the domain gap between synthetic and real hazy images. It uses a discriminator to distinguish the discrepancy between the synthetic image and the real one, which ignores the fact that the haze density in the real scene is unevenly distributed, resulting in color distortion in the dehazing results. Chen et al. [30] proposed a synthetic-to-real image dehazing framework that first pre-trained the dehazing model using synthetic datasets, and then fine-tuned it in an unsupervised manner using real hazy images to adapt to the domain shift between the training data and real hazy images. However, the lack of supervised training of real hazy–clear image pairs limits its ability to solve domain bias.

Therefore, we propose a haze transfer and feature aggregation network with a haze transfer subnetwork and a feature aggregation-based dehazing subnetwork. In the haze transfer subnetwork, a discriminator group (composed of one global and two chunking discriminators) is used to identify the haze difference between the synthetic and corresponding real hazy images at both the global and local levels. It aims to avoid non-homogeneity of haze that cannot be well characterized by overall discrimination. Meanwhile, the scene information transferred from a real hazy image to a synthesized hazy image can be weakened. In addition, the haze transfer subnetwork transfers the haze information from a real hazy image to a clear image to generate a new training sample pair to narrow the domain gap between the training sample and the real hazy image. The feature aggregation-based dehazing subnetwork considers the correlation between shallow and deep features. This correlation is used to enhance the role of deep and shallow related features in dehazing. The dehazing subnetwork also makes full use of the guidance of the clean image to keep the detailed information from being lost during dehazing. In general, the dehazing subnetwork produces a more comprehensive feature representation by aggregating the features at different levels under the guidance of clear image features, which helps to improve the dehazing performance of the network. The contributions of this study can be summarized as follows:

- Based on the hazy image generation physical model, we propose a haze transfer model for haze information from a real hazy image to a clear image, forming a training sample pair with a clear image. The transfer model considers not only the physical principles of hazy image generation but also the unevenness of haze density in real scenes.
- Aiming at hazy image recovery, we develop a feature aggregation network to highlight the information related to the content of clean images via correlation mining between shallow and deep features. Thus, the guidance of clear images can be used to prevent the loss of detailed information in dehazing.
- The proposed method overcomes the drawback that the model cannot be applied to real hazy images without performance loss because of lack of training sample pairs for real hazy images and clear images. The haze transfer subnetwork of the proposed method supports plug-and-play utilization in combination with other dehazing methods to solve the domain shifts between training samples and real hazy images.

2. Related work

Two popular single-image dehazing methods are considered in this study, including prior-based and deep learning-based methods.

2.1. Prior-based dehazing methods

Prior-based image dehazing methods [13–15,31–34] estimate the transmittance map and atmospheric light in the scattering model based on the prior information obtained from the clear image. Then, a clear image was recovered from the atmospheric scattering model based on the estimated results. He et al. [13] assumed that the pixels in a haze-free image are close to zero in at least one color channel. Based on this assumption, they proposed a dark channel prior to estimate the transmittance map. To remedy the shortcomings of the dark channel prior in the white area and the light channel prior in the black area, Jiang et al. [32] proposed an image dehazing method based on a superpixel adaptive dual channel prior. Fattal et al. [15] proposed a color line-based dehazing method based on the fact that the pixels of a small image block present a one-dimensional distribution in the RGB color space. Zhu et al. [14] proposed a color attenuation algorithm for removing haze from a single hazy image. The algorithm recovered the depth information of the hazy image by building a linear model of the scene depth of the hazy image and estimated the transmittance map and scene radiance using the atmospheric scattering model. Berman et al. [31] proposed a non-local prior-based dehazing algorithm, which assumed that each color cluster in a clear image becomes a haze-line in the RGB space in the presence of haze. Although those methods have been proven to be effective for image dehazing, their performance is often limited because the assumed prior might not be suitable for all real images.

2.2. Deep learning-based dehazing methods

In recent years, deep learning techniques have been widely used in media processing, such as pedestrian re-identification [35], human image generation [36], recommender systems [37], image deraining [38], and image dehazing [16–22,24–26,39]. Current deep learning-based dehazing methods can be divided into atmospheric scattering model-based dehazing methods and end-to-end image dehazing methods. The atmospheric scattering model-based dehazing methods [16–20] usually use neural networks to estimate parameters such as transmittance maps and then obtain clear images using the atmospheric scattering model. Cai et al. [16] proposed a four-part network including feature extraction, multi-scale mapping, local extrema, and nonlinear regression to estimate the transmittance map. Ren et al. [17] proposed a coarse-to-fine multi-scale deep neural network to estimate the transmittance map. Zhang et al. [18] proposed a densely connected pyramidal dehazing network that can jointly learn transmittance map, atmospheric light, and dehazed image. Pang et al. [20] proposed a binocular image dehazing framework (BidNet) that carried a stereo transform module to estimate the transmittance map. Because of error accumulation in the parameter estimation of the atmospheric scattering model, the aforementioned methods frequently produce inaccurate results.

The end-to-end image dehazing methods, on the other hand, have the advantage of avoiding parameter estimation for atmospheric scattering models [21,22,24–26]. Synthetic hazy images are often used as inputs to the network for training, with the corresponding clear images as labels. Ren et al. [21] pre-processed the original images using three methods: a white balance operation, image contrast enhancement, and gamma transform. The weights of the three processed images are learned directly through a gated fusion network and then fused to realize image dehazing. Liu et al. [22] proposed a GridDehazeNet with three modules: pre-processing, backbone, and post-processing. To adapt to different types of information in hazy images, Qin et al. [24] introduced a channel and pixel dual attention mechanism in a dehazing network. Dong et al. [25] proposed a U-Net-based multiscale enhanced dehazing network with dense feature fusion. Wu et al. [26] developed a dehazing network based on a similar autoencoder framework to preserve information flow and extend the receiver domain.

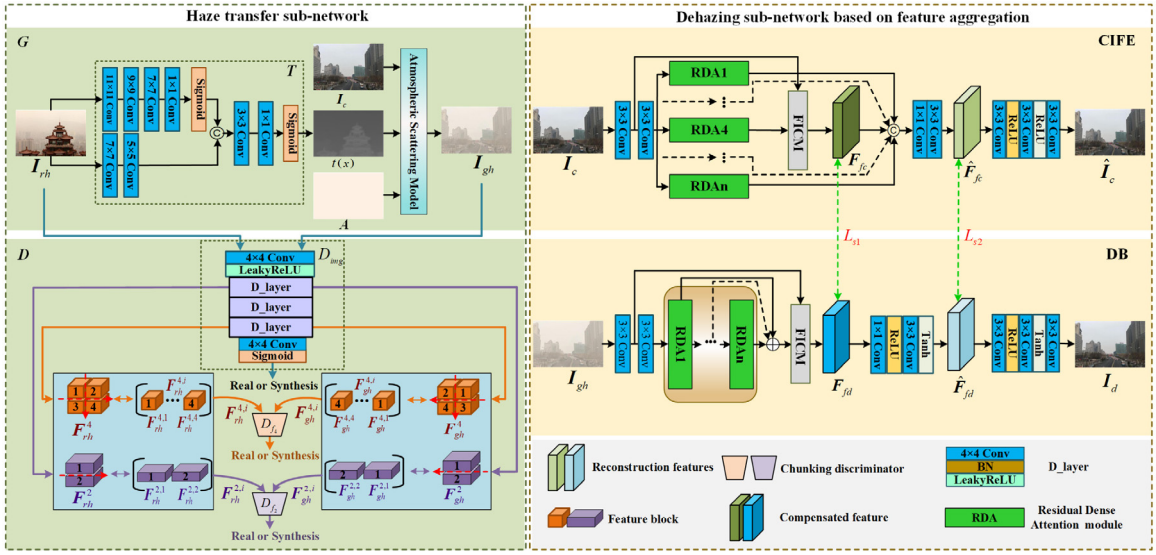

Fig. 1. Framework of the proposed dehazing method. ⊕ denotes summing operation. © represents concatenation operation.

However, when such algorithms are applied directly to real hazy images, the dehazing performance degrades dramatically due to domain shifts between the training data and hazy images in real scenes. It is challenging to capture hazy–clear image pairs from real scenes in practice.

3. Proposed method

The proposed method consists of two parts: a haze transfer subnetwork and a dehazing subnetwork based on feature aggregation. The framework is illustrated in Fig. 1.

The haze transfer subnetwork comprises a hazy image generator based on an atmospheric scattering model and a discriminator group. A hazy image generator is applied to transfer haze information from real hazy images to clear images. The discriminator group is used to keep the generated hazy images' haze information as close to that of real hazy images as possible. The feature aggregation-based dehazing subnetwork contains a Clear Image Feature Extractor (CIFE) and a Dehazing Branch (DB). CIFE guides the feature extraction in the DB. It ensures that the features extracted from the hazy image are consistent with the clear image. In the DB, we introduce the Residual Dense Attention (RDA) module and Fine-grained feature Information Compensation Module (FICM) to enhance useful information and prevent detailed information loss.

3.1. Haze transfer subnetwork

Supervised training with hazy images in real scenes might be the most direct way to improve image dehazing performance. However, it is difficult to apply because there is no corresponding clear image in real scenes. To this end, we propose a haze transfer model. Haze information in real scene is directly transferred to clear image, where the generated hazy image and corresponding clear image form training sample pair.

Assume that $\boldsymbol{I}_{rh}$ is a real hazy image, $\boldsymbol{I}_c$ is a clear image where $\boldsymbol{I}_{rh}$ and $\boldsymbol{I}_c$ are not paired, and $\boldsymbol{t}(x)$ is the transmittance map. T is the transmittance map estimation network. $\boldsymbol{t}(x)$ is estimated from T, and the hazy information in $\boldsymbol{I}_{rh}$ is transferred to $\boldsymbol{I}_c$. Thus, the generated hazy image $\boldsymbol{I}_{gh}$ can be obtained by

$$\boldsymbol{I}_{gh}(x) = \boldsymbol{I}_c(x)\boldsymbol{t}(x) + \boldsymbol{A}(1 - \boldsymbol{t}(x)) \tag{2}$$

where $\boldsymbol{A}$ is the atmospheric light estimated according to the literature [17]. T is designed based on the network structure of Multi-scale Convolutional Neural Networks (MSCNN) [17]. Convolutional kernels of varying sizes are directly used in T to process images or features of the same size in stages, allowing the features at different levels to be extracted effectively. Meanwhile, the pooling layer in the MSCNN is discarded to reduce the information loss.

The discriminators can be used directly for the consistency judgment of $\boldsymbol{I}_{gh}$ and $\boldsymbol{I}_{rh}$ to ensure the consistency of the distribution of haze in the generated hazy image $\boldsymbol{I}_{gh}$ and the real hazy image $\boldsymbol{I}_{rh}$. A discriminator group composed of an global discriminator D_{img} and two chunking discriminators D_{f_2} and D_{f_4} is built. The subscript 2 and 4 indicate that the features are divided into two and four blocks, respectively, as illustrated in Fig. 1. The global discriminator D_{img} renders the generated hazy image more similar in style to the real hazy images, whereas the chunking discriminators keep $\boldsymbol{I}_{gh}$ and $\boldsymbol{I}_{rh}$ features relatively comparable. Moreover, the chunking process considers the haze concentration difference in different areas of the real image, such that $\boldsymbol{I}_{gh}$ and $\boldsymbol{I}_{rh}$ could have similar haze concentrations. Therefore, the proposed haze transfer model could guarantee that the generated hazy image could be consistent with the real hazy image in terms of style, features, and haze density.

3.2. Dehazing subnetwork based on feature aggregation

The dehazing subnetwork mainly consists of the CIFE and DB, where the CIFE guides the DB to extract clearer features of the image in a supervised signal manner. Both CIFE and DB use a Residual Dense Block (RDB) [40] and Convolutional Block Attention (CBA) [41] as the basic structures to build the RDA module, as depicted in Fig. 2.

The RDA module combines the merits of CBA and RDB in its ability to focus on the salient features of an image while effectively preventing the removal of information at the pixel level. There are seven RDA modules in both the CIFE and DB.

Meanwhile, deep and shallow features extracted by deep neural networks represent different types of image information. Deep features contain information on global semantics with severe loss of detailed information, whereas shallow features contain rich detailed information and global semantic information that is not obvious. To fully utilize the complementary information between deep and shallow features, inspired by the works of non-local attention [42] and partial perceptual attention [43], we design the FICM, as shown in Fig. 3.

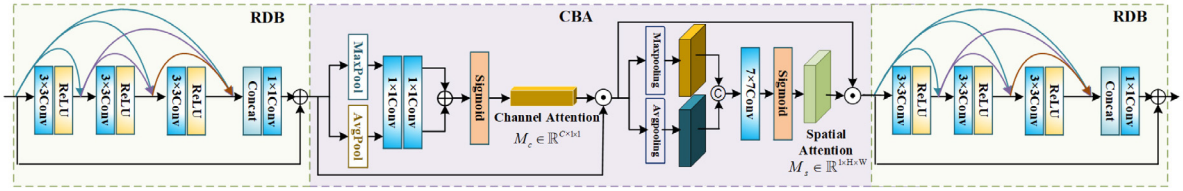

Fig. 2. Architecture of Residual Dense Attention (RDA) module, which consists of RDB-CBA-RDB. MaxPool and AvgPool denote channel-level pooling. MaxPooling and AvgPooling denote spatial-level pooling. ⊕ and ⊙ denote summing and element-wise multiplying operations, respectively. © denotes concatenation operation.

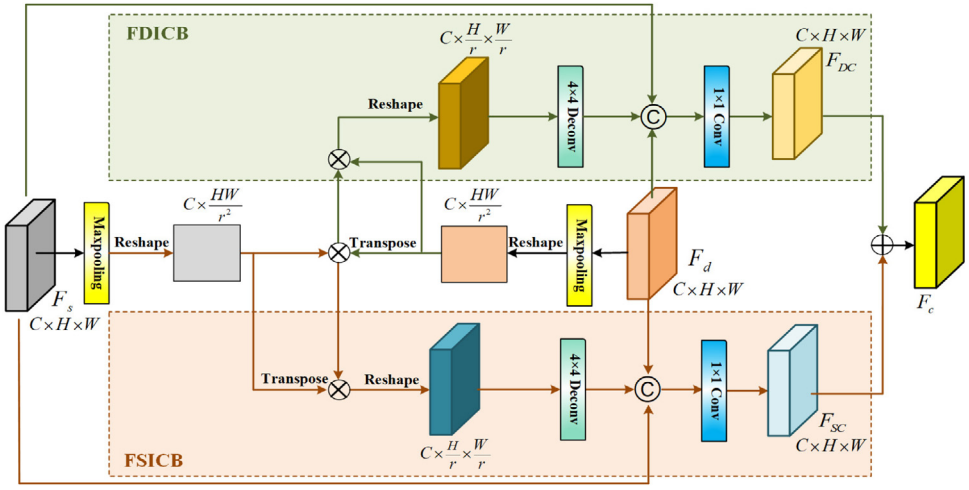

Fig. 3. Structure of FICM. Maxpooling denotes maximum pooling and ⊗ denotes matrix multiplication. C, H, and W denote dimensionality of features; $\boldsymbol{F}_s$ denotes shallow features, $\boldsymbol{F}_d$ denotes deep features; $\boldsymbol{F}_{DC}$ denotes features after detailed information compensation, $\boldsymbol{F}_{SC}$ denotes features after semantic information compensation, and $\boldsymbol{F}_c$ denotes features after comprehensive compensation. r is set to 4 in this study.

FICM includes a Fine-grained Detail texture Information Compensation Block (FDICB) and a Fine-grained Semantic Information Compensation Block (FSICB). By mining the correlation between shallow and deep features, FICM can enable shallow features to capture more global semantic information and compensate for lost detailed texture information in deep features. Because deep and shallow features contain information at different image levels, the proposed FICM embeds the ideological connotation of multiple knowledge representations [44].

3.3. Loss function

3.3.1. Loss function in haze transfer subnetwork

The objective of the haze transfer subnetwork is to learn a mapping from the real hazy image to the corresponding transmittance map and generate hazy images using clear images according to the atmospheric scattering model. Considering that the generated hazy images are as similar as possible to real images regarding color, brightness and haze concentration, the training of the haze transfer network consists of two steps: pre-training of the transmittance map estimation network and re-training after adding the discriminator group.

In the pre-training phase, the synthetic hazy images and the corresponding transmittance maps are used to pre-train network T. To ensure that the transmittance maps estimated by T are consistent with the labeled transmittance maps, the consistency loss [29,45] is used as a constraint:

$$L_{ct} = \|T(\boldsymbol{I}_{sh}) - \boldsymbol{t}_l(x)\|_1 \tag{3}$$

where $\boldsymbol{I}_{sh}$ is the synthetic hazy image, $\boldsymbol{t}_l(x)$ is the labeled transmittance maps corresponding to $\boldsymbol{I}_{sh}$, and T is the transmittance map estimation network.

The transmission map should be smooth in the regions of the same object or at the same depth and discontinuous across the boundaries of different objects or different depths [46]. Because the edges exist in discontinuous regions, we use the edge loss as a constraint to ensure that the estimated transmission map and the labeled transmission map have the same discontinuity and smoothness as possible at the corresponding locations to ensure the accuracy of the transmission map estimation. The edge loss is described by

$$L_{et} = \|E(T(\boldsymbol{I}_{sh})) - E(\boldsymbol{t}_l(x))\|_1 \tag{4}$$

where E denotes the edge detection operator, and the Canny operator is used in this study.

The Mean Square Error (MSE) is used to measure the reconstruction loss between the estimated transmittance maps and the labeled transmittance maps:

$$L_{rt} = \frac{1}{n}\|T(\boldsymbol{I}_{sh}) - \boldsymbol{t}_l(x)\|_F^2 \tag{5}$$

where, n denotes the number of pixels in the transmittance map.

In the re-training stage, the discriminator group composed of D_{img}, D_{f2} and D_{f4} makes the generated hazy images as similar as possible to the real ones in terms of style, features, haze density, etc. Network T coupled with Eq. (2) can be regarded as a generator (G), and the discriminator group is jointly trained using adversarial learning. The loss function of the discriminators in adversarial learning can be expressed as

$$L_{D_{img}} = \mathbb{E}[D_{img}(G(\boldsymbol{I}_{rh}, \boldsymbol{I}_c, \boldsymbol{A}))] + \mathbb{E}[D_{img}(\boldsymbol{I}_{rh}) - 1] \tag{6}$$

$$L_{D_{f2}} = \frac{1}{2}\sum_{i=1}^{2}\mathbb{E}[(D_{f2}(\boldsymbol{F}_{I_{rh}}^{2,i}) - 1)^2] + \frac{1}{2}\sum_{i=1}^{2}\mathbb{E}[(D_{f2}(\boldsymbol{F}_{I_{gh}}^{2,i}) - 0)^2] \tag{7}$$

$$L_{D_{f4}} = \frac{1}{2}\sum_{j=1}^{4}\mathbb{E}[(D_{f4}(\boldsymbol{F}_{I_{rh}}^{4,j})-1)^2] + \frac{1}{2}\sum_{j=1}^{4}\mathbb{E}[(D_{f4}(\boldsymbol{F}_{I_{gh}}^{4,j})-0)^2] \quad (8)$$

where $L_{D_{img}}$, $L_{D_{f2}}$ and $L_{D_{f4}}$ are the loss of D_{img}, D_{f2} and D_{f4}, respectively. $\boldsymbol{I}_{rh}, \boldsymbol{A}, \boldsymbol{I}_c$ denote the real hazy image, the atmospheric light of the real hazy image and the clear image, respectively. $\boldsymbol{I}_{rh}$ and $\boldsymbol{I}_c$ are not paired. $\boldsymbol{F}_{I_{rh}}^{2,i}$ and $\boldsymbol{F}_{I_{gh}}^{2,i}$ denote the ith block of the feature of the 2nd layer in D_{img} when $\boldsymbol{I}_{rh}$ and $\boldsymbol{I}_{gh}$ fed to D_{img}, respectively. $\boldsymbol{F}_{I_{rh}}^{4,j}$ and $\boldsymbol{F}_{I_{gh}}^{4,j}$ denote the jth block of the feature of the 4th layer in D_{img}, respectively. The loss of discriminator group D can be described by

$$L_D = L_{D_{img}} + L_{D_{f2}} + L_{D_{f4}} \quad (9)$$

The loss function of the generator G can be defined as

$$\begin{aligned} L_G = {} & \mathbb{E}[D_{img}(G(\boldsymbol{I}_{rh}, \boldsymbol{I}_c, \boldsymbol{A}))-1] + \frac{1}{2}\sum_{i=1}^{2}\mathbb{E}[(D_{f2}(\boldsymbol{F}_{I_{gh}}^{2,i})-1)^2] \\ & + \frac{1}{2}\sum_{j=1}^{4}\mathbb{E}[(D_{f4}(\boldsymbol{F}_{I_{gh}}^{4,j})-1)^2] + \frac{1}{n}\|T(\boldsymbol{I}_{rh}) - \widetilde{\boldsymbol{T}}(\boldsymbol{I}_{rh})\|_F^2 \end{aligned} \quad (10)$$

where T and $\widetilde{T}$ denote the pre-trained and re-trained transmittance map estimation networks, respectively. The training process of the haze transfer subnetwork is summarized in Algorithm 1.

Algorithm 1 Haze transfer subnetwork training

Input: $\boldsymbol{I}_{sh}$ // synthetic hazy images
$\boldsymbol{t}_l(x)$ // labeled transmittance maps corresponding to the synthetic hazy images
$\boldsymbol{I}_{rh}$ // real hazy images
$\boldsymbol{I}_c$ // clear images
Initialize: learning rate
(1) Pre-training transmittance map estimation network with $\boldsymbol{I}_{sh}$ and $\boldsymbol{t}_l(x)$ by Equations (3), (4) and (5)
(2) Re-training transmittance map estimation network with $\boldsymbol{I}_{rh}$ and $\boldsymbol{I}_c$
for i = 1:epoch
fixing generator G, training discriminator group D by Equation (9)
fixing discriminator group D, training generator G by Equation (10)
end for
Output: trained transmittance map estimation network

3.3.2. Loss function in dehazing subnetwork

The dehazing subnetwork uses the generated hazy images to train the dehazing model, which can be adapted to the real hazy image domain. CIFE is used to extract features from a clear image and then assist the DB in extracting clear image-related features in hazy images. Consistency and MSE losses are used to supervise the learning of the CIFE.

$$L_{cc} = \|\hat{\boldsymbol{I}}_c - \boldsymbol{I}_c\|_1 \quad (11)$$

$$L_{Mc} = \frac{1}{n}\|\hat{\boldsymbol{I}}_c - \boldsymbol{I}_c\|_F^2 \quad (12)$$

where $\hat{\boldsymbol{I}}_c$ is the image reconstructed with the features extracted by the CIFE, and $\boldsymbol{I}_c$ is the labeled clear image. n denotes the number of pixels in the image.

The DB is used to convert hazy images into clear images, and the consistency loss and MSE loss are used to constrain the dehazing results to be closer to clear images:

$$L_{cd} = \|\boldsymbol{I}_d - \boldsymbol{I}_c\|_1 \quad (13)$$

$$L_{Md} = \frac{1}{n}\|\boldsymbol{I}_d - \boldsymbol{I}_c\|_F^2 \quad (14)$$

where $\boldsymbol{I}_d$ is the dehazed images.

Perceptual loss is used to reduce the feature disparity between the dehazing result $\boldsymbol{I}_d$ and the clear labeled image $\boldsymbol{I}_c$ and to improve the visual effect of the dehazed images:

$$L_p = \|\phi(\boldsymbol{I}_d) - \phi(\boldsymbol{I}_c)\|_F^2 \quad (15)$$

where ϕ represents the feature maps obtained by the 16-layer VGG network [47].

In addition, the DB takes the features extracted from the CIFE as guidance and constraint. L_{s1} and L_{s2} are used to constrain the distance between features extracted from different layers of the CIFE and DB, respectively:

$$L_{s1} = \|\boldsymbol{F}_{fd} - \boldsymbol{F}_{fc}\|_1 \quad (16)$$

$$L_{s2} = \|\hat{\boldsymbol{F}}_{fd} - \hat{\boldsymbol{F}}_{fc}\|_1 \quad (17)$$

where $\boldsymbol{F}_{fd}$ and $\hat{\boldsymbol{F}}_{fd}$ are fine-grained features extracted from DB and reconstructed features, respectively. $\boldsymbol{F}_{fc}$ and $\hat{\boldsymbol{F}}_{fc}$ are fine-grained features extracted from CIFE and reconstructed features, respectively. The total loss of the DB can be expressed as

$$L_{total} = \alpha_1 L_{cd} + \alpha_2 L_{Md} + \beta_1 L_{s1} + \beta_2 L_{s2} + \gamma L_p \quad (18)$$

where α_1, α_2, β_1, β_2 and γ are hyper-parameters.

The training process of the dehazing subnetwork is summarized in Algorithm 2.

Algorithm 2 Dehazing subnetwork training

Input: $\boldsymbol{I}_{gh}$ // generated hazy images
$\boldsymbol{I}_c$ // clear images
Initialize: The learning rate, the decay rate
(1) Training CIFE with $\boldsymbol{I}_c$ by Equation (11) and (12)
(2) Training DB with $\boldsymbol{I}_{gh}$ and $\boldsymbol{I}_c$ by Equation (18)
Output: trained DB

4. Experiment

4.1. Experimental settings

The RESIDE dataset [48] is used to train the proposed network and generate the hazy images with the haze information in the real scene. RESIDE contains five subsets: synthetic outdoor training set (OTS), synthetic indoor training set (ITS), synthetic objective test set (SOTS), hybrid subjective test set (HSTS), and real-world task-driven testing set (RTTS). There are 72,135 synthetic hazy images in the OTS dataset. There are 13,990 indoor synthetic images in the ITS. SOTS includes 500 indoor synthetic hazy images, 500 outdoor synthetic hazy images, and 500 clear images. The HSTS stores ten outdoor synthetic hazy images with paired clear images and ten real hazy images. There are 4322 hazy images of real scenes in the RTTS.

In the haze transfer subnetwork training process, 350 synthetic hazy images and their corresponding transmittance maps are selected from the ITS to pre-train the network T. Then, 500 real hazy images are chosen from the RTTS, and 500 clear images are chosen from the OTS to re-train the haze transfer network. A total of 3822 clear images are randomly selected from the OTS, and the haze from the 3822 real hazy images in the RTTS is transferred to the clear images separately by the haze transfer

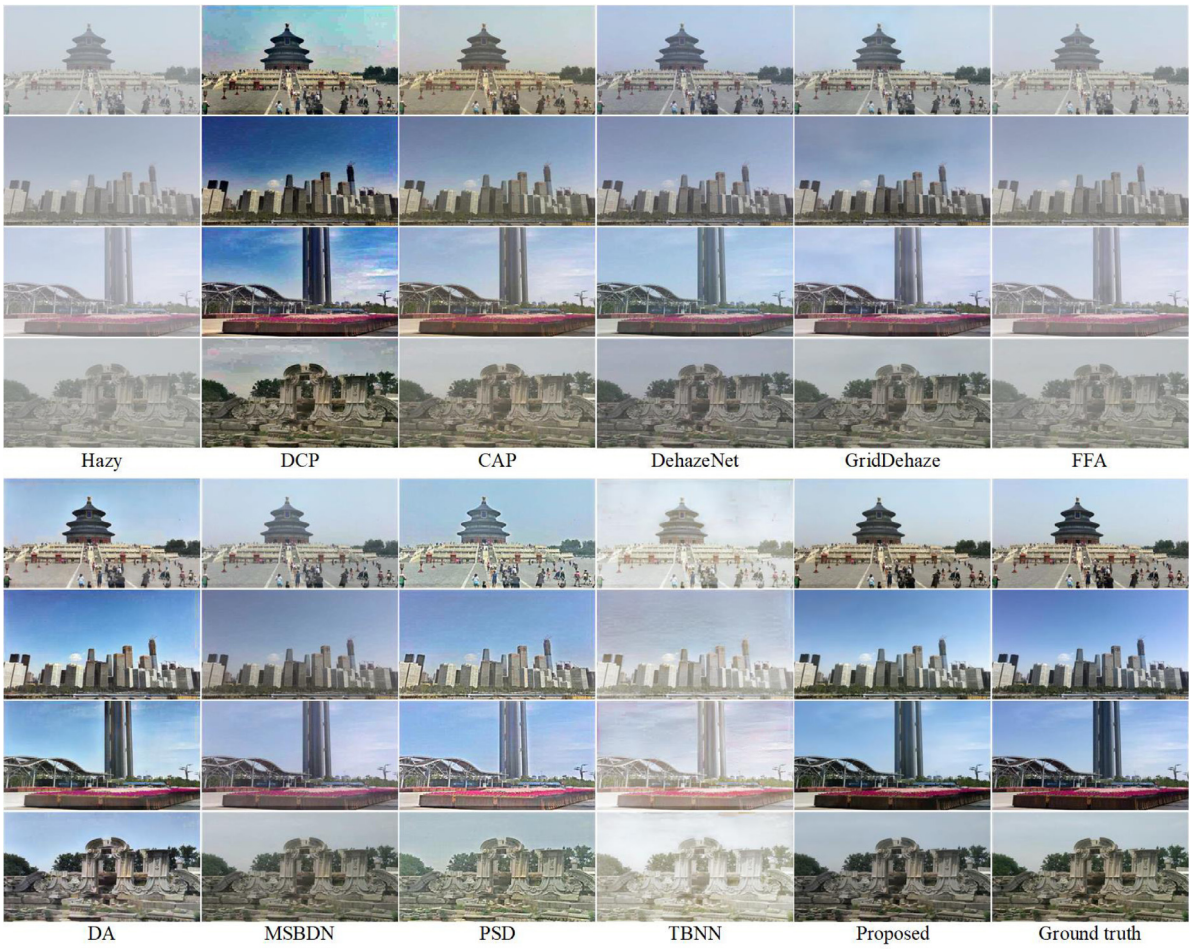

Fig. 4. Visual comparisons on GTset.

Table 1
Performance of different methods on GTset.

Methods	DCP	CAP	DehazeNet	GridDehaze	FFA	DA	MSBDN	PSD	TBNN	Proposed
PSNR (dB)	16.15	19.44	18.18	16.84	13.14	22.91	17.48	13.92	10.22	**26.10**
SSIM	0.7500	0.8165	0.8067	0.8032	0.7209	0.8308	0.8063	0.7477	0.6709	**0.8897**

subnetwork to obtain the Haze Transfer dataset (HTset). Thus, the HTset consists of 3822 synthetic hazy images and their corresponding 3822 clear images. Three thousand pairs of hazy–clear images in the HTset are used as the training set of the dehazing subnetwork, and 822 synthetic hazy images are used as the Generative Testing Set (GTset). In addition, 220 clear images (non-overlapping with previous images) are randomly selected from the OTS to train the CIFE. Except for the GTset, the other two datasets are also used to test the dehazing subnetwork, one is the Real Hazy Set (RHset). The RHset consists of 52 hazy images in real scenes: 18 images randomly selected from the RTTS, 26 images taken by cameras, and eight images selected from DCP [13] and DehazeNet [16]. The HAZERD dataset [49] contains 75 synthetic hazy images.

The Pytorch 1.5.0 framework and an NVIDIA GeForce RTX 3090 24 GB GPU are used in the experiments. For the haze transfer subnetwork, the learning rate is 1×10^{-4}, and the subnetwork is trained for 150 epochs. In the dehazing subnetwork, the CIFE is trained for 200 epochs with a learning rate of 1×10^{-5}. The DB is trained for 70 epochs with an initial learning rate of 1×10^{-3} and a decay rate of 0.5. The learning rate is adjusted after each 20 epoch, and Adam [50] is used as the optimizer for all the networks. The PSNR and SSIM [51] indexes are applied to evaluate the dehazing results. The unit of PSNR is decibel (dB).

4.2. Experiment on GTset

For the GTset, the proposed method is compared with current state-of-the-art image dehazing methods such as DCP [13], CAP [14], DehazeNet [16], GridDehaze [22], FFA [24], DA [29], MSBDN [25], TBNN [52], and PSD [30] for performance evaluation. DCP and CAP are prior-based methods, whereas others are deep learning-based methods. The average PSNR and SSIM of the dehazing results of all methods on the GTset are listed in Table 1.

Optimal results are highlighted in bold. It can be observed that the PSNR/SSIM on the GTset of the proposed method ranks over the other methods compared. The dehazing results of the different methods are exhibited in Fig. 4.

As shown in Fig. 4, the haze still exists in the dehazing results of the CAP, DehazeNet, GridDehaze, FFA, MSBDN, PSD, and TBNN methods. The FFA and TBNN methods have the most haze residue, and the PSD method presents a color distortion. The DCP method exhibits excessive dehazing, which leads to severe color distortion in the sky. The DA method has the problem of uneven dehazing, such as excessive dehazing in the sky area leading

Fig. 5. Visual comparisons on HAZERD.

Fig. 6. Detailed comparison of the dehazing results of DA and the proposed method. (a) and (c) Results generated by DA; (b) and (d) Results generated by the proposed method.

Table 2
Performance of different methods on HAZERD.

Methods	DCP	CAP	DehazeNet	GridDehaze	FFA	DA	MSBDN	PSD	TBNN	Proposed
PSNR (dB)	14.58	14.18	15.42	15.75	15.94	**17.65**	15.45	14.27	11.86	16.49
SSIM	0.7255	0.7085	0.7569	0.8044	0.7871	0.8043	0.8028	0.7230	0.7445	**0.8060**

to color distortion, while some local areas are not sufficiently dehazed. For example, the pedestrians in the image in the first row are still shrouded in the haze. The dehazing results of the proposed method are closest to the ground truth and achieve the best visual effect compared with the other methods.

4.3. Experiment on HAZERD

To verify the generalizability of the proposed method, we evaluate it on the HAZERD dataset [49]. Fig. 5 shows the dehazing results of the different methods, and their comparison is presented in Table 2.

As shown in Fig. 5, DCP, CAP, DehazeNet, and GridDehaze suffer from color distortion. The dehazing results obtained by DA, MSBDN, PSD, and TBNN are over-enhanced and unrealistic. FFA and the proposed method restore images with a better visual effect and are most similar to the ground truth. However, the dehazing performance of FFA is worse than that of the proposed method. As shown in Table 2, the proposed method achieves the highest value for SSIM and the suboptimal value for PSNR. To further compare the performance difference between DA and the proposed method, the marked regions of the dehazing results are enlarged to observe the details. The detailed comparison is shown in Fig. 6. The DA method gives blurred edges and artifacts compared with the proposed method, which implies that the proposed method delivers superior performance. Thus, it can be concluded that the proposed method outperforms other methods compared on HAZERD dataset.

4.4. Experiment on RHset

Experiments are conducted on RHset to further verify the proposed method by comparing it with the DCP [13], CAP [14], DehazeNet [16], GridDehaze [22], FFA [24], DA [29], MSBDN [25], PSD [30], and TBNN [52] methods. The results are shown in Fig. 7.

As shown in Fig. 7, the dehazing results of the DCP method have severe color distortion in the sky area, such as in the 1st, 3rd, and 9th columns of the 2nd row. Furthermore, the DCP method suffers from uneven dehazing, as seen in the buildings in the 1st column of the 2nd row, where haze has been removed in some areas but remains in others. Detailed information on the dehazing results obtained by the CAP method is lost, for example, the

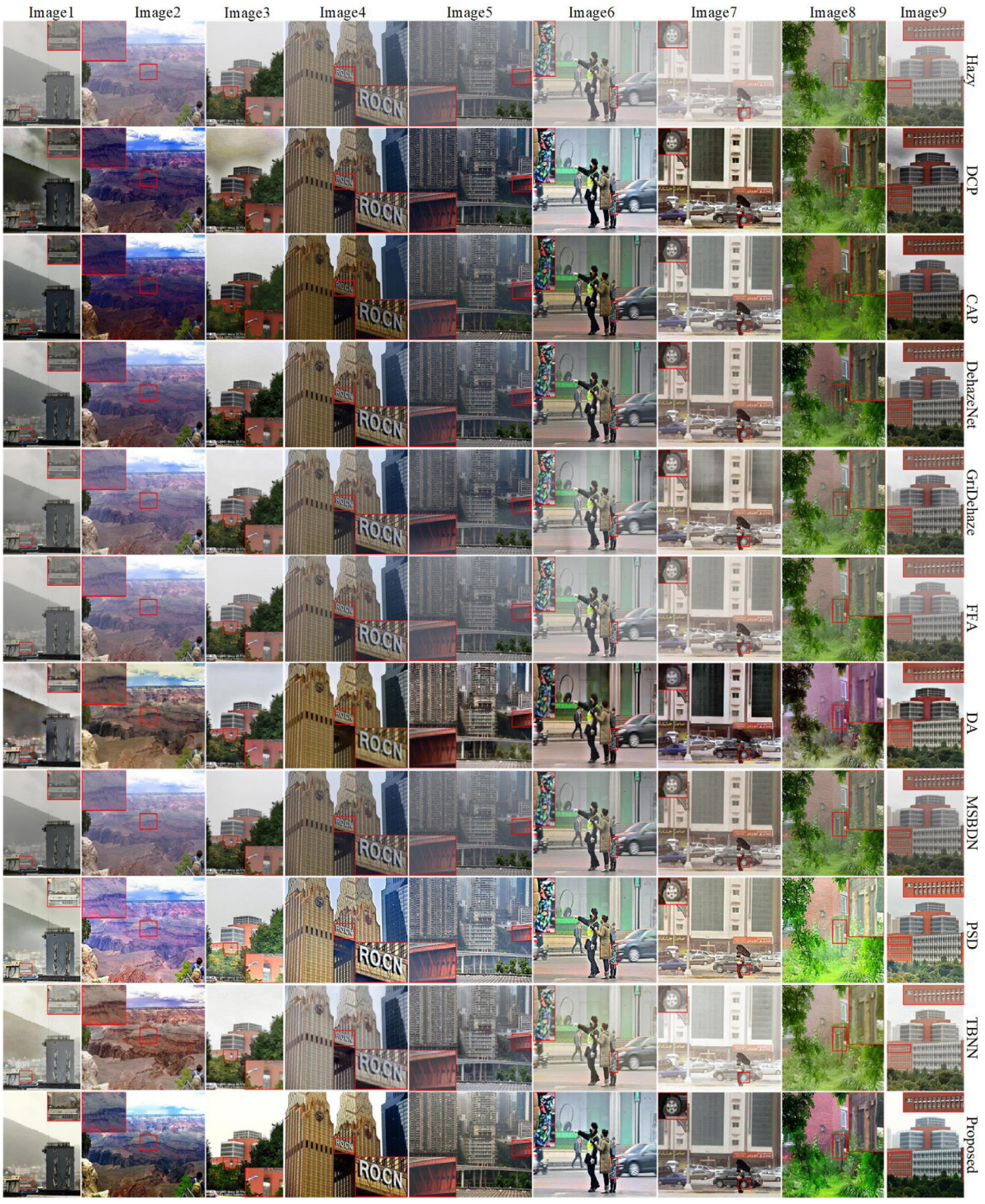

Fig. 7. Visual comparisons on RHset.

detailed information of the tree leaves in the 3rd column of the 3rd row is completely missing. The DehazeNet, GridDehaze, and MSBDN methods exhibit insufficient dehazing, and haze residues are still in their dehazing results. The PSD and TBNN methods also have poor effects, and their dehazing results exhibit color distortion to a certain degree, such as the 3rd column of the 9th row and the 2nd column of the 10th row. The FFA method has little dehazing effect on hazy images in real scenarios. The DA method has two problems: uneven dehazing and severe color distortion. For example, a large amount of haze remains around the building in the 1st column of the 7th row, and the dehazing results in the 2nd and 8th columns of the 7th row show serious color distortion. We have enlarged certain areas in the dehazing results for better observation:

- In the enlarged regions of the 1st column, the dehazing results of the GridDehaze, FFA, MSBDN, and TBNN methods have haze residues.
- In the 2nd column, the dehazing results of the DCP, CAP, PSD, and TBNN methods are over-enhanced. Haze residues are present in the dehazing results of the GridDehaze, FFA, and MSBDN methods.

Fig. 8. Hazy images synthesized by the proposed method. 1st row are real hazy images, 2nd row are generated hazy images, 3rd row are clear images.

Fig. 9. Effectiveness of the haze transfer subnetwork. (a) Hazy images, (b) GridDehaze trained with OTS, (c) GridDehaze trained with HTset, (d) TBNN trained with OTS, (e) TBNN trained with HTset, (f) Proposed model trained with OTS, (g) Proposed model trained with HTset.

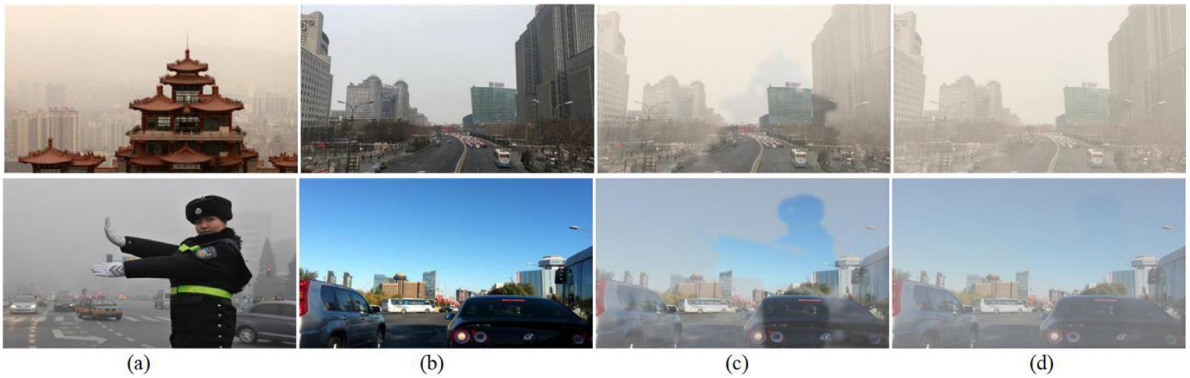

Fig. 10. Generated hazy images by the proposed method: (a) real hazy image, (b) clear image, (c) generated hazy images by the proposed method without discriminators, (d) generated hazy images by the proposed method with discriminators.

- In the 3rd column, except for the CAP and proposed methods, the dehazing results of the other compared methods show haze residues.
- In the 4th column, the dehazing results of the DA method exhibit edge blurring, and haze residues are present in the dehazing results of the TBNN, MSBDN, and FFA methods.
- In the 5th column, except for the DCP, DA, and proposed methods, the dehazing results of other comparison methods contain haze residues.
- In the 6th column, there are haze residues in the dehazing results for GridDehaze, FFA, and TBNN.
- In the 7th column, except for the proposed method, all other methods have haze residues on the wheels.
- In the 8th column, the dehazing results of the GridDehaze, FFA, MSBDN, and TBNN methods have haze residues, and DA suffers from color distortion.
- In the 9th column, the dehazing results of the GridDehaze, FFA, MSBDN, and TBNN methods contain haze residues.

The proposed method achieves an optimal visual effect from the overall and local dehazing results for hazy images in a real scene.

4.5. Ablation study

The ablation experiments are performed to confirm the contribution of each module in the proposed model.

4.5.1. Validation of haze transfer subnetwork

To address the problem of domain disparity between the existing synthetic dataset and real hazy images, we construct a haze transfer network to make the generated hazy images closer to the hazy images in real scenes. Some examples of hazy images generated by the haze transfer subnetwork are shown in Fig. 8. It can be observed that the haze information in real hazy images is transferred to clear images.

The GridDehaze [22] and TBNN [52] methods are selected to test hazy images in real scenes using trained models provided in their corresponding literature. The results are shown in Fig. 9.

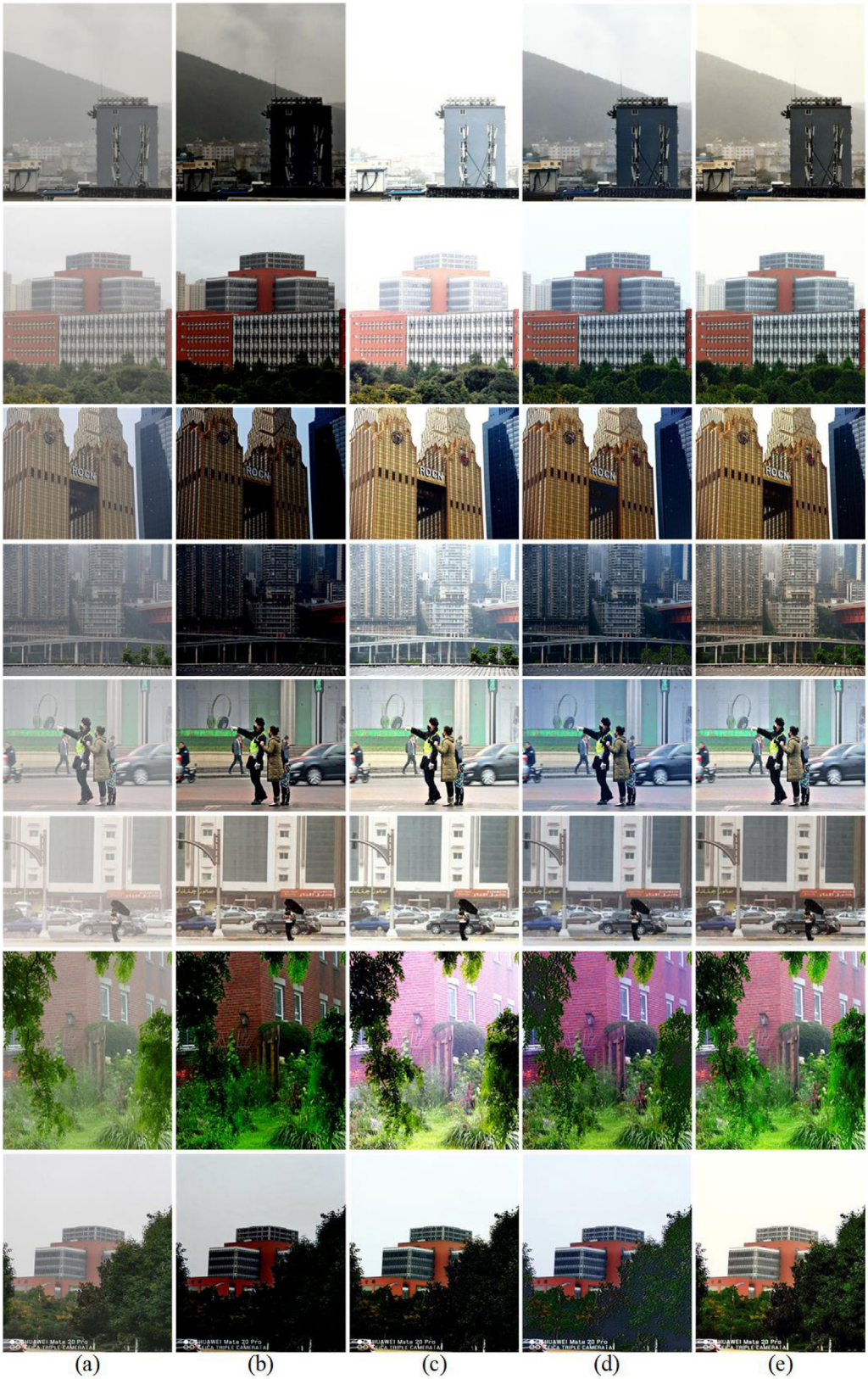

Fig. 11. Ablation experiment results of different modules on the RHset. (a) Hazy images, (b) Baseline, (c) Baseline+FICM, (d) Baseline+FICM+RDA, (e) Baseline +FICM+RDA+CIFE.

The dehazing models trained directly with the OTS dataset have almost no dehazing effect on the hazy images in real scenes. After being trained by the HTset, the dehazed models achieve better dehazing performances on the real hazy images, proving that the data generated by using the haze transfer subnetwork has better generalization performance. Therefore, the proposed

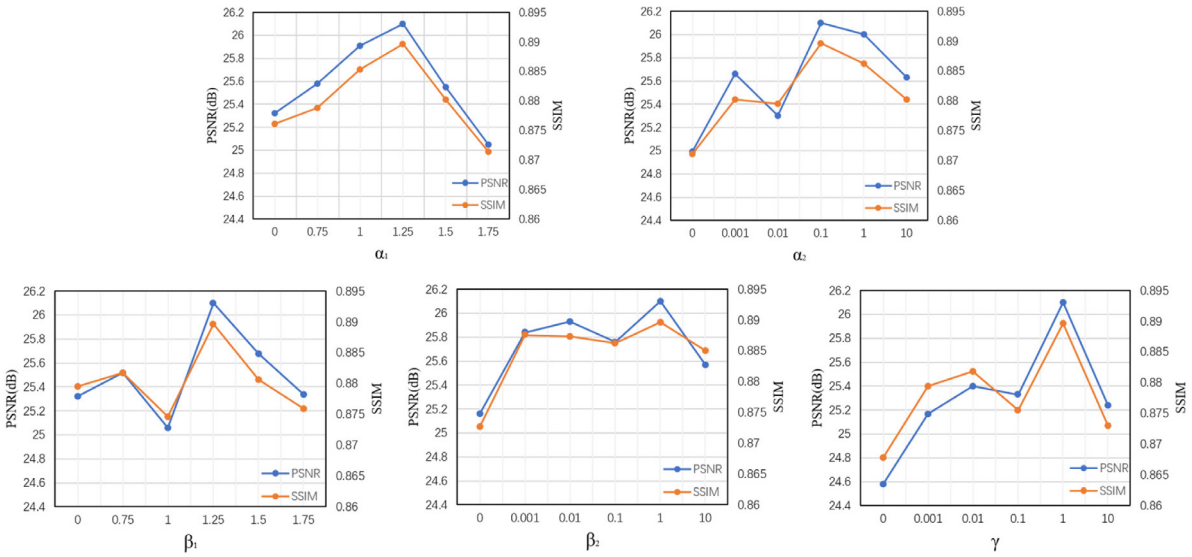

Fig. 12. Parameter analysis on the GTset. PSNR/SSIM curves with α_1, α_2, β_1, β_2, and γ.

Table 3
Ablation experiments of different modules on the GTSet.

Model	Baseline	Baseline+FICM	Baseline+FICM+RDA	Baseline+FICM+RDA+CIFE
PSNR (dB)	20.52	22.84	24.35	**26.10**
SSIM	0.8407	0.8578	0.8687	**0.8897**

haze transfer subnetwork supports plug-and-play, and it can be used in combination with other image dehazing methods without modification.

Moreover, we design a discriminator group in the haze transfer subnetwork to make the generated hazy image more natural, of which the haze distribution could be as similar as that of the real hazy image. Haze transfer subnetworks (with and without discriminators) are used to generate the hazy image. Fig. 10 shows the hazy images generated by the haze transfer subnetworks under these two conditions.

Comparing Fig. 10(c) with (d), some scene information is transferred from the real hazy image to the generated image without the discriminators. The discriminators can reduce the negative influence of the scene information of a real hazy image on a synthetic hazy image.

4.5.2. Validation of dehazing subnetwork

A series of experiments are conducted to verify the effectiveness of FICM, RDA, and CIFE in the dehazing network:

- **Baseline:** Replace RDA with RDB, remove the FICM and guidance of the CIFE for the DB in Fig. 1.
- **Baseline+FICM:** The FICM module is added to **Baseline** to verify the effectiveness of FICM.
- **Baseline+FICM+RDA:** All RDBs are replaced with RDAs based on **Baseline+FICM** to verify the effectiveness of the RDA structure.
- **Baseline+FICM+RDA+CIFE:** The guidance of CIFE is introduced based on **Baseline+FICM+RDA** to verify the effectiveness of CIFE.

The average PSNR and SSIM on the GTSet are shown in Table 3.

Comparing **Baseline** with **Baseline+FICM**, it can be seen that the values of PSNR and SSIM are greatly improved by adding the FICM to **Baseline**, which means that the FICM plays a significant role in the dehazing subnetwork. We discover that by replacing RDB with RDA in the dehazing subnetwork, the PSNR and SSIM values improve. Furthermore, the performance of the dehazing subnetwork is improved by introducing CIFE guidance.

Table 4
Performance of the proposed model on GTset with different number of RDA modules.

RDA modules	n = 5	n = 7	n = 9	n = 11
PSNR (dB)	25.08	**26.10**	26.02	25.72
SSIM	0.8738	**0.8897**	0.8849	0.8876
Params (M)	1.20	1.57	1.94	2.32

We conduct ablation experiments using hazy images in real scenes to observe the dehazing effect of the network after adding the FICM, RDA, and CIFE at **Baseline**. The results are shown in Fig. 11.

It can be seen that the color of the dehazed images obtained by **Baseline** is dark, and the contrast of the dehazed images is greatly improved by adding the FICM, but with color distortion. RDA is helpful in improving the brightness of the dehazed results. The model obtains optimal results by introducing guidance from the CIFE. To summarize, the proposed FICM, RDA, and CIFE can contribute to the improvement of dehazing performance.

4.5.3. Influence of the number of RDAs

In this section, we describe the ablation experiments that are carried out to select an appropriate number of RDA modules. The number of RDA modules in CIFE and DB is increased from 5 to 11. Models with different numbers of RDA modules are tested on the GTset. Table 4 shows the impact of the number of RDA modules (denoted by n) in the CIFE and DB on model performance. "Params" indicates the number of parameters.

As shown in Table 4, the PSNR and SSIM perform better when $n = 7$. When n is 9 or 11, the number of the parameters of the model increases, and the performance of the model decreases. Thus, we set $n = 7$.

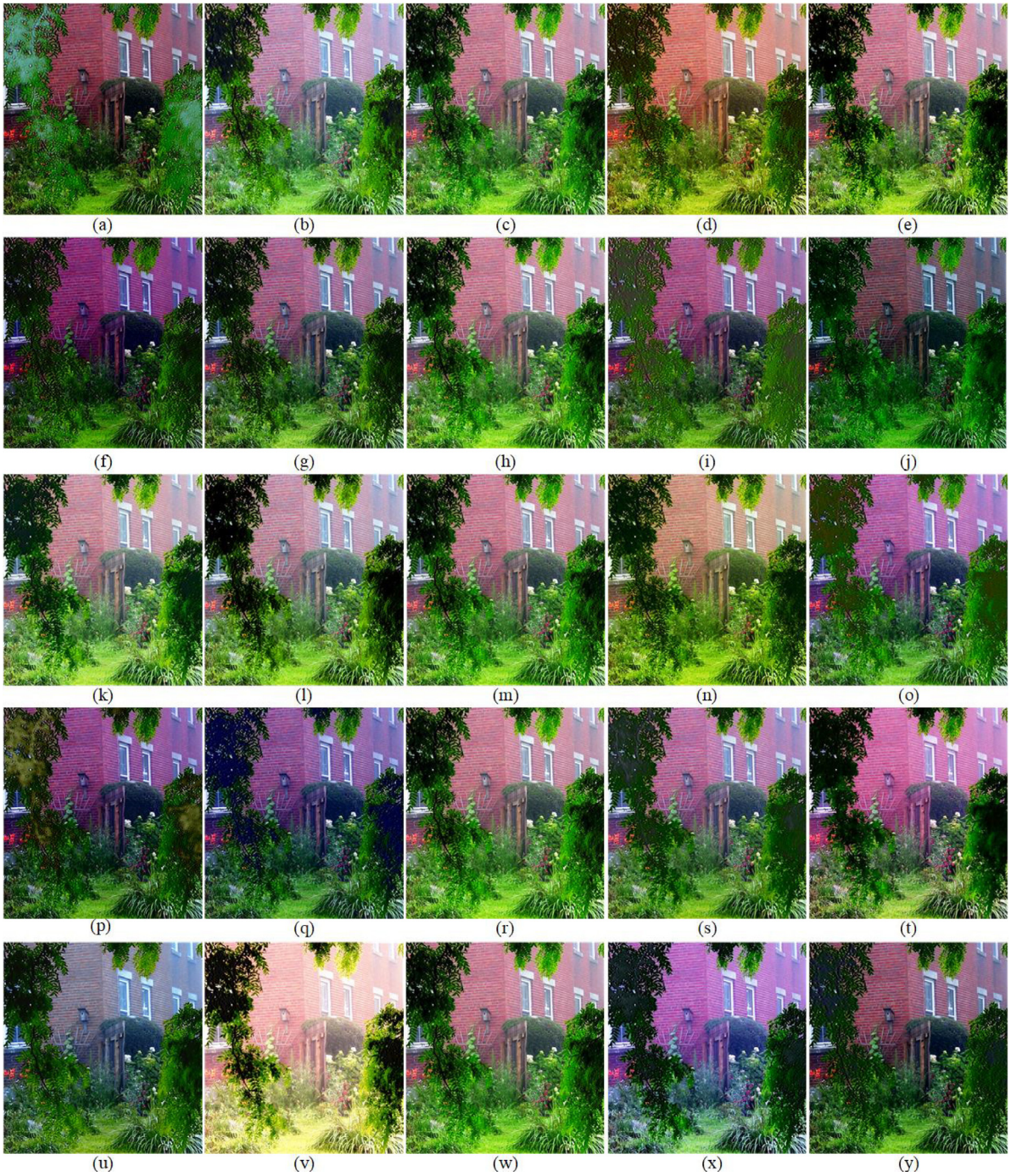

Fig. 13. Effect of hyper-parameters on the dehazing quality of the real hazy image. (a)–(e) $\alpha_1 = 0.75, 1, 1.25, 1.5, 1.75$, (f)–(j) $\alpha_2 = 0.001, 0.01, 0.1, 1, 10$, (k)–(o) $\beta_1 = 0.75, 1, 1.25, 1.5, 1.75$, (p)–(t) $\beta_2 = 0.001, 0.01, 0.1, 1, 10$, (u)–(y) $\gamma = 0.1, 1, 10, 100, 1000$.

4.6. Parameter selection and analysis

In Eq. (18), five hyper-parameters α_1, α_2, β_1, β_2, and γ need to be set. Each hyper-parameter is used to balance the contribution of the corresponding loss term to the DB. α_1 and α_2 control the contributions of consistency loss L_{cd} and reconstruction loss L_{Md}, respectively. β_1 and β_2 control the contribution of L_{s1} and L_{s2}, respectively, in the process of clear image guidance. γ is used to balance the perceptual loss L_p. We discuss the influence of hyper-parameters by changing one while fixing the others.

4.6.1. Parameter selection for synthetic hazy image dehazing

We use the GTset to analyze the influence of hyper-parameters on the dehazing results. Fig. 12 shows the PSNR/SSIM curves when each parameter changes.

As shown in Fig. 12, both PSNR and SSIM obtain the optimal values when $\alpha_1 = 1.25$, $\alpha_2 = 0.1$, $\beta_1 = 1.25$, $\beta_2 = 1$, and $\gamma = 1$. Therefore, we set $\alpha_1 = 1.25$, $\alpha_2 = 0.1$, $\beta_1 = 1.25$, $\beta_2 = 1$, and $\gamma = 1$ for synthetic hazy image dehazing.

4.6.2. Parameter selection for real hazy image dehazing

The real hazy image "Image8" in Fig. 7 is selected to illustrate the influence of hyper-parameters on the dehazing performance of the proposed model. With other parameters fixed,

Table 5
Model size, number of parameters and computational complexity of compared methods.

Methods	DCP	CAP	DehazeNet	GridDehaze	FFA	DA	MSBDN	PSD	TBNN	Proposed
Model size (MB)	–	–	0.03	3.70	25.39	208.35	119.00	126.00	192.77	7.55
Params (M)	–	–	0.01	0.96	4.46	54.59	31.35	31.41	50.35	1.57
FLOPs (G)	–	–	0.39	16.46	220.55	30.04	31.81	34.66	64.74	77.05

when $\alpha_1 = 0.75, 1, 1.25, 1.5, 1.75$, the dehazing results are shown in Fig. 13(a)–(e). It can be seen that the dehazing results at $\alpha_1 = 0.75, 1, 1.5, 1.75$, have color distortion and detail loss in comparison with that when $\alpha_1 = 1.25$. So we set $\alpha_1 = 1.25$. When $\alpha_2 = 0.001, 0.01, 0.1, 1, 10$, the dehazing results are shown in Fig. 13(f)–(j). From Fig. 13(f) to (j), it can be seen that the dehazed image has the best visual effect when $\alpha_2 = 0.1$. So we set $\alpha_2 = 0.1$. Similarly, it can be seen from Fig. 13(k) to (t) that the dehazing results have the optimal visual effect when $\beta_1 = 1.25$ and $\beta_2 = 0.1$. Therefore, we set $\beta_1 = 1.25$, $\beta_2 = 0.1$. The dehazing results for $\gamma = 0.1, 1, 10, 100, 1000$ are shown in Fig. 13(u)–(y). As shown in Fig. 13(u)–(y), the dehazing results exhibits color distortion when $\gamma = 0.1, 1, 100, 1000$. So we set $\gamma = 10$. To sum up, we set $\alpha_1 = 1.25$, $\alpha_2 = 0.1$, $\beta_1 = 1.25$, $\beta_2 = 0.1$ and $\gamma = 10$ for real hazy image dehazing.

4.7. Computational complexity

In this section, we compare the model size, number of parameters, and computational complexity of different methods. Table 5 presents the results. "-" indicates that no reported result is available, "Params" indicates the number of parameters, and "FLOPs" indicates the computational complexity.

As exhibited in Table 5, the model size and number of parameters of the proposed method are smaller, except for the DehazeNet and GridDehaze methods. However, the computational complexity of the proposed method is high. The FICM is used to calculate the similarity between each pixel in the deep and shallow features, which increases the computational complexity of the proposed model. Based on the visual effect and quantitative analysis of the dehazing results, it can be concluded that the proposed method achieves better performance with fewer parameters than the compared methods.

5. Conclusion

In this work, we proposed a method to implement image dehazing with haze transfer and feature aggregation. The discriminator group in haze transfer subnetwork was used to generate hazy images with similar haze information to real hazy images. The relationship between shallow and deep features was considered in feature aggregation-based dehazing subnetwork. The relationship was used to accentuate the information related to the content of clear image. At the same time, the guiding mechanism of clear image was used to extract useful features and prevent information loss during dehazing process. The proposed method outperforms existing dehazing methods in real hazy scenarios through experimental verification.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 62161015, No. 61966021), and Yunnan Natural Science Funds (Grant No. 2017FB094).

References

[1] Y. Pang, Y. Li, J. Shen, L. Shao, Towards bridging semantic gap to improve semantic segmentation, in: 2019 IEEE/CVF International Conference on Computer Vision, ICCV, 2019, pp. 4229–4238.
[2] S. Ma, Y. Pang, J. Pan, L. Shao, Preserving details in semantics-aware context for scene parsing, Sci. China Inf. Sci. 63 (2020) 120106.
[3] Z. Zhang, Y. Pang, CGNet: cross-guidance network for semantic segmentation, Sci. China Inf. Sci. 63 (2020) 120104.
[4] J. Nie, R.M. Anwer, H. Cholakkal, F.S. Khan, Y. Pang, L. Shao, Enriched feature guided refinement network for object detection, in: 2019 IEEE/CVF International Conference on Computer Vision, ICCV, 2019, pp. 9536–9545.
[5] Y. Pang, J. Xie, M.H. Khan, R.M. Anwer, F.S. Khan, L. Shao, Mask-guided attention network for occluded pedestrian detection, in: 2019 IEEE/CVF International Conference on Computer Vision, ICCV, 2019, pp. 4966–4974.
[6] J. Cao, Y. Pang, J. Han, X. Li, Hierarchical shot detector, in: 2019 IEEE/CVF International Conference on Computer Vision, ICCV, 2019, pp. 9704–9713.
[7] Y. Li, Y. Pang, J. Shen, J. Cao, L. Shao, NETNet: Neighbor erasing and transferring network for better single shot object detection, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2020, pp. 13346–13355.
[8] H. Li, J. Pang, D. Tao, Z. Yu, Cross adversarial consistency self-prediction learning for unsupervised domain adaptation person re-identification, Inform. Sci. 559 (2021) 46–60.
[9] J. Pang, D. Zhang, H. Li, W. Liu, Z. Yu, Hazy re-ID: An interference suppression model for domain adaptation person re-identification under inclement weather condition, in: 2021 IEEE International Conference on Multimedia and Expo, ICME, 2021, pp. 1–6.
[10] E.J. McCartney, Optics of the atmosphere: Scattering by molecules and particles, Phys. Today 30 (5) (1977) 76.
[11] S.G. Narasimhan, S.K. Nayar, Chromatic framework for vision in bad weather, in: Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, CVPR, 2000, pp. 598–605.
[12] S.G. Narasimhan, S.K. Nayar, Vision and the atmosphere, Int. J. Comput. Vis. 48 (3) (2002) 233–254.
[13] K. He, J. Sun, X. Tang, Single image Haze removal using dark channel prior, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2011) 2341–2353.
[14] Q. Zhu, J. Mai, L. Shao, A fast single image Haze removal algorithm using color attenuation prior, IEEE Trans. Image Process. 24 (11) (2015) 3522–3533.
[15] R. Fattal, Dehazing using color-lines, ACM Trans. Graph. 34 (1) (2014) 1–14.
[16] B. Cai, X. Xu, K. Jia, C. Qing, D. Tao, DehazeNet: An end-to-end system for single image Haze removal, IEEE Trans. Image Process. 25 (11) (2016) 5187–5198.
[17] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, M.-H. Yang, Single image Dehazing via multi-scale convolutional neural networks, in: European Conference on Computer Vision, 2016, pp. 154–169.
[18] H. Zhang, V.M. Patel, Densely connected pyramid Dehazing network, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2018, pp. 3194–3203.
[19] Y. Li, Q. Miao, W. Ouyang, Z. Ma, H. Fang, C. Dong, Y. Quan, LAP-net: Level-aware progressive network for image Dehazing, in: 2019 IEEE/CVF International Conference on Computer Vision, ICCV, 2019, pp. 3275–3284.
[20] Y. Pang, J. Nie, J. Xie, J. Han, X. Li, BidNet: Binocular image Dehazing without explicit disparity estimation, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2020, pp. 5930–5939.
[21] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, M.-H. Yang, Gated fusion network for single image Dehazing, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2018, pp. 3253–3261.
[22] X. Liu, Y. Ma, Z. Shi, J. Chen, GridDehazeNet: Attention-based multi-scale network for image Dehazing, in: 2019 IEEE/CVF International Conference on Computer Vision, ICCV, 2019, pp. 7313–7322.
[23] Y. Qu, Y. Chen, J. Huang, Y. Xie, Enhanced Pix2pix Dehazing network, in: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2019, pp. 8152–8160.
[24] X. Qin, Z. Wang, Y. Bai, X. Xie, H. Jia, FFA-net: Feature fusion attention network for single image Dehazing, in: Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, AAAI, 2020, pp. 11908–11915.
[25] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, M.-H. Yang, Multi-scale boosted Dehazing network with dense feature fusion, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2020, pp. 2154–2164.

[26] H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, L. Ma, Contrastive learning for compact single image Dehazing, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2021 pp. 10546–10555.
[27] C. Wang, H.-Z. Shen, F. Fan, M.-W. Shao, C.-S. Yang, J.-C. Luo, L.-J. Deng, EAA-Net: A novel edge assisted attention network for single image dehazing, Knowl.-Based Syst. 228 (2021) 107279.
[28] L. Li, Y. Dong, W. Ren, J. Pan, C. Gao, N. Sang, M.-H. Yang, Semi-supervised image Dehazing, IEEE Trans. Image Process. 29 (2020) 2766–2779.
[29] Y. Shao, L. Li, W. Ren, C. Gao, N. Sang, Domain adaptation for image Dehazing, in: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2020, pp. 2805–2814.
[30] Z. Chen, Y. Wang, Y. Yang, D. Liu, PSD: Principled synthetic-to-real Dehazing guided by physical priors, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2021, pp. 7180–7189.
[31] D. Berman, T. Treibitz, S. Avidan, Non-local image Dehazing, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016 pp. 1674–1682.
[32] Y. Jiang, C. Sun, Y. Zhao, L. Yang, Image dehazing using adaptive bi-channel priors on superpixels, Comput. Vis. Image Underst. 165 (2017) 17–32.
[33] Z. Zhu, H. Wei, G. Hu, Y. Li, G. Qi, N. Mazur, A novel fast single image Dehazing algorithm based on artificial multiexposure image fusion, IEEE Trans. Instrum. Meas. 70 (2021) 5001523.
[34] M. Zheng, G. Qi, Z. Zhu, Y. Li, H. Wei, Y. Liu, Image Dehazing by an artificial image fusion method based on adaptive structure decomposition, IEEE Sens. J. 20 (2020) 8062–8072.
[35] T. Si, F. He, H. Wu, Y. Duan, Spatial-driven features based on image dependencies for person re-identification, Pattern Recognit. 124 (2022) 108462.
[36] H. Wu, F. He, Y. Duan, X. Yan, Perceptual metric-guided human image generation, Integr. Comput.-Aided Eng. 29 (2) (2022) 141–151.
[37] Y. Pan, F. He, H. Yu, Learning social representations with deep autoencoder for recommender system, World Wide Web 23 (4) (2020) 2259–2279.
[38] R. Quan, X. Yu, Y. Liang, Y. Yang, Removing raindrops and rain streaks in one go, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR, 2021, pp. 9147–9156.
[39] S. Zhang, F. He, DRCDN: learning deep residual convolutional dehazing networks, Vis. Comput. 36 (9) (2020) 1797–1808.
[40] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, Y. Fu, Residual dense network for image super-resolution, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 2472–2481.
[41] S. Woo, J. Park, J.-Y. Lee, I.S. Kweon, CBAM: Convolutional block attention module, in: European Conference on Computer Vision, 2018, pp. 3–19.
[42] X. Wang, R. Girshick, A. Gupta, K. He, Non-local neural networks, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 7794–7803.
[43] W. Xiang, J. Huang, X.-S. Hua, L. Zhang, Part-aware attention network for person re-identification, in: Proceedings of the Asian Conference on Computer Vision, ACCV, 2020.
[44] Y. Yang, Y. Zhuang, Y. Pan, Multiple knowledge representation for big data artificial intelligence: framework, applications, and case studies, Front. Inf. Technol. Electron. Eng. 22 (12) (2021) 1551–1558.
[45] J.-Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: 2017 IEEE International Conference on Computer Vision, ICCV, 2017, pp. 2242–2251.
[46] W. Ren, J. Pan, H. Zhang, X. Cao, M.-H. Yang, Single image dehazing via multi-scale convolutional neural networks with holistic edges, Int. J. Comput. Vis. 128 (2020) 240–259.
[47] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014, arXiv preprint arXiv:1409.1556.
[48] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, Z. Wang, Benchmarking single-image Dehazing and beyond, IEEE Trans. Image Process. 28 (1) (2019) 492–505.
[49] Y. Zhang, L. Ding, G. Sharma, Hazerd: an outdoor scene dataset and benchmark for single image dehazing, in: 2017 IEEE International Conference on Image Processing, ICIP, 2017, pp. 3205–3209.
[50] D.P. Kingma, J.L. Ba, Adam: A method for stochastic optimization, in: International Conference on Learning Representation, ICLR, 2015.
[51] Z. Wang, A. Bovik, H. Sheikh, E. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[52] Y. Yu, H. Liu, M. Fu, J. Chen, X. Wang, K. Wang, A two-branch neural network for non-homogeneous Dehazing via ensemble learning, in: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPRW, 2021, pp. 193–202.