	0.952	0.939	0.949	0.944

V. CONCLUSION

This article proposes a two-in-one low-visibility enhancer (termed TOENet) for the ITMS, which successfully restores hazy and sandstorm images by a single network. The proposed CCEM reconstructs the relationship among the three RGB channels of damaged images. The multiscale channel attention features generated by CCEM are embedded into the backbone (i.e., the encoder-decoder network) of TOENet to extract additional latent feature information from low-visibility images. Extensive experimental results show that TOENet achieves superior enhancement results on both synthetic and real-world low-visibility test datasets. Besides, the running time and object detection experiments demonstrate that our method contributes more to higher level image analysis tasks under low-visibility environments in ITMS.

In conclusion, this article presents a two-in-one low-visibility enhancer for ITMS. Although our work achieves promising performance, it still faces several challenges. The further improvement of our method includes the following.

- 1) Although our TOENet successfully enhances hazy and sandstorm images in a single network, simultaneous enhancement of more types of images is currently not applicable. In future work, we will deeply analyze the features of more low-visibility images and attempt to enhance more types of images in a single network, which will be more meaningful in ITMS.
- 2) To overcome the inadequacy of paired hazy and sandstorm datasets, we consider introducing semi-supervised architecture and GANs in our future work.
- 3) Currently, the publicly available paired sandstorm datasets are few. In future work, we plan to collect more images of sandstorms in real-world scenes, create, and release large-scale paired datasets to further advance deep-learning-based sandstorm image enhancement research.

REFERENCES

- [1] A. Mehra, M. Mandal, P. Narang, and V. Chamola, “ReViewNet: A fast and resource optimized network for enabling safe autonomous driving in hazy weather conditions,” *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4256–4266, Jul. 2021.
- [2] B. Yu, Y. Chen, S.-Y. Cao, H.-L. Shen, and J. Li, “Three-channel infrared imaging for object detection in haze,” *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–13, Apr. 2022.
- [3] Y. Guo, Y. Lu, and R. W. Liu, “Lightweight deep network-enabled real-time low-visibility enhancement for promoting vessel detection in maritime video surveillance,” *J. Navigat.*, vol. 75, no. 1, pp. 230–250, Jan. 2022.
- [4] L. Wang, H. Qin, X. Zhou, X. Lu, and F. Zhang, “R-YOLO: A robust object detector in adverse weather,” *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, Dec. 2022.
- [5] S. M. Pizer et al., “Adaptive histogram equalization and its variations,” *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, Sep. 1987.
- [6] Y.-T. Kim, “Contrast enhancement using brightness preserving bi-histogram equalization,” *IEEE Trans. Consum. Electron.*, vol. 43, no. 1, pp. 1–8, Feb. 1997.
- [7] E. H. Land, “The retinex theory of color vision,” *Sci. Amer.*, vol. 237, no. 6, pp. 108–129, Dec. 1977.
- [8] X. Fu, P. Zhuang, Y. Huang, Y. Liao, X.-P. Zhang, and X. Ding, “A retinex-based enhancing approach for single underwater image,” in *Proc. IEEE ICIP*, 2014, pp. 4572–4576.
- [9] H. Wu and Z. Tan, “An image dehazing algorithm based on single-scale retinex and homomorphic filtering,” in *Proc. CSPS*. Springer, 2019, pp. 1482–1493.
- [10] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi, and N. Mazur, “A novel fast single image dehazing algorithm based on artificial multiexposure image fusion,” *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–23, Sep. 2020.
- [11] Z. Shi, Y. Feng, M. Zhao, E. Zhang, and L. He, “Normalised gamma transformation-based contrast-limited adaptive histogram equalisation with colour correction for sand-dust image enhancement,” *IET Image Process.*, vol. 14, no. 4, pp. 747–756, Feb. 2020.
- [12] 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.
- [13] J. Liu, R. W. Liu, J. Sun, and T. Zeng, “Rank-one prior: Real-time scene recovery,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 7, pp. 8845–8860, Dec. 2022.
- [14] 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.
- [15] 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. ECCV*. Springer, 2016, pp. 154–169.
- [16] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “AOD-Net: All-in-one dehazing network,” in *Proc. IEEE ICCV*, Oct. 2017, pp. 4770–4778.
- [17] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *Proc. IEEE CVPR*, 2017, pp. 2261–2269.
- [18] Y. Guo et al., “SCANet: Self-paced semi-curricular attention network for non-homogeneous image dehazing,” in *Proc. IEEE CVPRW*, 2023, pp. 1884–1893.
- [19] B. Li, X. Liu, P. Hu, Z. Wu, J. Lv, and X. Peng, “All-in-one image restoration for unknown corruption,” in *Proc. IEEE CVPR*, 2022, pp. 17431–17441.

- [20] J. Qu, Y. Gao, Y. Lu, W. Xu, and R. W. Liu, "Deep learning-driven surveillance quality enhancement for maritime management promotion under low-visibility weathers," *Ocean Coastal Manage.*, vol. 235, p. 106478, Mar. 2023.
- [21] Y. Dong, Y. Liu, H. Zhang, S. Chen, and Y. Qiao, "FD-GAN: Generative adversarial networks with fusion-discriminator for single image dehazing," in *Proc. AAAI*, 2020, vol. 34, no. 7, pp. 10729–10736.
- [22] X. Chai, J. Zhou, H. Zhou, and J. Lai, "PDD-GAN: Prior-based GAN network with decoupling ability for single image dehazing," in *Proc. ACM*, 2022, pp. 5952–5960.
- [23] X. Chen, Y. Li, C. Kong, and L. Dai, "Unpaired image dehazing with physical-guided restoration and depth-guided refinement," *IEEE Signal Process. Lett.*, vol. 29, pp. 587–591, Feb. 2022.
- [24] Y. Yang, H. Zhang, X. Wu, and X. Liang, "MSTFDN: Multi-scale transformer fusion dehazing network," *Appl. Intell.*, vol. 53, pp. 5951–5962, Jul. 2022.
- [25] Y. Song, Z. He, H. Qian, and X. Du, "Vision transformers for single image dehazing," Apr. 2022, *arXiv:2204.03883*.
- [26] Y. Si, F. Yang, Y. Guo, W. Zhang, and Y. Yang, "A comprehensive benchmark analysis for sand dust image reconstruction," Feb. 2022, *arXiv:2202.03031*.
- [27] J. Lei Ba, J. Ryan Kiros, and G. E. Hinton, "Layer normalization," Jul. 2016, *arXiv:1607.06450*.
- [28] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE ICCV*, Dec. 2015, pp. 1026–1034.
- [29] Y. Lu, Y. Guo, R. W. Liu, and W. Ren, "MTRBNet: Multi-branch topology residual block-based network for low-light enhancement," *IEEE Signal Process. Lett.*, vol. 29, pp. 1127–1131, Mar. 2022.
- [30] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. ECCV*, 2018, pp. 3–19.
- [31] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3522–3533, Nov. 2015.
- [32] D. Berman, T. Treibitz, and S. Avidan, "Air-light estimation using haze-lines," in *Proc. IEEE ICIP*, 2017, pp. 1–9.
- [33] D. Chen et al., "Gated context aggregation network for image dehazing and deraining," in *Proc. IEEE WACV*, 2019, pp. 1375–1383.
- [34] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, "FFA-Net: Feature fusion attention network for single image dehazing," in *Proc. AAAI*, 2020, vol. 34, no. 7, pp. 11908–11915.
- [35] R. W. Liu, Y. Guo, Y. Lu, K. T. Chui, and B. B. Gupta, "Deep network-enabled haze visibility enhancement for visual IoT-driven intelligent transportation systems," *IEEE Trans. Ind. Informat.*, vol. 19, no. 2, pp. 1581–1591, Feb. 2023.
- [36] Y. Zhu, G. Tang, X. Zhang, J. Jiang, and Q. Tian, "Haze removal method for natural restoration of images with sky," *Neurocomputing*, vol. 275, pp. 499–510, Jan. 2018.
- [37] X. Fu, Y. Huang, D. Zeng, X.-P. Zhang, and X. Ding, "A fusion-based enhancing approach for single sandstorm image," in *Proc. IEEE MMS*, 2014, pp. 1–5.
- [38] T. H. Park and I. K. Eom, "Sand-dust image enhancement using successive color balance with coincident chromatic histogram," *IEEE Access*, vol. 9, pp. 19749–19760, Jan. 2021.
- [39] J.-J. Jeon, T.-H. Park, and I.-K. Eom, "Sand-dust image enhancement using chromatic variance consistency and gamma correction-based dehazing," *Sensors*, vol. 22, no. 23, p. 9048, Nov. 2022.
- [40] A. Bartani, A. Abdollahpour, M. Ramezani, and F. A. Tab, "An adaptive optic-physic based dust removal method using optimized air-light and transfer function," *Multimedia Tools Appl.*, vol. 81, no. 23, pp. 33823–33849, Apr. 2022.
- [41] S. Lim and W. Kim, "DSLR: Deep stacked Laplacian restorer for low-light image enhancement," *IEEE Trans. Multimedia*, vol. 23, pp. 4272–4284, Nov. 2020.
- [42] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.
- [43] 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.
- [44] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug. 2011.
- [45] L. Zhang, Y. Shen, and H. Li, "VSI: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 10, pp. 4270–4281, Oct. 2014.
- [46] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'completely blind' image quality analyzer," *IEEE Signal Process. Lett.*, vol. 20, no. 3, pp. 209–212, Nov. 2012.
- [47] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE CVPR*, 2018, pp. 586–595.
- [48] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657–667, Feb. 2013.
- [49] N. Venkatanath, D. Praneeth, M. C. Bh, S. S. Channappayya, and S. S. Medasani, "Blind image quality evaluation using perception based features," in *Proc. IEEE NCC*, 2015, pp. 1–6.
- [50] B. Li et al., "Benchmarking single-image dehazing and beyond," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, Jan. 2019.
- [51] Z. Ge, S. Liu, F. Wang, Z. Li, and J. Sun, "YOLOX: Exceeding YOLO series in 2021," Aug. 2021, *arXiv:2107.08430*.
- [52] T.-Y. Lin et al., "Microsoft COCO: Common objects in context," in *Proc. ECCV*. Springer, 2014, pp. 740–755.



Yuan Gao is currently pursuing the B.E. degree with the School of Navigation, Wuhan University of Technology, Wuhan, China.

His research interests include deep learning, computer vision, and intelligent transportation systems.

Mr. Gao was awarded the National Scholarship by the Ministry of Education of the People's Republic of China in 2021. He was also awarded the Excellent Scholarship, the highest scholarship for undergraduate students in 2022.



Wenyu Xu received the B.E. degree in information engineering from the School of Mechanical Electronic and Information Engineering, China University of Mining and Technology, Beijing, China, in 2020.

She is currently working as an Information Engineer with Wuhan Baosight Software Company Ltd., Wuhan, China. Her research interests include computer vision and machine learning.



Yuxu Lu received the B.E. degree in navigation technology and M.S. degree navigation and information engineering from the School of Navigation, Wuhan University of Technology, Wuhan, China, in 2020 and 2022, respectively. He is currently pursuing the Ph.D. degree with The Hong Kong Polytechnic University, Hong Kong.

His research interests include computer vision, machine learning, and intelligent navigation systems.

Mr. Lu has received the Excellent Oral Presentation from the 11th International Conference on Machine Learning and Computing (ICMLC 2019).