

Exploring Fuzzy Priors from Multi-Mapping GAN for Robust Image Dehazing

Shengdong Zhang, Xiaoqin Zhang, Wenqi Ren, Li Zhao, En Fan, and Feng Huang

Abstract—Single image dehazing has been extensively studied. While Convolutional Neural Networks (CNNs) have driven notable progress in single image dehazing, their performance remains fundamentally constrained by the limited local receptive fields of convolutional operations, which impede the capture of global structural dependencies. In contrast, Generative Adversarial Networks (GANs) have demonstrated exceptional capabilities in image synthesis, offering global insights into structure, texture, and color. The fuzzy prior, a probabilistic knowledge acquired through adversarial training in GANs, plays a pivotal role in robust dehazing. Motivated by this, we propose the Fuzzy Prior Guided Dehazing Network (FPGDN). Our framework begins with a novel module that distills the fuzzy prior by translating an edge map into a color image, simultaneously capturing global structural, local textural, and color information. Subsequently, a dehazing network is constructed, leveraging this fuzzy prior. While the fuzzy prior captures rich color and texture features, the generated images may exhibit color shifts relative to the original scene. To remedy this, a CNN network is employed to capture local nuances and refine the dehazing outcome. Extensive experiments substantiate that the proposed FPGDN achieves superior dehazing performance on a variety of real and synthetic hazy images.

Index Terms—Fuzzy prior, dehazing, image restoration, edge map, texture information, GAN.

I. INTRODUCTION

HAZE is a common weather phenomenon caused by air particles, which can be described using following model:

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

where I represents the observed hazy image, J represents the clean image, and the transmission map t defines the

Manuscript received 18 December 2024; revised 23 March 2025, 1 May 2025, and 14 July 2024; accepted 4 August 2025. This work is supported by the Zhejiang Provincial Natural Science Foundation under Grant Numbers LTGG24F020005, LDT23F02024F02, and LGG22F010004, and Zhejiang Provincial Department of Education Research Program Project under Grant Number Y202351480, in part by the Taizhou Technology Plan Project under Grant Number 24gyb41, in part by the National Key Research and Development Program Project of China under Grant Number 2024YFC3306902, in part by the National Natural Science Foundation of China under Grant Numbers U24A20242 and 62475241. (Corresponding authors: Wenqi Ren; Xiaoqin Zhang.)

S. Zhang is with the School of Intelligent Engineering, Shaoxing University, Shaoxing 312000, Zhejiang, China (E-mail: shengdong1986@gmail.com).

X. Zhang and L. Zhao are with the College of Computer and Artificial Intelligence, Wenzhou University, Wenzhou 325035, China (E-mail: zhangxiaoqinnan@gmail.com; lizhao@wzu.edu.cn).

W. Ren is also with the School of Cyber Science and Technology, Sun Yat-sen University, Shenzhen 518110, China (E-mail: rwq.renwenqi@gmail.com).

E. Fan is with the School of Intelligent Engineering, Shaoxing University, Shaoxing 312000, Zhejiang, China (E-mail: e-fan@sxu.edu.cn).

F. Huang is with Zhejiang Province Taizhou Technician College, Taizhou 318000, Zhejiang, China (e-mail: huangxvfeng@hdu.edu.cn).

light reflected from the objects. The goal of single image dehazing is to restore the clean image J . Haze-induced contrast degradation not only obscures image details but also severely compromises the performance of high-level vision tasks, including object detection and semantic segmentation, by introducing erroneous feature representations. To reduce the effect of haze, many dehazing methods [1]–[7] have been proposed. Traditional dehazing methods [1], [8], [2] rely on priors extracted from hazy or clean images to restore the clean image. However, these methods lack robustness in complex scenarios, such as scenes containing white objects or regions with color distributions similar to the atmospheric light. To improve the robustness of transmission map estimation, the CNNs are employed to remove haze. Ren et al. designed a multi-scale model to estimate transmission map [9]. Cai et al. designed a light model for transmission map estimation [10]. However, these methods are affected by the accuracy of the estimated transmission map. To further enhance the quality of dehazing, end-to-end dehazing methods [5], [11]–[14] have been proposed. These methods establish a direct mapping between hazy images and their corresponding clean images. However, they are limited by the local nature of convolution and fail to capture long-range dependencies. Additionally, current learning-based dehazing methods are suffering from the domain shift problem, which hampers their ability to produce high-quality dehazed results on natural haze images. The application scenarios for industrial systems [15] are highly diverse, including intelligent traffic systems, surveillance cameras, and inspection robots. This diversity further diminishes the performance of existing learning-based dehazing methods.

To tackle these challenges, this paper introduces an innovative dehazing network designed to capture the variability of real-world conditions and mitigate the domain shift problem. Generative Adversarial Networks (GANs) [16] are capable of producing photo-realistic images that encompass global structure, fine details, and vibrant colors, which aligns well with the goal of dehazing ‘to restore a realistic image’. Leveraging the strengths of GANs, we propose incorporating their features to enhance the dehazing process. While GANs transform a noise map into a photo-realistic image, dehazing transforms a hazy image into a haze-free one. A straightforward approach would be to encode a hazy image into a latent code and then decode it into an image. However, the lack of real-world paired haze-free images makes this approach unfeasible. Instead, we identified a large dataset of simulated haze images. Although these simulated images differ from real haze, they can still be utilized. As demonstrated in Section III, the edge map extracted from a hazy image closely resembles that of

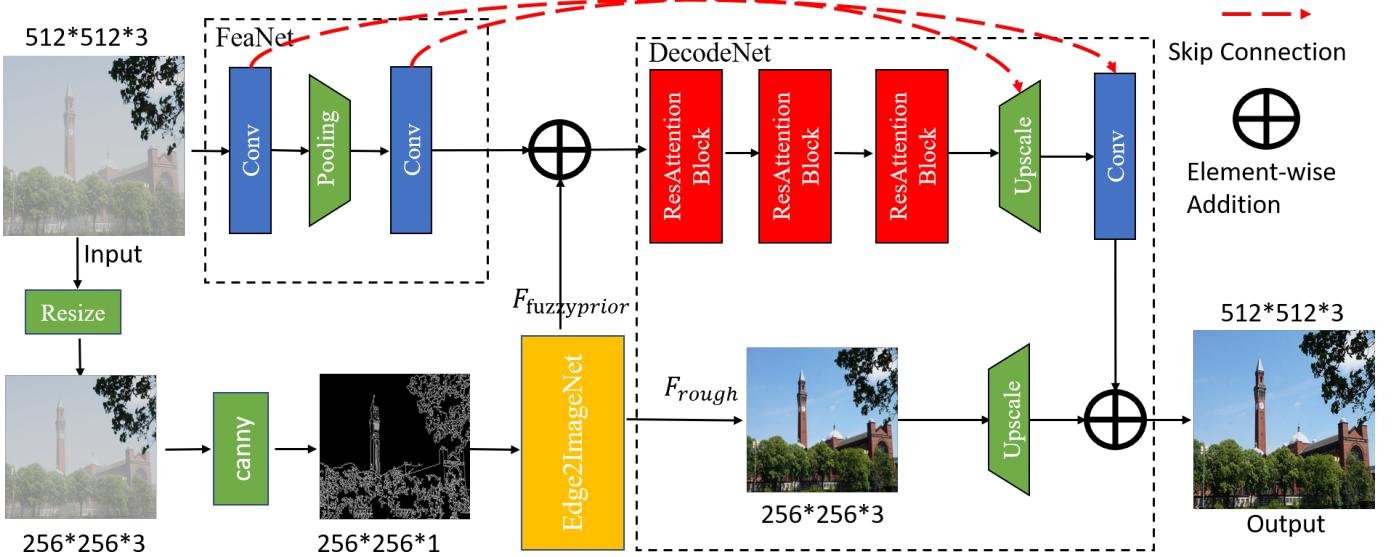


Fig. 1: The architecture of the proposed fuzzy prior Guided Dehazing Network (FPGDN). The proposed network contains three parts: FeaNet, Edge2ImageNet, and DecodeNet.

its corresponding clean image. Based on this observation, we propose a GAN model that translates an edge map into an image, thereby providing highly relevant features (fuzzy prior) for dehazing. Relying solely on edge maps for dehazing results can lead to ambiguous colors. To address this issue, we introduce a dehazing model that extracts features from hazy images and then restores the final dehazed result using these features in conjunction with contextual priors. The main contributions of this paper are summarized as follows:

- 1) We propose a GAN based model (Edge2ImageNet) to extract fuzzy prior, which contains global structure, local texture, and color information for dehazing. Traditional dehazing models, which rely solely on a small set of haze images, fail to capture the complex distribution of natural industrial images, resulting in suboptimal results. GAN has been used to synthesis real images, which proves that GAN can capture the distribution of natural images.
- 2) We propose to alleviate domain shift problem in edge domain. Edge maps are less affected by haze, which motivates us to address the domain shift problem in the edge domain.
- 3) Extensive experiments are conducted to show the high performance of the proposed method.

II. RELATED WORK

A. Single Image Dehazing

Single image dehazing is a hot topic in computer vision, which has received great attention owing to its wide applications. The existing single image dehazing methods can be divided into model-based and end-to-end dehazing methods. The model-based dehazing methods [1], [2], [10], [9] focus on the estimation of airlight and transmission map, which can be used to recover the final dehazing image. Dark channel prior (DCP) is the most famous model-based dehazing method,

which can obtain a good dehazed result for most hazy images. Cai et al. designed a multi-scale model [10] to estimate the transmission map. However, these methods cannot estimate the transmission map accurately, which leads to a poor dehazed result. To address this issue, end-to-end dehazing methods [3], [5], [12], [13], [6], [17]–[21] have been proposed, which obtain a dehazed result from an input hazy image directly. AECR [13] is a UNet-like dehazing method based on contrastive Learning. Liu et al. presented a multi-scale attention dehazing network [22]. Zhang et al. developed a novel dehazing model [5] that embeds the physical model into the network to obtain the dehazed result. AirNet is an unsupervised dehazing method that assumes that the degradation of the haze is uniform. However, the density of the haze is not uniform, which often results in poor dehazing quality. Dehamer [23] is a recent transformer-based dehazing method, which can capture long dependency. Gridformer [24] introduces a grid structure combined with transformer blocks to capture both local and global information for image restoration. MB-TaylorFormer [25] and its extension [26] leverage Taylor expansion to enhance the efficiency of transformer-based architectures. Hierarchical contrastive learning is employed to improve the quality of dehazed results. UIDM [7], an unsupervised dehazing method based on Fuzzy Clustering-Line Priors, further contributes to this field. However, all these methods overlook the structural information of clean images. GANs have demonstrated their effectiveness in capturing the distribution of clean images in tasks such as super-resolution [27] and deblurring [28]. Inspired by these successes, this paper proposes integrating GANs to learn the global structural information of clean images and transferring this knowledge to the dehazing network.

B. Edge Prior for Image Restoration

Edge prior is an important statistic law that can be used to improve the quality of the restored image. Edge prior

has shown its effectiveness in image inpainting [29], super-resolution [30], and dehazing [17], [5]. To employ the edge prior, Nazeri et al. designed a network to predict the missed details of the input edge map [29]. To improve super-resolution quality, Fang et al. developed a network to predict an edge map from the input low-resolution image [30]. Zhang et al. proposed a novel loss to preserve the edge information of the clean image. In addition, Ren et al. proposed a holistic edge network to refine the transmission map [17]. Unlike these methods, this paper proposes learning a fuzzy prior from the map between the edge map and the corresponding color image.

III. FUZZY PRIOR GUIDED DEHAZING NETWORK

The framework of the proposed Fuzzy Prior Guided Dehazing Network (FPGDN) is shown in Fig. 1. The proposed model consists of three parts. The first part is FeaNet, which extracts shallow features from the hazy input image; the second part is Edge2ImageNet, which extracts fuzzy prior from the edge map; the third part is DecodeNet, which generates the dehazed result based on the features from FeaNet and fuzzy prior from Edge2ImageNet. Specifically, the FPGDN consists of two stages. The first stage extracts features from the inputs, and the second stage generates the dehazed result. For a hazy input, it is resized to 1/2 of the original size, and an edge map is extracted from it. Then, local and color information is extracted from the hazy image via FeaNet, and global features are extracted via Edge2ImageNet. Finally, the DecodeNet reconstructs the dehazed image by upsampling the fused features and refining the texture details through a convolution operation. In the sections that follow, we elaborate on the motivations behind the proposed model and provide a detailed explanation of each module.

A. Motivations

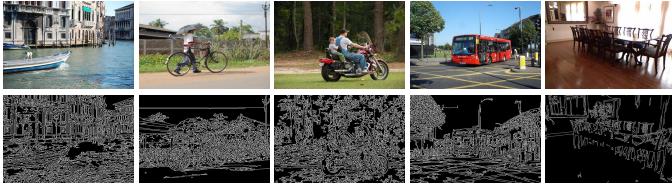


Fig. 2: A visual example of images and the corresponding edge maps. As shown, we can recognize objects from the edge maps.

Natural industry hazy images are diversity. However, the simulated haze images are with limited diversity. Furthermore, the simulated haze images are different from real haze images. To boost the adaptive ability of the proposed model, a GAN model is employed to capture the complex distribution of real images. GAN has shown its ability in image synthesise, which demonstrates GAN can capture the complex distribution of real images. The features from GAN can be used to synthesize real images, which can provide global structure, local texture and color information. As simulated haze color images are different from real color haze images, we seek to alleviate the

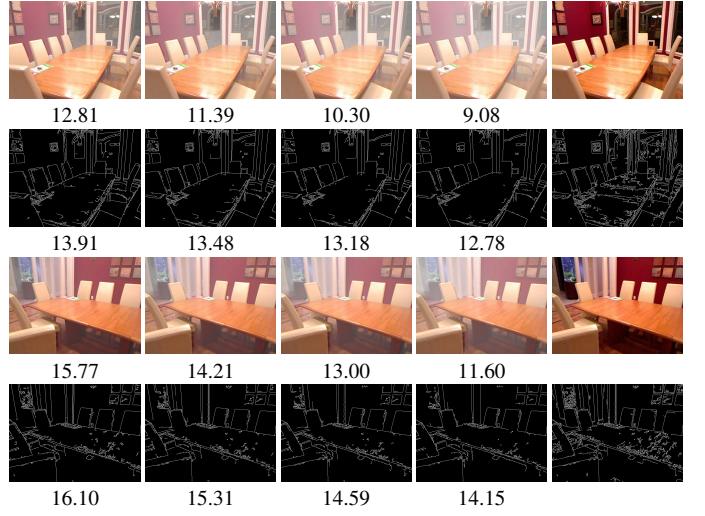


Fig. 3: Impact of haze on color images and edge maps. The PSNR (defined in Section III-A) between clean and hazy images drops significantly for color data but remains stable for edge maps, demonstrating edge invariance to haze.

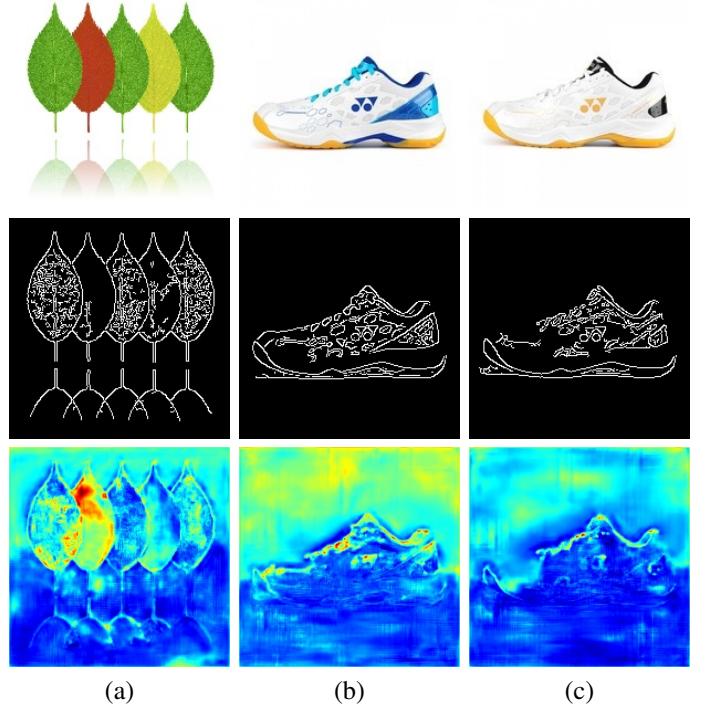


Fig. 4: A visual example of color ambiguity for similar edge maps. The color images are shown in the first row, and the corresponding color edge maps are shown in the second row, the third one shows the heatmap of Fuzzy Prior. As shown, similar edge maps with different color appearances lead to the blue color in the heatmap.

difference in edge domain. Edge maps exhibit higher similarity than original images, as shown in Fig. 3. We employ PSNR to quantify the difference between images, calculated as:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right), \quad (2)$$

where MSE is the Mean Squared Error: $MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - R(i, j)]^2$. Compared to color images, edge maps are less affected by haze. While the PSNR of color images drops significantly, that of edge maps decreases only slightly. This observation motivates our approach to restore color images from edge maps.

Natural scenes often exhibit color ambiguity due to contextual or environmental factors as shown in Fig. 4. For instance, the color of a leaf can vary across seasons (green in summer, yellow in autumn, brown in winter), yet its edge structure remains consistent as shown in Fig. 4(a). Moreover, Fig. 4(b) and (c) illustrate that the same edge map (e.g., shoe contours) can correspond to multiple plausible color appearances, highlighting the inherent ambiguity in edge-to-color mapping. To model this phenomenon, we propose a Gaussian Mixture Model (GMM) conditioned on edge maps, which captures the probabilistic relationship between edges and plausible color/texture variations:

$$P(C|E) = \sum_{i=1}^n w_i \cdot \mathcal{N}(C, \mu_i(E), \Sigma_i(E)) \quad (3)$$

where the weights w_i satisfy $\sum w_i = 1$, $\mathcal{N}(C, \mu_i(E), \Sigma_i(E))$ is the i -th Gaussian distribution with mean $\mu_i(E)$ and the covariance $\Sigma_i(E)$. Compared with Gaussian Model, GMM are universal approximators of smooth densities. By the Wiener's Tauberian theorem they can approximate any continuous distribution over color spaces \mathcal{R}^3 arbitrarily well as $n \rightarrow +\infty$. This justifies their capacity to model complex, discontinuous edge-to-color relationships.

To model the ambiguity and fuzzy [31], [32] between the edge map and color images, we introduce a fuzzy prior, which is a latent probability distribution over color hypotheses conditioned on edges. It models the ambiguity in color assignments (e.g., a tree edge could correspond to green, brown, or autumn hues) and is learned implicitly through adversarial training. Mathematically, it is defined as:

$$P_{fuzzy}(F|E) = \mathbb{E}_{\sim P(C|E)}[G(E, z)], \quad (4)$$

where E is the edge map, F is the feature maps from the last layer of the generator G , z is a latent noise vector sampled from a Gaussian Mixture distribution $P(C|E)$. To further explain the fuzzy prior, we show the heatmaps of fuzzy prior in the third row of Fig. 4. As shown, in regions with multiple colors corresponding to the edges, the heatmap appears blue, while for a single color corresponding to the edges, it appears yellow. High color diversity requires fuzzy priors to cover multiple peak distributions, leading to dispersed activation values and a reduced overall response. Conversely, when color certainty is high, the fuzzy prior follows a unimodal distribution. In this case, color restoration relies on local texture and brightness information, and the heatmap shows a

strong feature response. To address color ambiguity and fuzzy problem, we propose to employ the color information in the original hazy image to help the dehazing network obtain a meaningful result.

B. Details of Network Structure

The proposed model consists of three parts: FeaNet, Edge2ImageNet, and DecodeNet. The details of each module are described below.

Feature Extraction Network. The feature extraction network is designed to obtain representative features from the hazy image, which can provide additional constraints for image dehazing. This network is called FeaNet and consists of two convolution operations. The first convolution operation project a hazy image into feature space, and then a pooling operation is applied to feature map. After the pooling operation, we apply another convolution operation to refine the feature map.

$$F_{features}, F_h = FeaNet(I_{color}), \quad (5)$$

where $FeaNet$ represents the feature extraction network, I_{color} represents the color hazy image, and $F_{features}$ represents the shallow features, F_h represents the high resolution features.

Edge to Image Network. Edge information serves as a critical prior for image restoration [30]. However, directly reconstructing photorealistic color images from edge maps remains challenging due to inherent ambiguities in color and texture mapping. To address this fundamental limitation, we propose Edge2ImageNet, a specialized generative network that translates structural edge information into plausible color images while preserving geometric fidelity. As illustrated in Fig. 5, Edge2ImageNet employs a dual-branch architecture comprising:

- 1) **Encoder:** A convolutional network that progressively transforms the edge map into a latent representation. The encoder consists of multiple convolutional layers, with initial layers extracting shallow features and subsequent strided convolutions (stride=2) compressing spatial resolution while expanding channel depth.
- 2) **GAN-Based Decoder:** A StyleGANv2-inspired generator that synthesizes images from latent codes while extracting intermediate feature representations.

The encoding process is formally defined as:

$$z_{code}, F_{low} = Encoder(I_{edge}), \quad (6)$$

where $Encoder$ denotes our convolutional encoding network, I_{edge} represents the input edge map, and $L_{code} \in \mathbb{R}^{4 \times 4 \times 512}$ is the resulting latent code, F_{low} is the feature maps from shallower layers of Encoder.

This latent code is then transformed through eight fully connected layers to produce style vectors:

$$\mathbf{w} = MLP(z_{code}) \quad (7)$$

The decoder subsequently generates both a color image and feature-level priors:

$$F_{fuzzyprior}, F_{rough} = Decoder(\mathbf{w}, F_{low}), \quad (8)$$

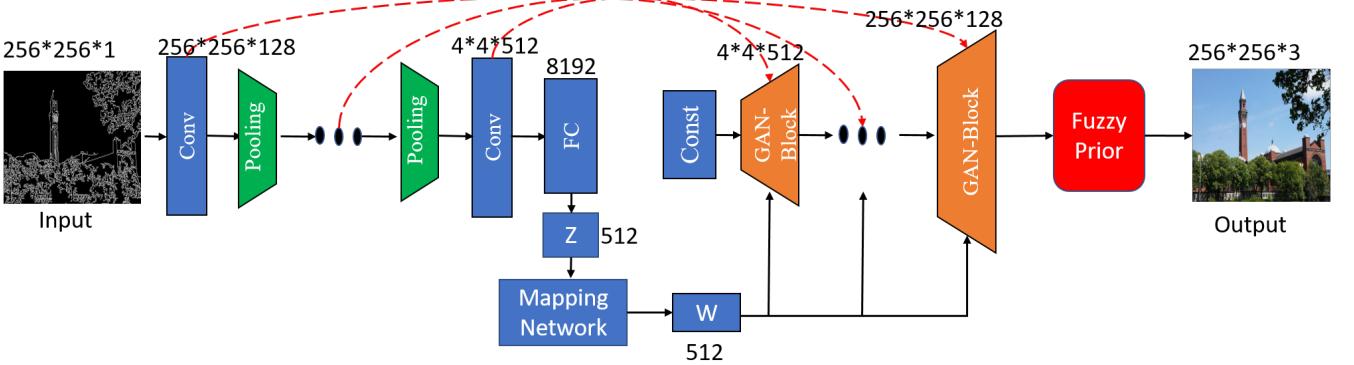


Fig. 5: The framework of the proposed Edge2ImageNet. Conv represents a convolution operation followed with a ReLU activation function. FC is supposed to be “Fully Connected”. “...” represents this block are repeated several times. We bring the Mapping Network and the GAN blocks from StyleGANv2, which is critical for improving dehazing quality.

where Decoder denotes our GAN-based synthesis network, $F_{\text{fuzzyprior}}$ represents the fuzzy prior extracted from penultimate layers, and F_{rough} denotes the output RGB image.

The full pipeline is compactly expressed as:

$$F_{\text{fuzzyprior}}, F_{\text{rough}} = \text{Edge2ImageNet}(I_{\text{edge}}), \quad (9)$$

where Edge2ImageNet integrates the complete edge-to-image translation workflow.

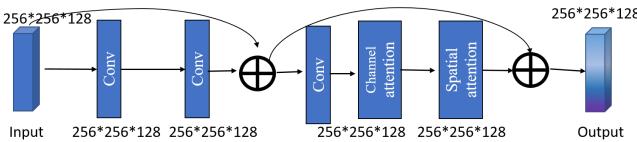


Fig. 6: The framework of the proposed ResAttionBlock.

Decode Network. After obtaining the fuzzy prior and encoded features from FeaNet, we use following equation to fuse them:

$$F_{\text{fuse}} = F_{\text{fuzzyprior}} + F_{\text{features}}, \quad (10)$$

where F_{fuse} represents the fused feature map. The decoding network takes the fused features, the shallow features from the FeaNet, and the F_{rough} as inputs, and then it outputs the final dehazed result.

$$\mathcal{F}(I) = \text{DecodeNet}(F_{\text{fuse}}, F_{\text{rough}}, F_h), \quad (11)$$

where F_h represents the output of the first convolution layer, $\mathcal{F}(I)$ represents the final dehazed result reconstructed by the proposed model. To boost the dehazing ability of the proposed model, we adopt the channel attention and spatial attention from FFA-Net [19] to form ResAttionBlocks. The detail of ResAttionBlock is shown in Fig. 6. As shown, we use two convolution operations to refine the feature maps and then add them to the original input feature maps. We use spatial and channel attention to further boost the representative ability of the proposed ResAttionBlocks.

C. Training Losses

To train the Edge2ImageNet, two training losses are adopted: the content loss \mathcal{L}_{con} loss and the adversarial loss \mathcal{L}_{adv} . The \mathcal{L}_{con} loss is defined as:

$$\mathcal{L}_{\text{con}} = \frac{1}{N} \sum_{i=1}^N \|\mathcal{F}(E_i, \Theta) - C_i\|_1, \quad (12)$$

where N is the number of training samples, $\|\cdot\|_1$ is L_1 norm, C defines the ground-truth color image, E_i denotes the input edge map, and Θ represents the weights of the learned filters of Edge2ImageNet.

To restore a high-quality color image, the \mathcal{L}_{adv} is defined in a manner similar to that in the prior GAN works:

$$\mathcal{L}_{\text{adv}} = \frac{1}{N} \sum_{i=1}^N \log(1 - D(\mathcal{F}(E_i, \Theta))), \quad (13)$$

where \mathcal{F} represents the proposed Edge2ImageNet, Θ represents the weights of the learned filters, and D represents the discriminator. The final loss for Edge2ImageNet can be defined as:

$$\mathcal{L}_{\text{Edge2ImageNet}} = \mathcal{L}_{\text{con}} + \lambda \mathcal{L}_{\text{adv}} \quad (14)$$

where λ controls the importance of \mathcal{L}_{adv} . After training the Edge2ImageNet, the discriminator can be updated by the following loss:

$$\max_D \frac{1}{N} \sum_{i=1}^N (\log(1 - D(\mathcal{F}(E_i, \Theta))) + \log(D(C_i))). \quad (15)$$

After the Edge2ImageNet is trained, it is embedded into the proposed model, and then the proposed model is fine-tuned with the restoration loss:

$$\mathcal{L}_{\text{rest}} = \frac{1}{N} \sum_{i=1}^N \|\mathcal{D}\mathcal{N}(I_i, E_i, \Omega) - J_i\|_1, \quad (16)$$

where $\mathcal{D}\mathcal{N}$ denotes the whole dehazing network, Ω represents the weights of the learned filters of the proposed model, and J denotes the clean images.

TABLE III: Quantitative comparison results of various dehazing methods and the proposed method on the outdoor images from the RESIDE dataset. The best result is marked with red color.

	DCPDN [5]	MSCNN [9]	DehazeNet [10]	DCP [1]	CAP [8]	AOD-Net [3]	PhysicsGAN [39]	UIDM [7]
PSNR	20.24	18.64	19.41	19.13	22.27	19.06	20.04	25.56
SSIM	0.89	0.78	0.85	0.81	0.62	0.90	0.92	0.97
	GASDN [40]	GDN [22]	EPDN [11]	GFN [4]	DeHamer [23]	AirNet [12]	SGID-PFF [6]	FPGDN
PSNR	28.13	30.08	23.82	21.55	24.85	25.61	30.20	31.76
SSIM	0.96	0.98	0.89	0.84	0.95	0.91	0.97	0.97



Fig. 7: Visual comparisons of various dehazing methods on the RESIDE dataset.

Fig. 7. As demonstrated, the SGID-PFF yields dehazed results that appear overly dark. In comparison, AirNet, DeHamer, and AOD-Net produce dehazed images that retain a light haze. Additionally, DCP generates dehazed images with noticeable color distortion, while DCPDN generates dehazed images that are over-enhanced. In contrast, the dehazed results from our proposed method are closely similar to the ground truth haze-free images.

C. Visual Evaluation on Natural Haze Images

The effectiveness of the proposed model has been demonstrated on simulated haze images. However, the natural haze images are different from the simulated ones. So, it is critical to show the dehazing performance on natural haze images. To perform such an experiment, this paper chooses some natural haze images that are difficult to process by traditional and learning-based dehazing methods.

We choose two challenging hazy images, which contain dense haze. For these two images, we choose sixteen state-of-the-art dehazing methods to show the high performance of the proposed method. These two images are typical transportation application scenarios, which contain roads or trains. Traditional dehazing methods [1], [2] based on hand-crafted prior cannot obtain high quality dehazing results. While CAP [8] cannot remove haze well and retain a lot of haze in dehazed results. Learning-based dehazing methods include supervised and unsupervised dehazing methods. As shown in Fig. 8, we can see that JarvisIR [42], DehazeXL [42], AEGR-Net [13], AirNet, FFA-Net, GDN, MSBDN, Dehamer [23] cannot remove haze well and retain a lot of haze in final dehazed results. The EPN can remove haze better. However, it also removes the image details from the dehazed results. DA [43] reduces the difference between simulated and real hazy images, and restores a high bright dehazed results. However, we can still observe the haze from the dehazed results. PhysicsGAN [39]

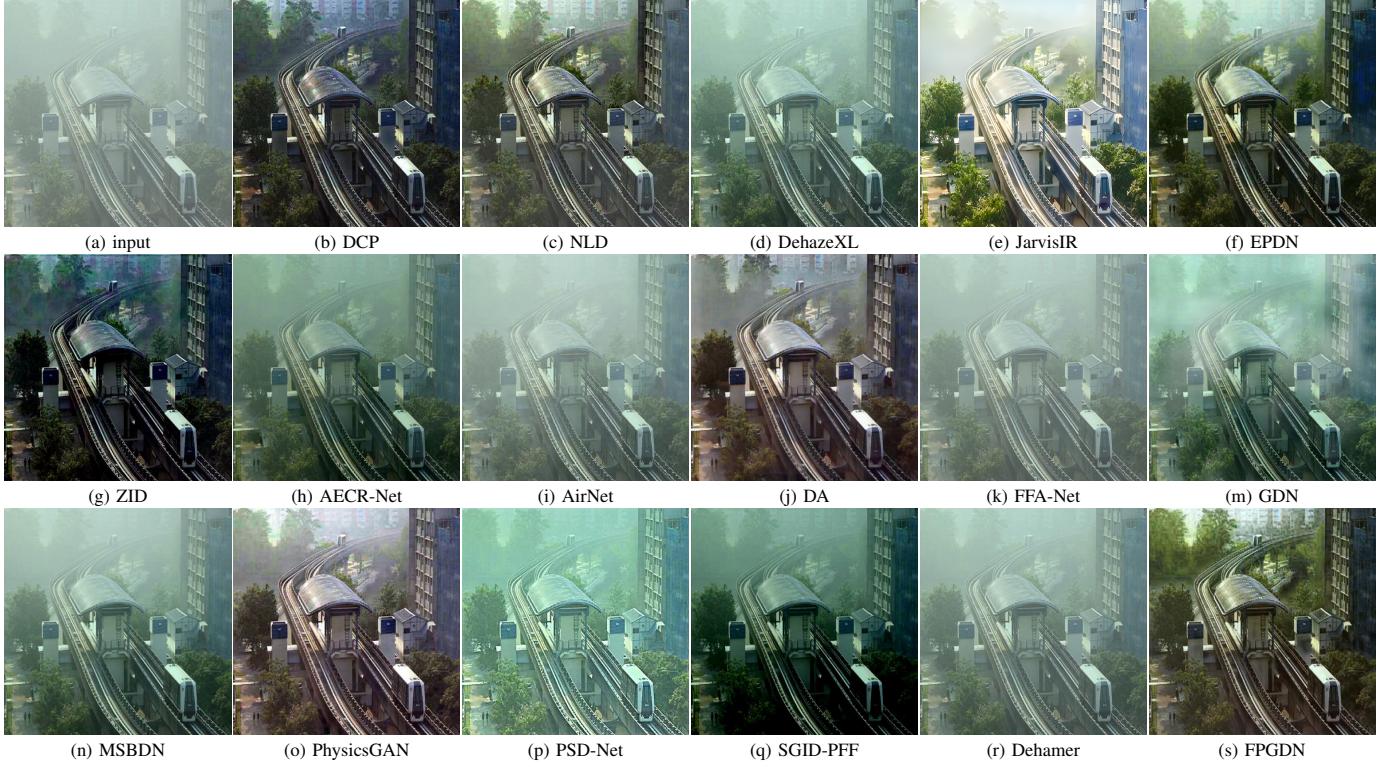


Fig. 8: Visual results of some recently dehazing methods and the proposed method. As shown, we note the dehazed results of JarvisIR, DehazeXL, CAP, AEZR, AirNet, FFA-Net, GDN, MSBDN, and Dehamer contain a lot of haze. The dehazed results of NLD, DA, PhysicsGAN, and EPDN have less haze. However, these methods often fail to remove haze from far areas. The dehazed results of DCP, ZID, and EPDN can remove more haze. However, these dehazed results often contain color distortions. In contrast, the proposed method removes haze more completely.

fails to remove haze from far areas. SGID-PFF has the problem of overestimate or underestimate haze problem, which results in the problem of over-enhancing the hazy image or keeping haze in dehazed results. PSD improves the dehazing results by combining traditional priors. However, we can see that the dehazed results of PSD-Net [14] contain color distortion. The dehazed results of ZID [36] can better remove the haze. However, the dehazed results of ZID contains a lot of color distortion. DehazeXL and JarvisIR both struggle to eliminate haze in distant regions, leaving substantial residual fog that obscures fine details. In contrast, Visual dehazed results demonstrate that FPGDN removes haze more completely while preserving fine details, such as textures in distant areas.

TABLE IV: Quantitative no-reference metric results for the dehazed images produced by various dehazing methods. The best result is marked with red color.

Metric	NLD	DA	EPDN	AirNet	Dehamer	FPGDN
FADE ↓	0.39	0.44	0.32	1.36	1.17	0.25
NIQE ↓	3.53	4.27	3.75	4.14	3.96	3.89

To further demonstrate the superiority of our proposed method, we have conducted a no-reference quality assessment. We evaluated the visual quality of the dehazed images using no-reference metrics, specifically the Fog Aware Density Evaluator (FADE) and the Natural Image Quality Evaluator (NIQE). The quantitative results of these evaluations are

presented in Table IV for dehazed results shown in Fig. 8. In terms of FADE scores, our proposed model yields a more haze-free result compared to other methods. In terms of NIQE scores, our method achieves a performance that ranks third among the evaluated approaches. In the future, we intend to utilize NIQE as a guidance mechanism to further enhance the quality of dehazing.

We show another application scene of the dehazing in Fig. 9. As shown, the dehazed results of traditional dehazing methods, such as DCP and NLD, often have the problem of over-enhancement. The traditional learning-based method (CAP) cannot reomove haze completely and contains much haze in the dehazed result. Deep-learning-based methods, such as PhysicsGAN, MSBDN, JarvisIR, and DehazeXL, have the similar problem as the traditional learning-based method. Some learning based methods, such as EPDN and SGID-PFF can remove haze better. However, the dehazed results have a problem of losing image details. Unsupervised dehazing method also cannot remove haze completely. On contrast, the proposed method can remove haze better while preserving the details.

To demonstrate superiority on real-world hazy images, we evaluate our method on the O-Haze dataset – featuring haze synthesized by professional machinery. Quantitative results (Table V) indicate that DehazeXL [42] ranks second and DNMGDT [44] third in PSNR; our proposed method estab-

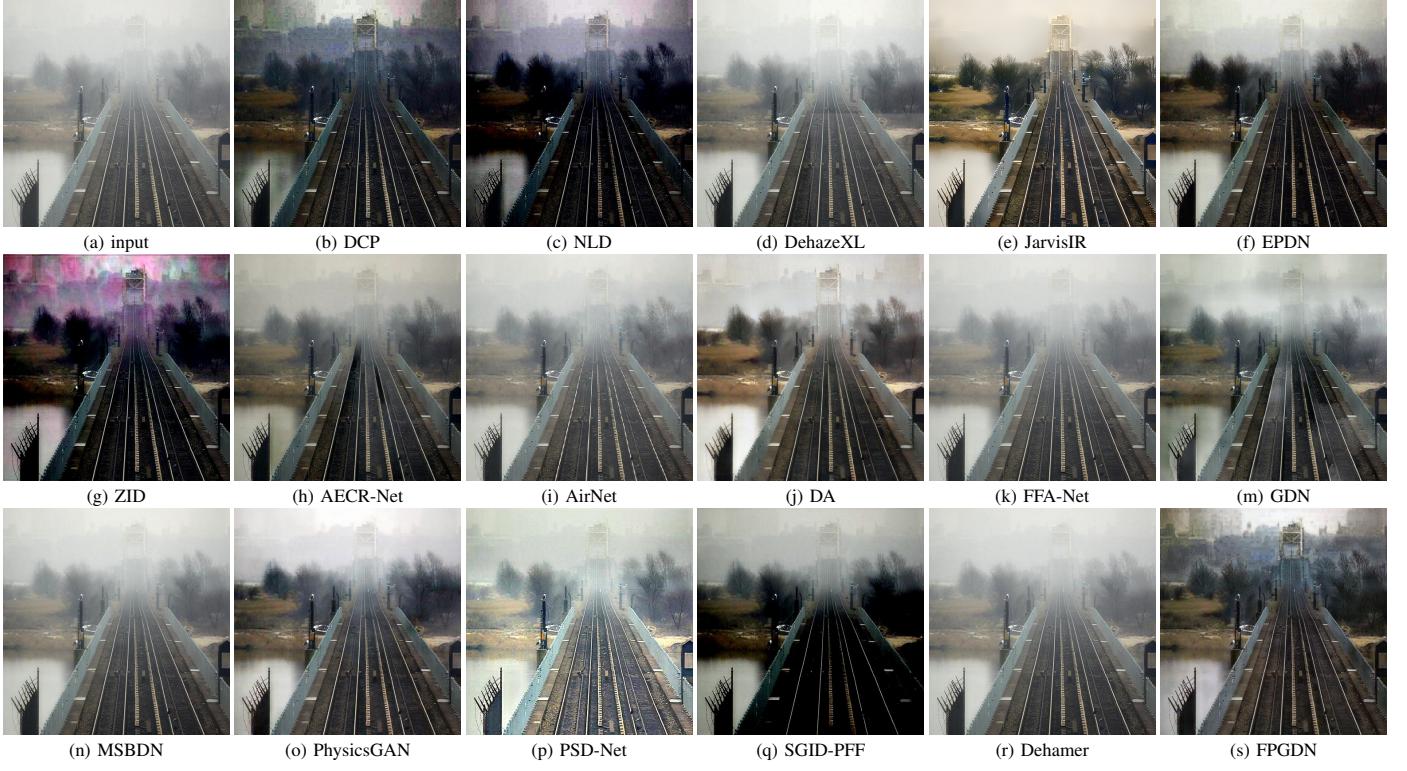


Fig. 9: Visual results of some recently dehazing methods and the proposed method. As shown, we note the dehazed results of JarvisIR, DehazeXL, CAP, AEZR-Net, AirNet, FFA-Net, GDN, MSBDN, and Dehamer contain a lot of haze. The dehazed results of NLD, DA, PhysicsGAN, and EPDN have less haze. However, these methods often fail to remove haze from far areas. The dehazed results of DCP, ZID, and EPDN can remove more haze. However, these dehazed results often contain color distortions. In contrast, the proposed method removes haze more completely.

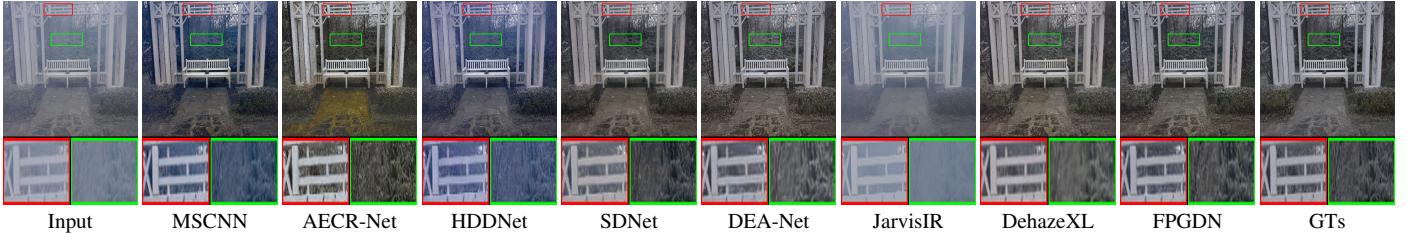


Fig. 10: Visual comparisons of various dehazing methods on the O-HAZE dataset [41]. Critical regions (red/green boxes insets) are magnified to highlight differences in haze removal and detail preservation.

TABLE V: Quantitative comparison on O-Haze.

	DCP	MSCNN	AOD-Net	DCPDN	GPN	HDDNet	SDNet	AEZR-Net	DEA-Net	JarvisIR	DehazeXL	DNMGDT	FPGDN
PSNR	16.59	19.07	16.72	15.62	18.30	19.28	19.38	18.70	16.32	14.31	21.49	19.55	21.54
SSIM	0.74	0.77	0.68	0.62	0.72	0.77	0.77	0.77	0.70	0.65	0.73	0.74	0.78

lishes a new state-of-the-art, surpassing all competitors in both PSNR and SSIM. Visual comparisons are presented in Fig. 10, with critical regions enlarged for detailed analysis. As shown, AEZR-Net exhibits noticeable color distortion, particularly in areas with high contrast and complex textures. JarvisIR [45], HDDNet [46], and MSCNN retain significant residual haze, especially in distant objects and background regions. DehazeXL [42], SDNet [47] and DEA-Net [48] show improved results but introduce artifacts. DehazeXL produces blur in certain areas (highlighted by green boxes), while DEA-Net and SDNet cannot incompletely remove haze in some

regions (indicated by the red box). In contrast, our method generates visually superior results with complete haze removal, preserved structural details, and natural color fidelity. The dehazed images produced by our method closely resemble the ground truth, demonstrating the effectiveness of our approach in handling real-world hazy conditions.

D. Ablation Studies

To prove the effectiveness of the proposed module, a series of variable models are designed. Firstly, the effectiveness of the proposed Edge2ImageNet is demonstrated.

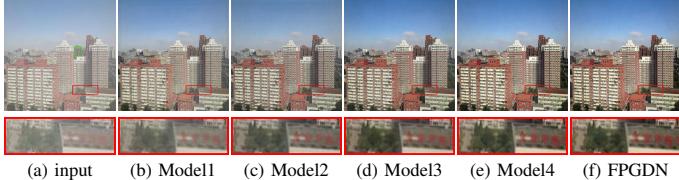


Fig. 11: The visual results of different models. The dehazed results obtained by other models often retain haze or color distortions. The proposed method can remove haze more completely and obtain a more natural dehazing result.

TABLE VI: Quantitative PSNR and SSIM results on the synthetic dataset from the RESIDE dataset.

Metric	Model1	Model2	Model3	Model4	FPGDN
PSNR	23.35	29.49	26.99	30.57	31.76
SSIM	0.86	0.93	0.84	0.95	0.97

Edge2ImageNet is designed to capture long dependencies and provide additional prior information, which is conducive to image dehazing. To this end, two models are designed: Model1, which does not incorporate Edge2ImageNet, and Model2, which includes Edge2ImageNet but without pre-training. Secondly, the effectiveness of GMM and attention blocks is evaluated. Model3 is formed by eliminating the attention blocks from Model2, which reduces the model's representational capacity. Model4 is created by replacing GMM with GM pretrained from our proposed model, resulting in less diversity of the generated images. The dehazing performance of these models is presented in Table VI. The Edge2ImageNet is critical for improving dehazing performance by capturing the long dependency, which helps the model understand the haze distribution. The experimental result demonstrates that the attention mechanism critically exploits the fuzzy prior and local features for dehazing. GMM increases the diversity of the output from Edge2ImageNet and further boost the dehazing performance. The Edge2ImageNet with pre-trained can further enhance the dehazing performance by transferring the knowledge from natural haze-free images to the dehazing problem, which helps restore a natural dehazing result. Some visual results of these models are shown in Fig. 11. The comparative analysis presented in Fig. 11 reveals significant performance disparities among different model configurations. Model1 exhibits residual haze artifacts and noticeable chromatic aberrations in its output, indicating limitations in atmospheric scattering correction. The incorporation of Edge2ImageNet in Model2 demonstrates substantial improvements, effectively mitigating haze contamination while preserving color consistency through enhanced edge-aware feature extraction. The omission of attention blocks leads to a marked deterioration in visual fidelity, underscoring the critical role of spatial attention mechanisms in maintaining structural coherence during the dehazing process. Notably, the integration of GMM demonstrates superior haze removal capability by leveraging statistical priors derived from clean reference images, achieving higher SSIM scores compared to baseline methods. Furthermore, our proposed pre-training

strategy leveraging transfer learning from a pre-trained model yields qualitatively superior results, producing haze-free outputs with enhanced natural appearance and better preservation of fine details.

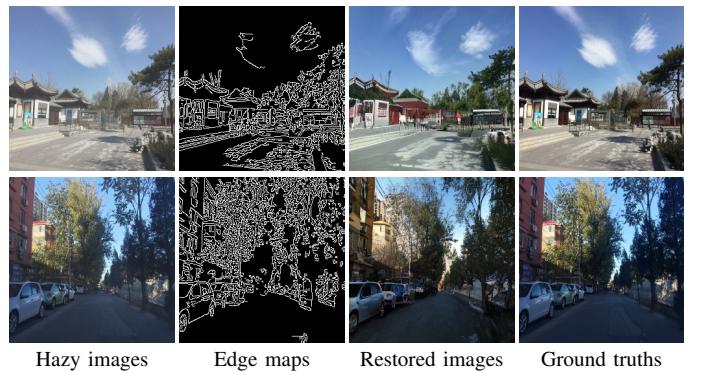


Fig. 12: A visual example of the outputs of the Edge2ImageNet alongside hazy inputs, edge maps, and ground truths (GT). As shown, the Edge2ImageNet restores the overall structures and fine details. Note: Color inconsistencies in F_{rough} are intentionally discarded; only $F_{fuzzyprior}$ is used for final dehazing.

Furthermore, we investigate the function of Edge2ImageNet. As illustrated in Fig. 12, while the generated images F_{rough} preserve global structural consistency with GT, they exhibit color ambiguities (e.g., trees rendered as green vs. brown) and textural deviations. These discrepancies arise because edge maps lack photometric cues, leading to underconstrained color synthesis. Crucially, such pixel-level inaccuracies in F_{rough} could indeed hinder direct haze removal if used naively. To mitigate this, our framework decouples the role of Edge2ImageNet. First, Fuzzy Prior Extraction: Instead of using F_{rough} directly, we leverage the feature-level probabilistic prior $F_{fuzzyprior}$ (Eq. (3)), which encodes a distribution of plausible color/texture hypotheses conditioned on edges (modeled via GMM in Section III-A). This prior captures multi-modal solutions (e.g., trees as green or brown) without committing to incorrect pixel values. Second, Guided Refinement via DecodeNet: $F_{fuzzyprior}$ is fused with local features from the hazy image (Eq. (10)). DecodeNet then resolves ambiguities by integrating local cues (texture, contrast) from the hazy input, suppressing color drifts while preserving structural coherence. Therefore, Edge2ImageNet's primary contribution is providing this uncertainty-aware representation ($F_{fuzzyprior}$) (Section III-A). F_{rough} serves only as an intermediate visualization of structural accuracy; its color inaccuracies are intentionally discarded. Our framework treats $F_{fuzzyprior}$ as a generative hypothesis space, which DecodeNet refines by applying constraints from the hazy image. This decoupling prevents propagating F_{rough} 's pixel-level artifacts to the final output. The efficacy of this approach is substantiated by significant quantitative gains (Table VI: +6.14 dB PSNR vs. without Edge2ImageNet), confirming its role in enhancing dehazing robustness across diverse real-world scenarios (e.g., varying



Shengdong Zhang received his Ph.D. degree from Wuhan University, Wuhan, China, in 2020. He received the B.Sc. degree in computer science and technology from Shanxi University, Shanxi, China, in 2009, and the M.Sc. degree in mechanics from Peking University, Beijing, China, in 2012. His research interests include deep learning, computer vision, and image processing.



Feng Huang received the Ph.D. at The School of Computer Science and Technology, Hangzhou Dianzi University, Hangzhou, China, in 2025. He received the BEng degree in Electronic Information Engineering from Zhejiang University of Science and Technology, Hangzhou, China, in 2013, and the MEng degree in Circuits and Systems from Hangzhou Dianzi University in 2016. He is currently an associated professor at Zhejiang Province Taizhou Technician College, Taizhou, Zhejiang, China. His research interests include Image Segmentation, Image Dehazing, Video Action Segmentation, and Action Quality Assessment.



Xiaoqin Zhang received the B.Sc. degree in electronic information science and technology from Central South University, China, in 2005 and Ph.D. degree in pattern recognition and intelligent system from the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China, in 2010. He is currently a professor in Wenzhou University, China. His research interests are in pattern recognition, computer vision and machine learning. He has published more than 100 papers in international and national journals, and international conferences, including IEEE T-PAMI, IJCV, IEEE T-IP, IEEE T-NNLS, IEEE T-C, ICCV, CVPR, NIPS, IJCAI, AAAI, ACM MM and among others.



level vision problems.

Wenqi Ren received the Ph.D. degree from Tianjin University, Tianjin, China, in 2017. From 2015 to 2016, he was supported by the China Scholarship Council and working with Prof. Ming-Husan Yang as a joint-training Ph.D. Student with the Electrical Engineering and Computer Science Department, University of California at Merced. He is currently an Associate Professor with the School of Cyber Science and Technology, Shenzhen Campus of Sun Yat-sen University, Shenzhen, China. His research interests include image processing and related high-



Li Zhao received the B.Sc. degree in automation in 2005 and MEng degree in control theory and control engineering in 2008 from Central South University, China. She is currently an assistant researcher in Wenzhou University. Her research interests are in pattern recognition, computer vision, and machine learning.



En Fan received the B.Sc. degree in electronic information science and technology from Hubei Engineering University, Xiaogan, China, the M.Sc. degree in signal and information processing from Nanchang Hangkong University, Nanchang, China, and the Ph.D. degree in signal and information processing from Xidian University, Xi'an, China, in 2015. He is currently a Lecturer with Shaoxing University, Shaoxing, China. His current research interests include sensor fusion, navigation, and intelligence information processing.