

Automatic Model-Based Dataset Generation for High-Level Vision Tasks of Autonomous Driving in Haze Weather

Meiqi Wang , Tianqi Su , Siyi Chen , Wenhan Yang , *Member, IEEE*, Jiaying Liu , *Senior Member, IEEE*, and Zhongfeng Wang , *Fellow, IEEE*

Abstract—Improving the performance of high-level computer vision tasks in adverse weather (e.g., haze) is highly critical for autonomous driving safety. However, collecting and annotating training sets for various high-level tasks in haze weather are expensive and time-consuming. To address this issue, we propose a novel haze generation model called HazeGEN by coupling the variational autoencoder and the generative adversarial network to automatically generate annotated datasets. The proposed HazeGEN leverages a shared latent space assumption based on an optimized encoder–decoder architecture, which guarantees high fidelity in the cross-domain image translations. To ensure that the generated image can truly facilitate high-level vision task performance, a semisupervised learning strategy is developed for HazeGEN to efficiently learn the useful knowledge from both the real-world images (with unsupervised losses) and the synthetic images generated following the atmosphere scattering model (with supervised losses). Extensive experiments and ablation studies demonstrate that training the model with our generated haze dataset greatly improves accuracy in high-level tasks such as semantic segmentation and object detection. Furthermore, one important but under-exploited issue is investigated to find out whether the developed dataset can be a good substitute for the real ones. Results show that the generated dataset has the most similar performance to the real-world collected haze dataset on multiple challenging industrial scenarios compared with prior works.

Manuscript received 11 July 2022; revised 14 October 2022 and 5 November 2022; accepted 17 November 2022. Date of publication 28 November 2022; date of current version 13 July 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 62174084 and Grant 62104097, and in part by the High-Level Personnel Project of Jiangsu Province under Grant JSSCBS20210034. Paper no. TII-22-3037. (Meiqi Wang and Tianqi Su contributed equally to this work.) (Corresponding authors: Jiaying Liu; Zhongfeng Wang.)

Meiqi Wang, Tianqi Su, Siyi Chen, and Zhongfeng Wang are with the School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China (e-mail: mqwang@smail.nju.edu.cn; sutianqi@smail.nju.edu.cn; sychen@smail.nju.edu.cn; zfwang@nju.edu.cn).

Wenhan Yang is with the Peng Cheng Laboratory, Shenzhen 518066, China (e-mail: yangwenhan@pku.edu.cn).

Jiaying Liu is with the Wangxuan Institute of Computer Technology, Peking University, Beijing 100871, China (e-mail: liujiaying@pku.edu.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TII.2022.3224958>.

Digital Object Identifier 10.1109/TII.2022.3224958

Index Terms—Autonomous driving, haze, object detection, semantic segmentation, semisupervised learning, shared latent space, variational autoencoder–generative adversarial network (VAE–GAN).

I. INTRODUCTION

DEEP neural networks (DNNs) have become successful in many computer vision tasks [1], [2], e.g., image recognition, object detection, semantic segmentation, etc. However, when facing adverse weather, such as haze/fog, the captured images lose details and are affected by color casting, which severely degrades the DNN models' performance on high-level tasks and influences the reliability and efficiency of vision-based systems such as autonomous vehicles (see Fig. 1). Thus, improving the robustness and efficiency of the DNN model in adverse weather attracts significant research attention [3].

One of the solutions to enhance the image quality is to use the preprocessing image enhancement module, such as dehazing module [4]. However, this solution cannot satisfy practical needs due to the lack of consideration of the resource constraint of industrial autopilot systems, which severely slows down the processing process and violates the need for the real-time applications. Furthermore, existing image preprocessing methods might not necessarily be effective for high-level tasks in real-world, because the commonly adopted evaluation metrics (e.g., peak signal-to-noise ratio and structural similarity index) mainly target the signal fidelity between the dehazed images and the corresponding clean ones, which might fail to stand for the performance on high-level tasks [5].

Using a particular annotated dataset to train the model can be a more practicable way for high-level tasks. To obtain the required dataset, some efforts are made to simulate a foggy dataset. Sakaridis et al. [5] and Gao et al. [6] used the atmosphere scattering model [7] to generate a hazy dataset as defined as follows:

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

The hazy image $I(x)$ is constructed from haze-free image $J(x)$, transmission map $t(x)$, and global atmospheric light $A(x)$. However, the synthetic $t(x)$ and $A(x)$ might not be reasonable at the given scenes, which leads to artifacts, such as color casting and noises. Gong et al. [8] tried to learn a direct mapping across

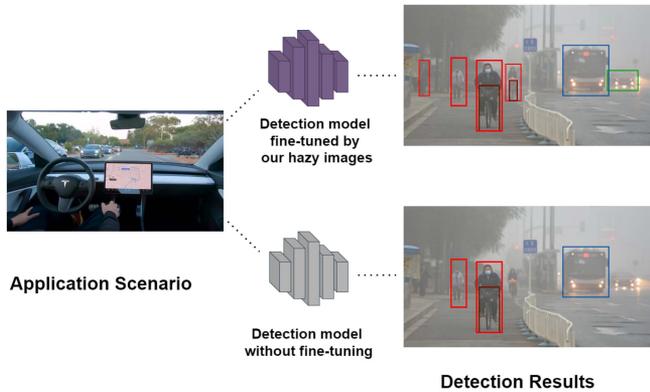


Fig. 1. Application scenario of the proposed model: the detection model without fine-tuning on our generated haze dataset fails to detect all targets, which may cause severe security issues in autonomous driving systems.

different domains by using supervised learning with the existing synthetic dataset. However, without the adaption to real hazy images, the output synthetic hazy image has gaps with the real one, leading to high-level tasks' performance degradation in real scenarios.

In summary, the existing haze dataset generation approaches still have two neglected issues: 1) The majority of prior works rely more on synthetic data and physics model than real haze training images, which makes it challenging to adapt to complex and changing scenarios in real-world industrial environments. 2) Conventional dataset generation approaches might introduce artifacts such as color casting and noises at the same time, which fails to preserve image details and degrades high-level task performance in autonomous driving.

To address the issues mentioned above, multiple innovations are combined together to build a deeply optimized image-to-image translation framework that can truly facilitate the high-level tasks performance improvement. Overall, we propose a haze generation model called HazeGEN trained in a semisupervised way. To directly learn from the real-world haze images that have no paired "ground truth" (i.e., the pixel-to-pixel paired clean images), multiple unsupervised-based losses are developed to leverage the knowledge from the real haze distribution. To better preserve the details of the original clean images, several supervised training losses are also leveraged based on the paired training images synthesized following the atmosphere scattering model [see (1)]. To jointly learn from both unpaired and paired images, we enforce the constraint that the two kinds of images share the same latent space so that the images from the clean, real haze, and synthetic haze domains can be mapped into the same latent code and transferred flexibly. Based on this assumption, we introduce the VAE [9] as the image generator, the effectiveness of which is further enhanced by the adversarial training in GAN to avoid the high-level task performance degradation caused by the low-resolution images generated by VAE-only [10]. With all these components, we generate our photo-realistic haze dataset for high-level task training.

In conclusion, our contributions are highlighted as follows.

- 1) One end-to-end haze generation model is developed, which can automatically generates haze dataset that can

be used to truly boost the high-level task performance. It is noted that the laborious dataset creation process for industrial applications can be eliminated in a more practical way in this work by: 1) demonstrating the superiority of the proposed HazeGEN over not only the previous synthetic haze datasets, but also the prior dehazing works and the real-world collected datasets; 2) measuring the similarity between our generated images and the real-world annotated haze images directly on the the high-level tasks in haze weather, rather than at the pixel level or visual level as in prior works.

- 2) We are the first to integrate the shared latent space based VAE-GAN into the semisupervised learning to generate haze datasets. The proposed optimizations at the architecture and algorithm levels enable highly effective learning from both real haze images and prior knowledge for haze modeling with few artifacts, which truly improve the high-level task performance in haze weather.
- 3) The proposed model is extensively evaluated on two most common tasks (object detection and semantic segmentation) in multiple challenging industrial scenarios, e.g., the real-world haze with diverse distributions (including both the light and dense haze) and the haze-like weather (including dark snow dust and heavy snow). Results show that the performance is increased by as high as 10.03% and 5.93% for object detection and semantic segmentation, respectively. The proposed method greatly surpasses the state-of-the-art (SOTA) hazy generation methods in accuracy, and outperforms the SOTA dehazing methods in both speed and accuracy. These results demonstrate its superiority in supporting the real-time and safety requirements of autonomous driving under various weather conditions.

II. RELATED WORK

A. Dataset Generation Methods in Haze Weather

Many works are proposed to synthesize datasets for haze weather. Zoph et al. [11] used a data augmentation method to generate annotated datasets. However, the method ignores the difference between the style transformation and the mapping from the 3-D physical world into the 2-D image domain. Li et al. [12] used the atmosphere scattering model. However, it introduces artificial flaws and color distortion due to the unavoidable measurement error. The defects lead to a degradation of the performance on high-level vision tasks. Li et al. [13] proposed an end-to-end learned model-based simulation method based on the generative adversarial networks (GAN). However, the GAN models cannot estimate the posterior and are unstable to train. Another method called Cycle-GAN [14] is proposed by utilizing unpaired training data in real-haze and clean days. But without a supervised training constraint, this method leads to an undesired conversion style (e.g., the loss of important details) [15].

B. High-Level Vision Tasks in Adverse Weather

High accuracy of high-level vision tasks such as object detection and segmentation are highly important for safe autonomous

driving. However, most existing works focus on improving their performance in clear days and only recent works explore adverse conditions. Qin et al. [16] applied a preprocessing dehazing model. However, they only paid attention to enhancing the image quality rather than the high-level tasks' performance. Vishwanath et al. [17] proposed an unsupervised physical-model-based detection framework in hazy and rainy conditions. However, inevitable measuring error in the physical model makes synthesized images include undesirable artifacts, affecting the high-level tasks' performance. Chen et al. [18] assumed that the adversarial weather conditions result in the domain shift, and they propose a domain adaptive Faster-RCNN to tackle the domain shift problem. However, the synthetic dataset used in the method exhibits a clear difference in appearance with real-world images, which harms object detection performance in the real scene. To conclude, previous attempts are still far from the goal to generating natural hazy images that can benefit high-level vision tasks.

C. Shared Latent Space

The basic assumption behind the image translations between two domains is that some latent factors, such as the high-level embedding features [19], [20], are shared by the source domain and the target domain. Li et al. [21] indicated that a shared manifold for observations can be learnt based on the Gaussian process latent variable model. Furthermore, [22] and [23] used the shared latent space assumption for handwritten pattern conversion and face aging tasks, demonstrating this assumption's generalizability. Inspired by these works, we propose the HazeGEN framework based on the shared latent space assumption, to map images from clean, synthetic, and real haze domains to the same latent space and recover images from this space.

III. METHOD

In this section, the model architecture of the proposed HazeGEN is introduced first, followed by the loss functions for the semisupervised training to show how the information in real haze scenarios is learnt and how the artifacts are suppressed to ensure the high performance of high-level tasks.

A. Model Structure

When approaching the image translation problem from the probabilistic modeling view, the major difficulty is learning a joint distribution from the marginal distributions in two different domains since there exist an infinite number of joint distributions that can lead to the specified marginal distributions [24]. To address this ill-posed issue, we need to make an additional assumption about the joint distribution.

Inspired by the works [21], [22], [23], we make the shared-latent space assumption. In the assumed space, there exists a shared latent code for any given pair of images x_1 and x_2 . This code allows us to recover both images from it. Following the shared-latent space assumption, we build the domain transformations in this work as illustrated in Fig. 2.

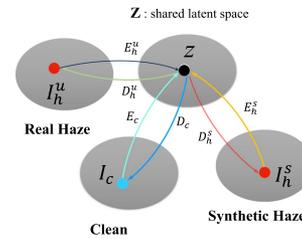


Fig. 2. Shared latent space assumption among the real haze, synthetic haze, and clean image domains.

Specifically, the generator (GEN) of VAE and the discriminator (DIS) from GAN are combined to build the image translation framework as shown in Fig. 3. In each GEN , an encoder–decoder (E – D) architecture is adopted to encode the input image from one domain into the latent space z and decode it into the target domain. The embedding in the shared latent space z can be regarded as a compact, high-level embedding vector of a scene, while the encoded result $E(I)$ can be considered as a particular sample in z . With the shared latent space, $E(I)$ is then reconstructed in a cross-domain or in-domain way with the corresponding decoder D . The reconstructed image $D(E(I))$ is then sent to the discriminator (DIS) to be distinguished whether the generated images can be regarded as real images from the input domain. SubNetwork roles are analyzed to demonstrate how the proposed HazeGEN is different from prior work for the similar task such as [25]. Note that $\{E_h^s, D_h^s, DIS_h^s\}$ and $\{E_h^u, D_h^u, DIS_h^u\}$ are separately illustrated to distinguish the parameters iteratively updated by supervised and unsupervised learning, but they share the same sets of VAE-GAN modules.

In each GEN (see Fig. 4), the encoding part mainly contains two stride-2 convolutions and four residual blocks. We use the stride-convolution layer to down-sample the feature maps by 1/2. The decoding part consists of four residual blocks and two transposed convolutions. We use the transposed-convolution layer to up-sample the features by a factor of 2. A skip connection connects the encoder and decoder in the symmetric layer at different resolutions. The number of output channels are shown in Fig. 4. Each residual block consists of a convolution layer, an instance normalization layer (IN) [26], and a ReLU activation layer. The discriminator (see Fig. 5) is constructed with a convolution layer, an instance normalization layer, and a ReLU. Motivated by Pix2PixHD [27], the output is down-sampled to three scales to calculate the discriminator loss, which ensures the hierarchical feature similarity is reserved in the generated images.

B. Loss Function

The proposed semisupervised learning can be divided into two stages: the supervised training and the unsupervised training, which are performed in a step-by-step way. The whole process is illustrated in Algorithm 1.

1) *Loss in Unsupervised Training Stage*: The unsupervised training stage is supposed to extract real hazy distribution from unpaired real hazy images. In this stage, the real weather image

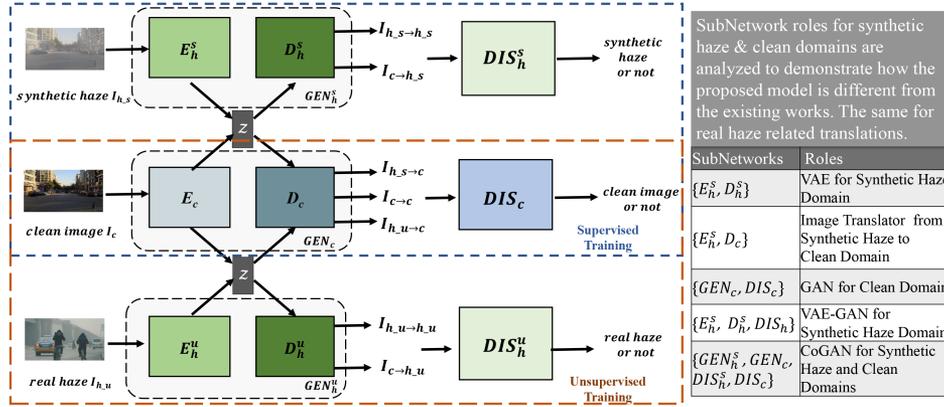


Fig. 3. The overall architecture of the HazeGEN which includes the VAE generator (GEN) and discriminator of GAN (DIS). GEN contains an encoder E and a decoder D . E and D are connected by a code in the shared latent space z . Different connections of E and D are responsible for in-domain transformation (e.g. $I_{c \rightarrow c} = D_c(E_c(I_c))$) or cross-domain transformation (e.g. $I_{c \rightarrow h,s} = D_h^s(E_c(I_c))$). The shared modules for synthetic and real haze are separately illustrated to distinguish the parameters iteratively updated during the supervised and unsupervised stages, respectively.

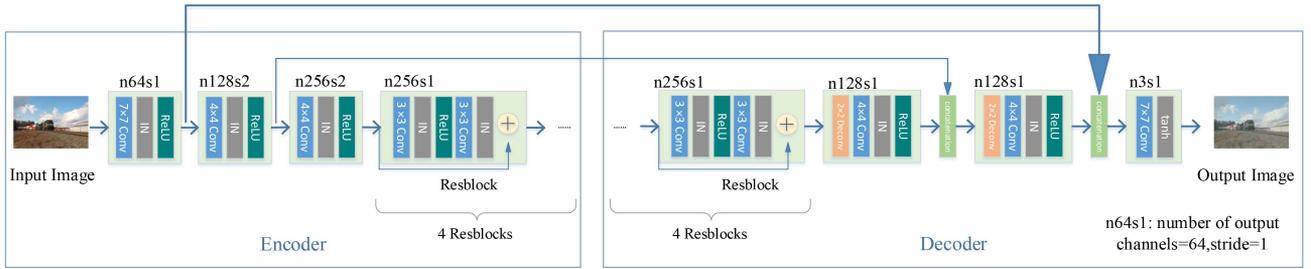


Fig. 4. Overall architecture for generator (GEN).

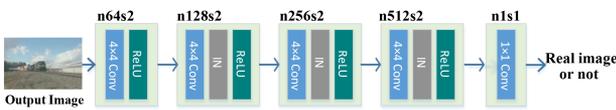


Fig. 5. Overall architecture for discriminator (DIS).

$I_{h,u}$ is encoded into the shared latent space z via an encoder E_h^u , and can be decoded into either the clean domain or the original real haze domain via a different decoders D_c or D_h^u . The generated results are then sent to the discriminator to distinguish whether the generated image is the real data from one specific domain. We use VAE loss L_{VAE} to evaluate the reconstruction likelihood and extract the encoder's distribution representation. GAN loss L_{GAN} is further leveraged to ensure the quality of the outputs through adversarial training. Following the setting of [25], we set the same loss weights for L_{GAN} and L_{VAE} . The combined VAE/GAN loss function in the unsupervised stage can be written as

$$L_{VAE/GAN_{unspv}}(I_{h,u}) = L_{VAE}(I_{h,u}) + L_{GAN \rightarrow h}(I_{h,u}) \quad (2)$$

$$L_{VAE}(I_{h,u}) = D_{KL}(q(E_h^u(I_{h,u})|I_{h,u})||p(z)) + \alpha \|D_h^u(E_h^u(I_{h,u})) - I_{h,u}\|_1 \quad (3)$$

$$L_{GAN \rightarrow h}(I_{h,u}) = \mathbb{E}[\log(DIS_h^u(I_{h,u}))] + \mathbb{E}[\log(1 - DIS_h^u(D_h^u(E_h^u(I_{h,u}))))]. \quad (4)$$

The VAE optimizes the encoder by imposing a prior over the latent space on $p(z)$, while making the decoded results $D_h^u(E_h^u(I_{h,u}))$ as close as possible to the original input $I_{h,u}$. $E_h^u(I_{h,u})$ represents the latent space information z encoded by E_h^u . The Kullback–Leibler divergence D_{KL} evaluates the distance between the encoded distribution $q(E_h^u(I_{h,u})|I_{h,u})$ and the distribution prior of zero Gaussian $p(z)$. The last term in L_{VAE} stands for in-domain L_1 loss to reconstruct the image with high fidelity.

To further improve the high-level tasks' performance, we adopt the total variation (TV) loss in the unsupervised stage to encourage spatial smoothness, which can enhance visibility by removing unwanted details and preserving important details such as edges:

$$L_t = \|\nabla_h D_h^u(E_h^u(I_{h,u}))\|_1 + \|\nabla_v D_h^u(E_h^u(I_{h,u}))\|_1, \quad (5)$$

where ∇_h and ∇_v represent the horizontal and vertical differential operation matrices, respectively.

The total loss in unsupervised stage can be written as

$$L_{\text{unspv}} = L_{\text{VAE/GAN}}(I_{h_{-u}}) + \eta L_t \quad (6)$$

where η is a hyperparameter that determines the contribution of the TV loss L_t .

2) **Loss in Supervised Training Stage:** The main issue in the supervised stage is to learn the atmosphere scattering model in the paired training set, keep consistency between the inputs and avoid image ill-rendering caused by unsupervised training. To guide the transformation and learn valuable information, we employ VAE/GAN loss in the supervised stage

$$L_{\text{VAE/GAN}}(I_{h_{-s}}, I_c) = L_{\text{VAE}}(I_{h_{-s}}) + L_{\text{GAN} \rightarrow h}(I_{h_{-s}}) + L_{\text{VAE}}(I_c) + L_{\text{GAN} \rightarrow c}(I_c) \quad (7)$$

where $I_{h_{-s}}$ and I_c stand for synthetic hazy image and its corresponding clean image, respectively. Cross-domain L_1 loss is adopted to ensure that the cross-domain transformed image $D(E(I))$ is close to the ground truth (i.e., the corresponding pixel-to-pixel paired haze/clean images before transformation)

$$L_1 = \|D_h^s(E_c(I_c)) - I_{h_{-s}}\|_1 + \|D_c(E_h^s(I_{h_{-s}})) - I_c\|_1. \quad (8)$$

The total loss function in supervised-learning stage is

$$L_{\text{spv}} = L_{\text{VAE/GAN}}(I_{h_{-s}}, I_c) + \beta L_1 \quad (9)$$

where β reflects the importance of L_1 .

3) **Shared Loss in Both Training Stages:** Since shared latent space assumption implies the cycle-consistency constraint [28], we employ cycle consistency loss in both supervised and unsupervised training stages to ensure that the reconstructed image generated by the in-domain and cross-domain encoder-decoder is the same as the original image

$$L_{\text{cyc}} = \|D_c(E_h(D_h(E_c(I_c)))) - I_c\|_1 + \|D_h(E_c(D_c(E_h(I_h)))) - I_h\|_1 \quad (10)$$

where E_h/D_h denotes the encoder/decoder for either synthetic or real haze domain.

Note that the loss mentioned above mainly deals with images' similarity at the pixel level. The per-pixel losses do not capture perceptual differences between output and ground-truth images. Therefore, we apply perceptual loss which is based on a pretrained VGG-16 to measure the perceptual similarity in feature space

$$L_p = \|\phi(D_c(E_h(I_h))) - \phi(I_c)\|_2 + \|\phi(D_h(E_c(I_c))) - \phi(I_h)\|_2 \quad (11)$$

where $\phi(\cdot)$ represents the feature maps obtained by the relu5_3 layers within the VGG-16 network. The total loss function can be expressed as

$$L = L_{\text{unspv}} + L_{\text{spv}} + \delta L_{\text{cyc}} + \gamma L_p. \quad (12)$$

The hyperparameters δ and γ control the importance of the corresponding L_{cyc} and L_p , respectively.

Algorithm 1: The Process of the Semi-Supervised Learning for HazeGEN.

Input: The initial parameters of the GEN_c , GEN_h^s , GEN_h^u , clean images I_c , synthetic hazy images $I_{h_{-s}}$ for supervised learning, and real hazy images $I_{h_{-u}}$ for unsupervised learning.

Output: Updated GEN_c , GEN_h^s .

for $iter=0$; $iter < iter_{max}$; $iter++$ **do**

if $iter \% 2 = 0$ **then**

 Randomly sample paired images from the image dataset $I_c, I_{h_{-s}}$;

 Get the generated hazy images $I_{c \rightarrow h_{-s}}$ and clean images $I_{h_{-s} \rightarrow c}$ with GEN_h^s and GEN_c , respectively;

 Compute the supervised losses L_{spv} , and the shared losses L_p, L_{cyc} with $I_{c \rightarrow h_{-s}}, I_{h_{-s} \rightarrow c}, I_c, I_{h_{-s}}$;

 Update the GEN_h^s and GEN_c with the above losses;

else

 Randomly sample unpaired images from the image dataset $I_c, I_{h_{-u}}$;

 Get the generated hazy images $I_{c \rightarrow h_{-u}}$ and clean images $I_{h_{-u} \rightarrow c}$ with GEN_h^u and GEN_c , respectively;

 Compute the unsupervised losses L_{unspv} , and the shared losses L_p, L_{cyc} with $I_{c \rightarrow h_{-u}}, I_{h_{-u} \rightarrow c}, I_{h_{-u}}, I_c$;

 Update the GEN_h^u and GEN_c with the above losses;

IV. APPLICATION STUDY

In this section, we illustrate the effectiveness of our HazeGEN in two of the most common applications for autonomous driving, i.e., object detection and semantic segmentation. The datasets are generated based on the existing annotated clear day dataset. They are then used for training different backbones to improve these models' performance in haze weather. Two widely used metrics, mean Intersection-Over-Union (mIoU) and mean Average Precision (mAP), are used to measure the performance. mIoU is the average intersection over union across all classes, which is the most commonly used metric for segmentation accuracy. Its formulation is

$$\text{mIoU} = \frac{1}{n} \sum_{i=1}^n \frac{\text{TP}(i)}{\text{TP}(i) + \text{FP}(i) + \text{FN}(i)} \quad (13)$$

where n is the total number of categories. TP , FP , and FN represent true-positive, false-positive, and false-negative predictions, respectively. mAP gives average accuracy of predicted object locations across all object predictions. Its definition is

$$\text{mAP} = \frac{1}{n} \sum_{i=1}^n \text{AP}_k(i) \quad (14)$$

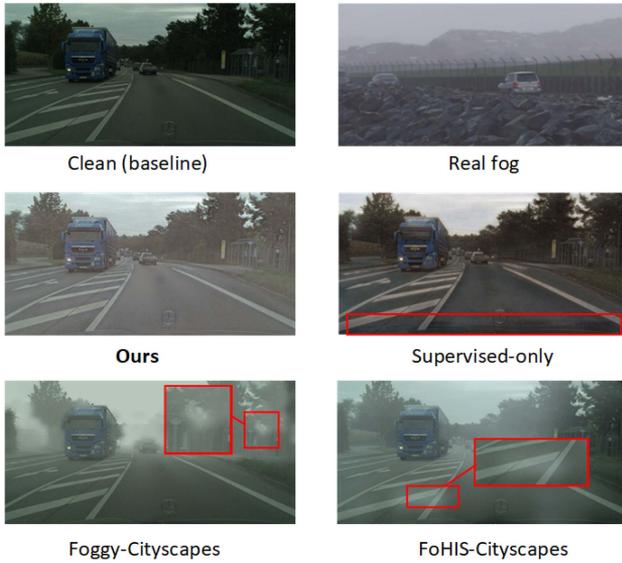


Fig. 6. Visual comparisons of different generated datasets. The defects of other methods are labeled by red boxes. The “Supervised-only” generated haze can be barely seen and ill-rendering problems exist on the car logo. Haze synthesized by other methods of Foggy/FoHIS-Cityscapes is unnaturally concentrated, and their picture details are corrupted, which lead to lower detection/segmentation accuracy. In our method, the smoother distribution and better preserved details benefit the high-level vision tasks.

where n stands for the number of categories, and AP_i is the average precision (AP) of the class i . Higher mIoU or mAP denotes better accuracy.

A. Implementations and Results for Dataset Generation

We train the proposed data generation network HazeGEN in Pytorch framework and use ADAM optimizer for training. The initial learning rate is 0.0001 and decreases by $0.75 \times$ every 30 000 iterations. The model is trained for 180 000 iterations until it gets convergence. The weight decay is set as 0.00001. The semisupervised training is performed by randomly sampling 4000 the labeled data from synthetic outdoor training set (the synthetic haze data used for supervised training stage) and 4000 unlabeled images from realistic hazy images in RESIDE [29] dataset (real haze data used for unsupervised training stage). The input images are resized to the size of 256×256 . The data augmentation strategies, including horizontal and vertical flipping, and random cropping are applied. We follow [28] to preset the hyperparameters and then adjust them to get the best experimental results. After the adjustment, the loss weights are finally set as $\alpha = 1$, $\beta = 1000$, $\delta = 10$, $\gamma = 0.5$, and $\eta = 0.5$. The network is trained on the NVIDIA V100 GPU.

The visual results are shown in Fig. 6. It is worth mentioning that we observe that when more depth information is adopted (e.g., in Foggy-Cityscapes), the unnaturally concentrated haze distribution and corrupted details occur, which lead to lower detection/segmentation accuracy as will be demonstrated in the following experiments. In our generated images, by contrast, there are smoother fog distribution and better-preserved details after adding haze, so that they can truly facilitate high-level

Algorithm 2: Details of the Semantic Segmentation.

Train

Input:

Synthesized Cityscapes I_{train} and its annotation J_{train} ; Pre-trained MobileNetV3;

Output:

Fine-tuned MobileNetV3;

Initialization:

Fine-tuned epoch = 5; Optimizer = SGD;
Learning rate = 10^{-6} ; Momentum = 0.9;

Training:

```
for  $i=0; i < \text{Fine-tuned epoch}; i++$  do
     $J'_{train} = \text{MobileNetV3}(I_{train});$ 
    optimize MobileNetV3 with loss;
```

Test

Input:

Real fog from Foggy Driving I_{test_d} ;
Synthesized fog from Foggy Cityscapes I_{test_c} ;
Fine-tuned MobileNetV3;

Output:

Segmentation results J_{test_d} and J_{test_c} ;

Test:

```
 $J_{test_d} = \text{MobileNetV3}(I_{test_d})$ 
 $J_{test_c} = \text{MobileNetV3}(I_{test_c})$ 
```

task performance improvement. The reason is that we propose a semisupervised training method that can learn from both real and synthetic haze images in an end-to-end way. As a result, instead of directly leveraging the depth information of the atmosphere scattering model, the optimal implicit depth prior is imposed by the learning algorithm automatically from the perceptual of benefiting machine vision algorithms.

B. High-Level Task Performance on Real-World Images

To validate whether the proposed biGM can really meet the practical needs, we test semantic segmentation and object detection performance on images collected in real scenarios.

1) *Segmentation*: Overall, the pretrained segmentation model is finetuned with our generated haze dataset and previous synthetic haze dataset, and the comparisons are made to evaluate the finetuned models’ performance in real foggy weather. The details of the semantic segmentation experiments are shown in “Algorithm 2.”

The segmentation model DeepLabv3+[32] is constructed by using MobileNetV3 as the backbone. We adopt annotated dataset Cityscapes [33] as the basic clean image dataset to generate the corresponding foggy training set. Three representative data generation methods, atmosphere-scattering-model based FoHIS [34], depth-map-based Foggy-Cityscapes [5], and the more recent domain-adaptation-based TUNIT [30] are selected in comparisons. Besides, we make comparison with the simpler data augmentation methods of contrast reduction and the lately popular method of contrastive learning based segmentation method [31]. We test the fine-tuned MobileNetV3 model with Foggy Driving [5] dataset, which consists of a fine-annotation

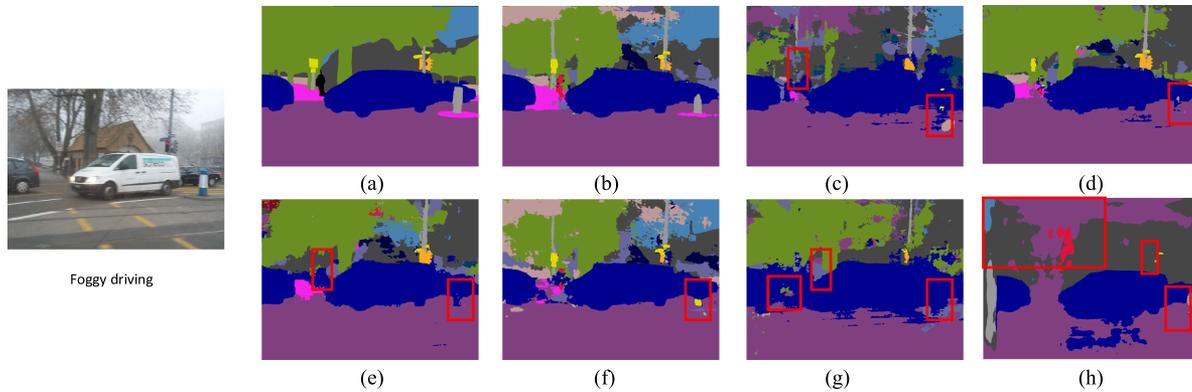


Fig. 7. Visual comparisons of the segmentation. Our segmentation result is closest to the ground truth with all of the critical instances for automobile driving being successfully segmented. The predicting errors in the other methods are labeled by red boxes, among which the fire hydrant (the grey area) cannot be correctly predicted. Other works also have difficulty in distinguishing the pedestrian on the sidewalk (the red area).

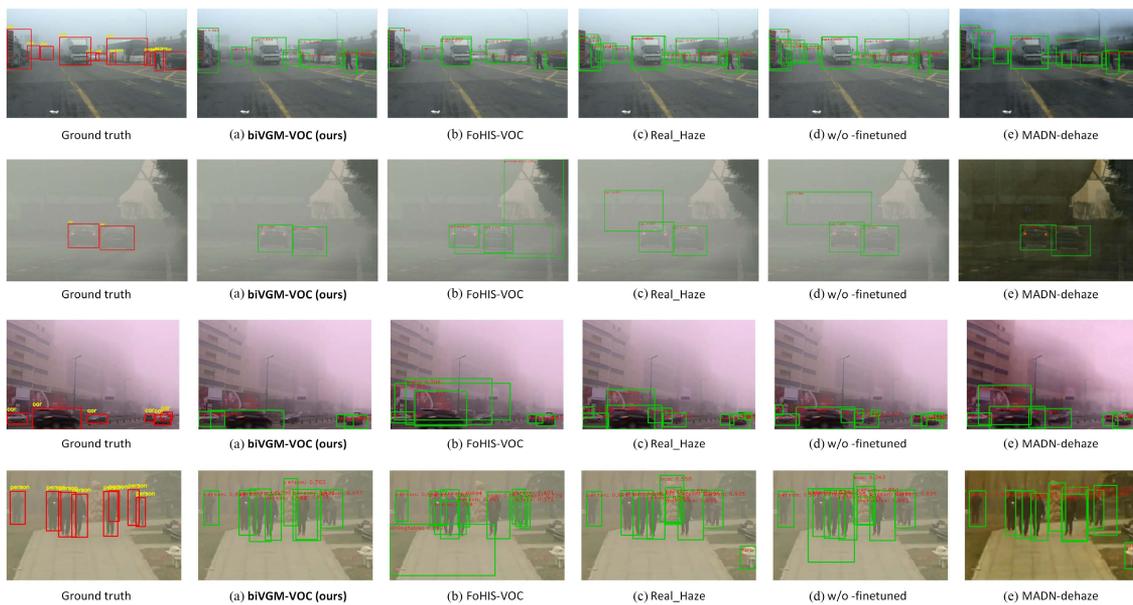


Fig. 8. Visual comparisons for the object detection. We finetune the Faster-RCNN with different hazy datasets. It can be observed that the detection model finetuned by our hazy dataset is the closest to the detection ground truth under both light and dense haze distributions. Models trained by other hazy dataset make wrong predictions in real hazy weather.

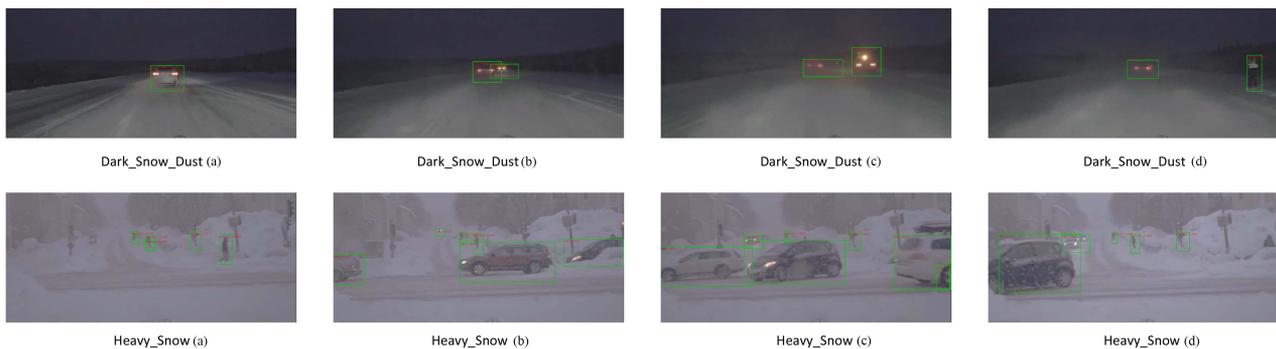


Fig. 9. Our detection results on real-world haze-like images collected from the real autonomous driving system [35] to verify the generation capacity of the proposed model. We evaluate our detection system under different light and weather conditions.

TABLE I
QUANTITATIVE EVALUATION RESULTS OF THE SEGMENTATION RESULTS ON REAL HAZE DATASET FOGGY DRIVING

Category	Methods	Foggy Driving Fine (mAP/ mIoU)	Time Cost(s)	Foggy Driving Coarse (mAP/ mIoU)	Time Cost(s)
Haze generation & Retrain	Clean (baseline1)	50.63/38.24		48.31/34.07	
	Contrast Reduction (baseline2)	48.23/35.65		49.23/32.61	
	Ours	57.81/43.56		56.72/41.08	
	Supervised-only	33.55/22.85	0.1205	46.39/25.67	0.1021
	FoHIS-Cityscapes	47.78/34.23		48.52/32.61	
	Foggy-Cityscapes	52.89/38.72		52.49/34.58	
Dehaze & Segmentation pipeline	TUNIT-Cityscapes [30]	52.03/14.62		48.66/12.24	
	FFA-dehaze	50.75/38.28	2.9907	48.17/33.75	1.8270
	MADN-dehaze [2]	36.56/25.14	2.0131	46.40/28.83	1.4341
Contrastive learning	Contrastive-segmentation [31]	46.98/22.24	/	38.30/24.29	/

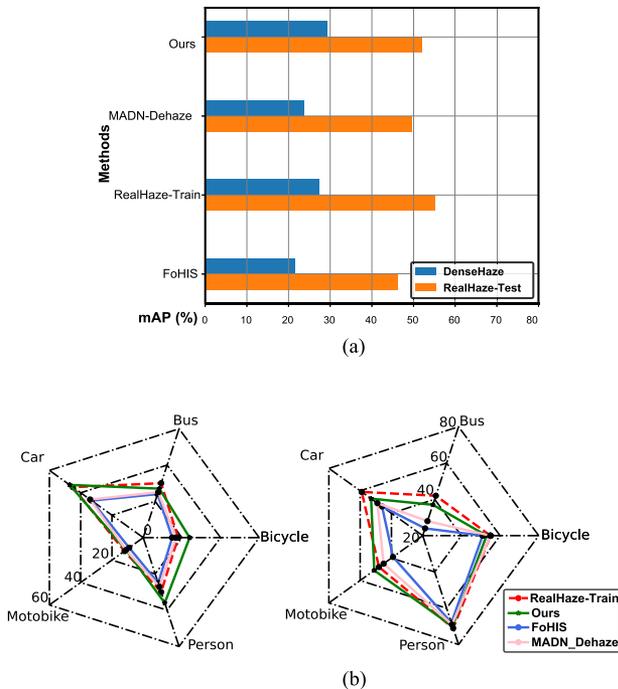


Fig. 10. Our generated dataset achieves the highest mAP compared to other works (even higher than RealHaze-Train on DenseHaze) as shown in (a), and has the closest accuracy compared to the RealHaze-Train (b). (a) Accuracy tested on DenseHaze and RealHaze-Test. The methods include training datasets or dehaze methods. (b) Per-Class Accuracy on dense synthetic haze (left) and real haze (right).

set and coarse-annotation set for the evaluation of segmentation in the real-world fog circumstance. The quantitative results are listed in Table I and the visualization of segmentation is shown in Fig. 7. It is shown in Table I that our haze dataset generation method surpasses other models in both mAP and mIoU by large margins. Also, our segmentation result in Fig. 7 is the closest to the ground truth with all of the critical instances for automobile driving successfully predicted. It should be noted that the segmentation model trained on the atmosphere-scattering-model based images FoHIS-cityscapes has even worse performance than the clean dataset and the dataset augmented with the simple contrast reduction method due to the artificial defects. These results verify that our generated dataset can better improve the high-level vision task performance.

The main reason for our work’s superiority is that the previous synthetic datasets Foggy-Cityscapes and FoHIS-Cityscapes, and

the images produced by “Supervised-only” learning are generated only by using atmosphere scattering model, which ignores the real hazy distribution. As for contrastive learning, it pays more attention to the high-level features instead of the pixel-level details, while the latter are necessary to build a hazy dataset that can well facilitate the segmentation task. By contrast, our method can better simulate the details and distribution of real-world hazy scenes owing to the semisupervised training strategy, which is effective in segmentation model finetuning.

Considering that dehazing is also one common way to preprocess hazy images for higher high-level task performance, we also conduct experiments with the state-of-the-art dehazing methods, MADN-dehaze [2] and FFA-dehaze [16]. The fog is removed via dehazing module before being put into the MobileNetV3. According to the Table I, our method outperforms the dehazing segmentation pipeline in both segmentation accuracy and processing speed since our method does not need the time-consuming preprocessing involved in dehazing method. This result proves that our method is more appropriate to be applied to real-time tasks for autonomous driving.

2) *Details of the Object Detection*: To validate the generalization capacity of the proposed work, we further evaluate the HazeGEN generated dataset on the object detection model.

Training and testing details are shown in Algorithm 3. We choose ResNet101-based Faster-RCNN as the detection model. FoHIS and HazeGEN methods are applied to synthesize HazeGEN-VOC and FoHIS-VOC foggy datasets based on VOC2007 dataset [36]. By using these datasets, we fine-tuned the pretrained Faster-RCNN to establish a stable CNN-based detector for haze weather. We evaluated the model’s detection accuracy on Real-world Task-driven Testing Set (RTTS) dataset, which has 4322 natural hazy images with five annotated object classes.

The experiment results are shown in Table II, Figs. 8 and 9, which prove that our method achieves satisfactory detection results in mAP and outperforms other dataset generation methods and the previous dehazing-detection methods. Fig. 8 shows that the models trained by other hazy datasets make too many wrong predictions, while ours makes satisfying predictions that are the closest to the ground truth for different haze distributions (under the same settings). Fig. 9 shows the promising detection results on images collected from real-world autonomous driving system [35] for another two haze-like conditions, dark snow dust and heavy snow. It can be observed that different light and weather conditions have tiny affect on our model’s detection

TABLE II
THE OBJECTION DETECTION RESULTS (MEASURED BY AVERAGE PRECISION (AP%)) ON NATURAL HAZE RTTS DATASET

Category	Methods	Bicycle	Bus	Car	Motorbike	Person	Average
Haze generation and Retrain	HazeGEN-VOC (ours)	54.19	31.51	52.56	51.66	70.37	52.05
	FoHIS-VOC	51.58	24.53	46.19	39.93	68.79	46.12
Dehaze and Detection pipeline	FFA-dehaze	54.33	27.52	49.56	45.25	69.38	49.21
	MADN-Dehaze	55.02	28.07	49.01	45.70	69.97	49.55

Algorithm 3: Details of the Semantic Segmentation.

Train

Input:

Synthesized VOC2007 I_{train} and its annotation J_{train} ; Pre-trained Faster-RCNN;

Output:

Fine-tuned Faster-RCNN;

Initialization:

Fine-tuned epoch = 5; Optimizer = SGD;
Learning rate = 10^{-3} ; Monument = 0.9;

Training:

for $i=0; i < \text{Fine-tuned epoch}; i++$ **do**
 $J'_{train} = \text{Faster-RCNN}(I_{train});$
 optimize Faster-RCNN with loss;

Test

Input:

Real-world Task-driven Testing Set I_{test} ;
Fine-tuned Faster-RCNN;

Output:

Object detection results J_{test} ;

Test:

$J_{test} = \text{Faster-RCNN}(I_{test})$

accuracy, which verifies the robustness of the proposed HazeGEN against domain shift. These results are consistent with those of segmentation, which demonstrates that the proposed innovations, the semisupervised learning and the shared latent space assumption for the VAE-GAN hybrid model, play an important part in improving high-level task performance.

C. Performance Compared With Real-World Collected Dataset for Challenging Industrial Scenarios

In this section, we evaluate different models' detection mAP and per-class AP on two challenging industrial scenarios for comparison with the results of real-world collected haze dataset to validate whether our synthetic hazy dataset can be a substitute for the real-world dataset. These two scenarios are the real-world haze images that have diverse distributions (RealHaze-Test) from RTTS, and the dense synthetic haze images (DenseHaze) from Foggy-Cityscape.

In Fig. 10(a), it is obvious that the model trained on our dataset surpasses other models and has the closest accuracy with RealHaze-Train. Specifically, for the detection mAP on RealHaze-Test, ours surpasses its counterparts FoHIS and MADN-Dehaze by a performance increase as high as 10%. For dense haze testing, the model trained on our generated haze performs even better than the RealHaze-Train dataset. These

TABLE III
ABLATION STUDY FOR LOSS FUNCTION ON REAL HAZE SCENARIOS

Methods	Foggy Driving Fine (mAP/ mIoU)	Foggy Driving Coarse (mAP/ mIoU)
Ours	57.81/43.56	56.72/41.08
w/o L_t	56.18/41.04	51.50/33.64
w/o L_1	52.56/38.49	51.42/32.77
w/o L_p	54.81/39.75	52.34/33.11
w/o L_{cyc}	56.09/40.71	54.13/33.91
w/o L_{GAN}	49.39/34.72	52.80/32.02

results demonstrate that our generated dataset can be used to substitute RealHaze dataset in complex industrial scenarios with various haze densities.

To step further, from the perclass accuracy in Fig. 10(b), we can find that the model trained on our generated dataset has the closest accuracy to the model trained on the RealHaze-Train. By contrast, other methods have a large accuracy degradation compared with the model trained with real haze images. This demonstrates that our method can better simulate the characteristics of real haze, which means that the realistic HazeGEN-based produced images can be a good substitute for real-world annotated dataset, which greatly alleviates the expensive and time-consuming effort in collecting and annotating dataset for high-level tasks in adverse weather.

D. Ablation Study

The ablation studies are performed to analyze how proposed training losses and model contribute to the improvement in high-level task performance in real scenarios. We remove the important losses or model components and evaluate their influences on the segmentation task as shown in Table III. It can be seen that the proposed/applied semisupervised losses and the important components do take effect in improving the final performance. Specifically, L_t for preserving important details in unsupervised training and L_p to measure the perceptual differences obviously improve the final performance. The advantages of the shared latent space assumption (evaluated by removing the cycle consistency loss L_{cpc} and L_1 loss implied by this assumption) and the coupling of VAE and GAN (evaluated by removing the GAN loss L_{GAN}) are also demonstrated clearly.

V. CONCLUSION

In this article, we proposed an end-to-end dataset generation method to automatically generate training datasets for high-level tasks in haze weather. We creatively combined the semisupervised training strategy, shared latent space, and coupled VAE-GAN to efficiently learn from both real-world haze

distribution and atmosphere scattering model. Experiment results demonstrated significant performance improvement in various haze scenarios for the common high-level vision tasks of autonomous driving, i.e., the object detection and semantic segmentation. Compared with other methods, HazeGEN-based generated dataset can be a much better substitute for the real-world collected dataset, which demonstrates its great application potential in improving the safety of real autonomous driving systems.

REFERENCES

- [1] Y. Han, W.-G. Han, and K. Zeng, "Improved multi-domain convolutional neural network for intelligent tracking of marangoni effect process sequence images," *IEEE Trans. Ind. Informat.*, vol. 18, no. 5, pp. 3397–3405, May 2022.
- [2] T. Jia, J. Li, L. Zhuo, and G. Li, "Effective meta-attention dehazing networks for vision-based outdoor industrial systems," *IEEE Trans. Ind. Informat.*, vol. 18, no. 3, pp. 1511–1520, Mar. 2022.
- [3] M. J. Mirza, M. Masana, H. Possegger, and H. Bischof, "An efficient domain-incremental learning approach to drive in all weather conditions," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2022, pp. 3001–3011.
- [4] Y. Gao, Q. Li, and J. Li, "Single image dehazing via a dual-fusion method," *Image Vis. Comput.*, vol. 94, 2020, Art. no. 103868.
- [5] C. Sakaridis, D. Dai, and L. Van Gool, "Semantic foggy scene understanding with synthetic data," *Int. J. Comput. Vis.*, vol. 126, no. 9, pp. 973–992, 2018.
- [6] R. Gao, Q. Guo, F. Juefei-Xu, H. Yu, and W. Feng, "AdvHaze: Adversarial haze attack," 2021, *arXiv:2104.13673*.
- [7] G. S. Narasimhan and S. K. Nayar, "Vision and the atmosphere," *Int. J. Comput. Vis.*, vol. 48, no. 3, pp. 233–254, 2002.
- [8] R. Gong, D. Dai, Y. Chen, W. Li, D. P. Paudel, and L. Van Gool, "Analogical image translation for fog generation," in *Proc. AAAI Conf. Artif. Intell.*, 2021, vol. 35, pp. 1433–1441.
- [9] C. Doersch, "Tutorial on variational autoencoders," *Stat.*, vol. 1050, p. 13, 2016.
- [10] A. Razavi, A. Van den Oord, and O. Vinyals, "Generating diverse high-fidelity images with vq-vae-2," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 14837–14847.
- [11] B. Zoph et al., "Learning data augmentation strategies for object detection," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 566–583.
- [12] J. Li, L. Zhuo, H. Zhang, G. Li, and N. Xiong, "Effective data-driven technology for efficient vision-based outdoor industrial systems," *IEEE Trans. Ind. Informat.*, vol. 16, no. 7, pp. 4344–4354, Jul. 2020.
- [13] X. Li, K. Kou, and B. Zhao, "Weather gan: Multi-domain weather translation using generative adversarial networks," 2021, *arXiv:2103.05422*.
- [14] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2223–2232.
- [15] H. Sun, Y. Zheng, and Q. Lang, "Domain adaptation for synthesis of hazy images," *J. Comput. Commun.*, vol. 9, no. 10, pp. 142–151, 2021.
- [16] X. Qin, Z. Y. Wang, X. Bai Xie, and H. Jia, "Ffa-net: Feature fusion attention network for single image dehazing," in *Proc. AAAI Conf. Artif. Intell.*, 2020, vol. 34, pp. 11908–11915.
- [17] A. VishwanathP. SindagiR. Oza Yasarla, and V. M. Patel, "Prior-based domain adaptive object detection for hazy and rainy conditions," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 763–780.
- [18] Y. Chen, W. Li, C. Sakaridis, D. Dai, and L. Van Gool, "Domain adaptive faster r-cnn for object detection in the wild," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3339–3348.
- [19] J. Li, M. Jing, K. Lu, L. Zhu, and H. T. Shen, "Locality preserving joint transfer for domain adaptation," *IEEE Trans. Image Process.*, vol. 28, no. 12, pp. 6103–6115, Dec. 2019.
- [20] W. Li, L. Duan, D. Xu, and I. W. Tsang, "Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 6, pp. 1134–1148, Jun. 2014.
- [21] J. Li, B. Zhang, and D. Zhang, "Shared autoencoder gaussian process latent variable model for visual classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 9, pp. 4272–4286, Sep. 2018.
- [22] T. Sumi, B. K. Iwana, H. Hayashi, and S. Uchida, "Modality conversion of handwritten patterns by cross variational autoencoders," in *Proc. IEEE Int. Conf. Document Anal. Recognit.*, 2019, pp. 407–412.
- [23] E. Pantraki and C. Kotropoulos, "Face aging as image-to-image translation using shared-latent space generative adversarial networks," in *Proc. IEEE Glob. Conf. Signal Inf. Process.*, 2018, pp. 306–310.
- [24] T. Lindvall, "Lectures on the coupling method," 2002.
- [25] R. Gao et al., "Zero-VAE-GAN: Generating unseen features for generalized and transductive zero-shot learning," *IEEE Trans. Image Process.*, vol. 29, pp. 3665–3680, 2020.
- [26] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," 2016, *arXiv:1607.08022*.
- [27] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8798–8807.
- [28] M.-Y. Liu, T. Breuel, and J. Kautz, "Unsupervised image-to-image translation networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 700–708.
- [29] B. Li et al., "Benchmarking single-image dehazing and beyond," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, Jan. 2019.
- [30] K. Baek, Y. Choi, Y. Uh, J. Yoo, and H. Shim, "Rethinking the truly unsupervised image-to-image translation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 14154–14163.
- [31] W. Wang, T. Zhou, F. Yu, J. Dai, E. Konukoglu, and L. Van Gool, "Exploring cross-image pixel contrast for semantic segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 7303–7313.
- [32] L.-C. Chen, Y. G. Zhu, F. Papandreou Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 801–818.
- [33] M. Cordts et al., "The cityscapes dataset for semantic urban scene understanding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 3213–3223.
- [34] N. Zhang, L. Zhang, and Z. Cheng, "Towards simulating foggy and hazy images and evaluating their authenticity," in *Proc. Int. Conf. Neural Inf. Process.*, 2017, pp. 405–415.
- [35] M. Bijelic et al., "Seeing through fog without seeing fog: Deep multimodal sensor fusion in unseen adverse weather," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 11682–11692.
- [36] M. Everingham and J. Winn, "The pascal visual object classes challenge 2007 (voc2007) development kit," Univ. Leeds, Tech. Rep., 2007.



Meiqi Wang received the B.S. degree in electronic information science and technology in 2018 from Nanjing University, Nanjing, China, where she is currently working toward the Ph.D. degree in electronics and communications engineering.

She was a Visiting Student with the Wangxuan Institute of Computer Technology, Peking University, Beijing, China. Her current research interests include image processing and algorithm-hardware co-design for efficient and

robust deep learning systems.

Miss Wang was a recipient of the IEEE SOCC 2020 Best Paper Award and the IEEE SiPS 2019 Best Paper Award (the First Prize).



Tianqi Su received the B.S. degree in electronic information science and technology in 2020 from Nanjing University, Nanjing, China, where he is currently working toward the master's degree in electronics and communications engineering.

His current research interests include image restoration and algorithm-hardware codesign for robust deep learning systems.



Siyi Chen received the B.S. and master's degrees in electronic information science and technology from Nanjing University, Nanjing, China, in 2019 and 2022, respectively. Her current research interests include image/video processing/restoration and bad weather restoration.



Wenhan Yang (Member, IEEE) received the B.S. degree and Ph.D. degree (Hons.) in computer science from Peking University, Beijing, China, in 2012 and 2018, respectively.

He is currently an Associate Researcher with Peng Cheng Laboratory, Shenzhen, Guangdong, China. He has authored more than 50 technical articles in refereed journals and proceedings, and holds 9 granted patents. His current research interests include image/video processing/restoration, bad weather restoration, human-machine collaborative coding.

Dr. Yang was a recipient of the IEEE ICME-2020 Best Paper Award, the IFTC 2017 Best Paper Award, the IEEE CVPR-2018 UG2 Challenge First Runner-up Award, and the MSA-TC Best Paper Award of ISCAS 2022. He was the Candidate of CSIG Best Doctoral Dissertation Award in 2019. He served as the Area Chair of IEEE ICME-2021 and 2022, the Session Chair of IEEE ICME-2021, and the Organizer of IEEE CVPR-2019/2020/2021 UG2+ Challenge and Workshop.



Jiaying Liu (Senior Member, IEEE) received the Ph.D. degree (Hons.) in computer science from Peking University, Beijing, China, in 2010.

She is currently an Associate Professor, Boya Young Fellow with the Wangxuan Institute of Computer Technology, Peking University, China. She has authored more than 100 technical articles in refereed journals and proceedings, and holds 60 granted patents. Her current research interests include multimedia signal processing, compression, and computer vision.

Dr. Liu is a Senior Member of CSIG, and CCF. She was a Visiting Scholar with the University of Southern California, Los Angeles, California, from 2007 to 2008. She was a visiting Researcher with Microsoft Research Asia, in 2015 supported by the Star Track Young Faculty Award. She has served as a member of Multimedia Systems and Applications Technical Committee (MSA TC), and Visual Signal Processing and Communications Technical Committee (VSPC TC) in IEEE Circuits and Systems Society. She is also a member of Image, Video, and Multimedia (IVM) and Signal and Information Processing Theory and Methods (SIPTM) Technical Committee in APSIPA. She received the IEEE ICME 2020 Best Paper Award and IEEE MMSP 2015 Top10% Paper Award. She has also served as the Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CIRCUIT SYSTEMS FOR VIDEO TECHNOLOGY and *Journal of Visual Communication and Image Representation*, the Technical Program Chair of IEEE ICME-2021/ACM ICMR-2021, the Area Chair of CVPR-2021/ECCV-2020/ICCV-2019, and the CAS Representative at the ICME Steering Committee. She was the APSIPA Distinguished Lecturer (2016–2017).



Zhongfeng Wang (Fellow, IEEE) received the B.E. and M.S. degrees from the Department of Automation, Tsinghua University, Beijing, China, in 1988 and 1990, respectively. He received the Ph.D. degree in electrical and computer engineering from the University of Minnesota, Minneapolis, MN, USA, in 2000.

He has been working for Nanjing University, China, as a Distinguished Professor, since 2016. Previously, he worked for Broadcom Corporation, California, from 2007 to 2016 as a leading VLSI architect. Before that, he worked for Oregon State University and National Semiconductor Corporation. He is a world-recognized expert on Low-Power High-Speed VLSI Design for Signal Processing Systems. He has authored or coauthored more than 200 technical papers. His current research interests are in the area of optimized VLSI design for digital communications and deep learning.

Dr. Wang was a recipient of the multiple best paper awards received from the IEEE technical societies, among which is the VLSI Transactions Best Paper Award of 2007. He has edited one book VLSI and held more than 20 U.S. and China patents. In the current record, he has had many papers ranking among top 25 most (annually) downloaded manuscripts in IEEE TRANSACTIONS ON VLSI SYSTEMS. In the past, he has served as Associate Editor for IEEE TRANSACTIONS ON TCAS-I, T-CAS-II, AND T-VLSI for many terms. He has also served as TPC member and various chairs for tens of international conferences. Moreover, he has contributed significantly to the industrial standards. So far, his technical proposals have been adopted by more than 15 international networking standards. In 2015, he was elevated to the Fellow of IEEE for contributions to VLSI design and implementation of FEC coding.