

CoA: Towards Real Image Dehazing via Compression-and-Adaptation

Long Ma¹ Yuxin Feng² Yan Zhang² Jinyuan Liu¹ Weimin Wang¹
 Guang-Yong Chen³ Chengpei Xu^{4,*} Zhuo Su²

¹Dalian University of Technology ²Sun Yat-sen University ³Fuzhou University ⁴University of New South Wales

malone94319@gmail.com, {fengyx26, zhangy2779}@mail2.sysu.edu.cn, atlantis918@hotmail.com,

Chengpei.Xu@unsw.edu.au, wangweimin@dlut.edu.cn, gychen@fzu.edu.cn, suzhuo3@mail.sysu.edu.cn

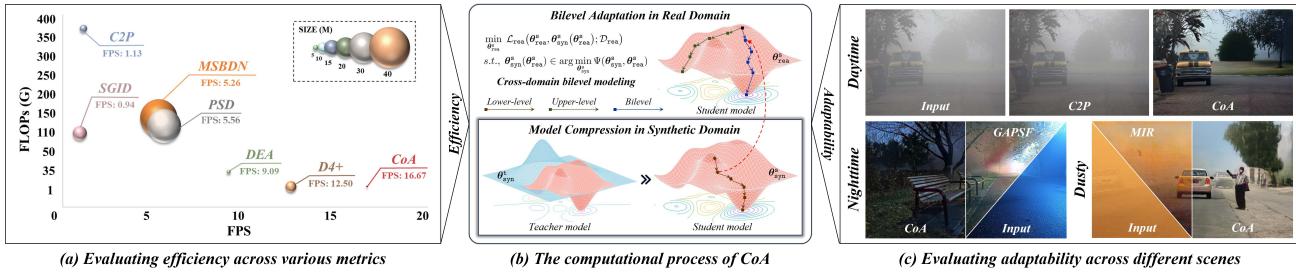


Figure 1. **Performance evaluation.** The proposed CoA incorporates model compression in synthetic domain for efficiency and bilevel adaptation in real domain for adaptability, as illustrated in the central sub-figure (b). The left sub-figure (a) presents efficiency across various metrics, clearly showing that our CoA outperforms others by a significant margin. The right sub-figure (c) shows adaptability across different scenes, where it is evident that our CoA consistently performs excellently.

Abstract

Learning-based image dehazing algorithms have shown remarkable success in synthetic domains. However, real image dehazing is still in suspense due to computational resource constraints and the diversity of real-world scenes. Therefore, there is an urgent need for an algorithm that excels in both efficiency and adaptability to address real image dehazing effectively. This work proposes a **Compression-and-Adaptation (CoA)** computational flow to tackle these challenges from a divide-and-conquer perspective. First, model compression is performed in the synthetic domain to develop a compact dehazing parameter space, satisfying efficiency demands. Then, a bilevel adaptation in the real domain is introduced to be fearless in unknown real environments by aggregating the synthetic dehazing capabilities during the learning process. Leveraging a succinct design free from additional constraints, our CoA exhibits domain-irrelevant stability and model-agnostic flexibility, effectively bridging the model chasm between synthetic and real domains to further improve its practical utility. Extensive evaluations and analyses underscore the approach's superiority and effectiveness. The code is publicly available at <https://github.com/fyxnl/COA>.

*Corresponding author.

1. Introduction

Image dehazing tasks aim to utilize the atmospheric optical model and advanced computing theory to achieve efficient restoration of scene information and image details through inverse imaging solution or data-driven model learning. It has drawn much attention in multiple emerging computer vision areas recently [46] [27] [28]. Similar to other low-level vision tasks [32] [11], image dehazing technology has evolved from an early stage focused on improving synthetic data metrics to a direction aimed at effectively generalizing across various real-world haze scenarios. Current approaches face dual challenges: efficiency-oriented methods achieve real-time processing but lack scene adaptability, while adaptability-focused techniques dynamically adjust strategies at the cost of high computational complexity that hinders deployment in resource-constrained environments. This paper addresses these limitations through novel architectural innovations in both computational efficiency and adaptive scene understanding.

1.1. Related Works

Efficiency-oriented Dehazing Methods. This work focuses on enhancing computational efficiency through optimization algorithms, model simplification, or machine learning methods to enable faster processing under limited resources. Broadly, dehazing methods can be divided

into image enhancement-based, physics-based, and image layering-based approaches. Image enhancement-based dehazing methods [24] treat the task as an image improvement problem, increasing contrast and color accuracy with techniques like histogram equalization and color correction. Physics-based methods [17], [28] are more specific, leveraging hazy image priors and constraints to aid inverse problem solving. Image layering-based approaches, such as that of Zhang *et al.* [41], increase efficiency by breaking down the image into scales or layers, progressively refining transmittance maps and enhancing detail using multilevel pyramid structures and dense connectivity. Despite their computational and real-time advantages, they struggle with stability and consistent quality, especially in complex real-world haze conditions, often resulting in partial dehazing and insufficient color clarity.

Adaptability-focused Dehazing Methods. Adaptive dehazing methods tailor their strategies based on specific image attributes to improve dehazing effectiveness. These methods encompass three main types: local feature-based, atmospheric scattering model-based, and domain adaptation-based approaches. Local feature-based methods [15] adaptively adjust dehazing intensity according to local contrast and brightness variations, which helps restore image details effectively. Atmospheric scattering model-based adaptive dehazing methods [20] dynamically adjust the transmittance estimation by analyzing different regions of the image to achieve more accurate restoration. These methods can, to some extent, balance both efficiency and adaptability. Domain adaptation-based dehazing methods [26, 37] have gradually emerged in the field of image dehazing in recent years. These methods learn a shared feature representation, enabling the model to simultaneously adapt to the features of both the source domain and the target domain. Although these methods yield superior dehazing results, their complex models and dynamic image-specific adjustments often increase computational costs, potentially impacting real-time applicability.

1.2. Our Contributions

To address the challenges of efficiency and adaptability in real image dehazing, this work proposes a new **Compression-and-Adaptation (CoA)** scheme as shown in Fig. 1 (b). The predefined model is compressed in the synthetic domain to ensure computational efficiency (refer to Fig. 1 (a)). Further, a bilevel adaptive learning is developed to leverage the dehazing capability acquired in the synthetic domain to improve adaptability (see Fig. 1 (c)) in the real domain. Our key contributions are summarized as

- Following the divide-and-conquer paradigm, we propose a new CoA learning strategy that first compresses the predefined model and then adapts it to unseen scenes, offering an efficient solution for real image dehazing. To the

best of our knowledge, this is the first work to successfully integrate both efficiency and adaptability, achieving the best of both worlds in real image dehazing.

- To adapt to unlabeled, diverse real-world scenes, we propose a new cross-domain bilevel model designed to learn dehazing parameters for the real domain constrained by synthetic dehazing learning. Additionally, we derive a bilevel adaptive learning scheme that effectively harnesses the synthetic dehazing capacity, preventing instability and enhancing adaptability.
- Benefiting from succinct design free from additional constraints, CoA shows two key properties on three standard real benchmarks. It shows stability across synthetic domains, improving 14.2% on average across four metrics (evaluated on three domains). CoA also offers flexibility with dehazing models, reducing computational cost by at least 73% in parameters (tested on three methods).

2. The Proposed Method

In this section, we first rethink real image dehazing, highlighting two key practical needs. We then propose model compression and bilevel adaptation to meet these needs, followed by an overview of the network architecture and training loss.

2.1. Rethinking Real Image Dehazing

Despite the strong performance of advanced image dehazing techniques [31, 36, 44] on synthetic datasets [8, 21, 30], these datasets often rely on simplifying assumptions, resulting in a substantial gap between synthetic and real-world domains. Addressing the core challenges in real-world image dehazing involves two key aspects:

- **Efficiency:** Edge devices often have limited computational resources compared to systems with ample capacity, requiring dehazing models to be highly efficient.
- **Adaptability:** Real-world imaging conditions vary dynamically. Dehazing models must therefore be adaptable to maintain reliable performance.

The latter is more challenging, relying on effective learning mechanisms like unsupervised learning [12]. However, current methods struggle with unsupervised learning due to difficulty in capturing statistical regularities, making transferring dehazing capabilities from the synthetic to real domain a viable alternative. To provide an intuitive presentation of the goal upon the above demands, we build

$$\begin{aligned} & \min_{\theta_{\text{real}}} f(\theta_{\text{real}}(\theta_{\text{syn}})), \\ & \text{s.t., } \begin{cases} \kappa(\theta_{\text{real}}) < \kappa(\theta_{\text{syn}}), & (\text{efficiency}) \\ \zeta(\theta_{\text{real}}) > \zeta(\theta_{\text{syn}}), & (\text{adaptability}) \end{cases} \end{aligned} \quad (1)$$

where θ_{syn} and θ_{real} represent parameters in synthetic and real domain, respectively. The function $\kappa(\cdot)$ and $\zeta(\cdot)$ denote the evaluator for efficiency and adaptability, respectively.

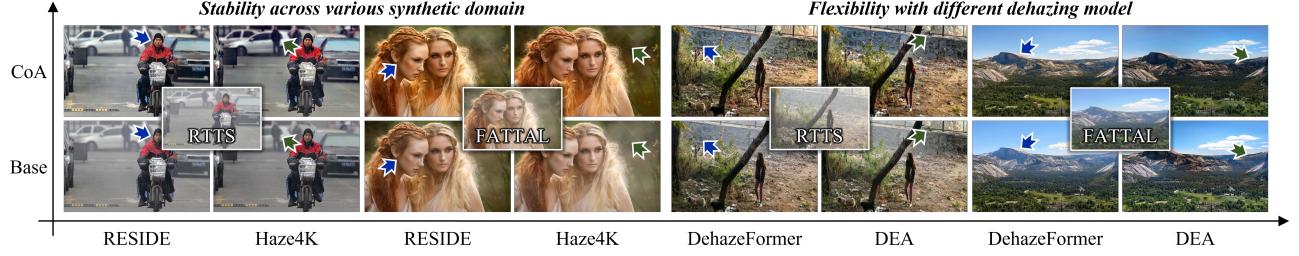


Figure 2. **Qualitative comparisons of algorithmic properties.** Arrows indicate specific regions that highlight visible differences.

To achieve this, we propose a divide-and-conquer strategy: the first phase focuses on model compression for efficiency, while the second phase employs bilevel adaptation for adaptability. Details are provided below.

2.2. MoC:Model Compression in Synthetic Domain

The main challenge in model compression is transferring dehazing capabilities from a large-scale parameter space to a smaller one. Here we propose a composite loss function¹ with context alignment to realize it.

$$\min_{\theta_{\text{syn}}^s} \mathcal{L}_{\text{syn}}(\theta_{\text{syn}}^s) + \mathcal{L}_a(\theta_{\text{syn}}^s, \theta_{\text{syn}}^t), \quad (2)$$

where θ_{syn}^t and θ_{syn}^s are the pre-trained teacher model and desired student model in the synthetic domain, respectively. The function \mathcal{L}_{syn} and \mathcal{L}_a represent a group of supervised losses and the context alignment loss, respectively.

During model compression, the student model follows the teacher model’s training strategy on paired synthetic data. To maximize the retention of the teacher model’s capabilities, we introduce a context alignment loss (\mathcal{L}_a) that aligns the outputs of the teacher and student encoders at corresponding positions. By weighting the loss functions across layers, this approach enables fine-grained feature alignment, improving knowledge transfer and model optimization. The layered alignment ensures both global feature distributions and layer-specific details are accurately matched, resulting in more precise compression of the teacher model into the student model.

2.3. BiA:Bilevel Adaptation to Real Domain

After compressing the model for efficiency, we address adaptation to real-world scenes by proposing a cross-domain bilevel model. This is followed by a bilevel adaptive learning scheme to enable smooth transition from the synthetic to real domain.

2.3.1. Cross-Domain Bilevel Modeling

The gap between the synthetic and real domains makes adaptation challenging. The dehazing parameters in the synthetic and real domains (θ_{rea}^s) have a nested subordinate

Algorithm 1 Learning via Compression-and-Adaptation

Require: The pre-trained teacher model θ_{syn}^t , the initial parameter θ_{syn}^s , step-sizes $\eta_{\text{syn}}^m, \eta_{\text{syn}}^b$, coefficient α , iteration numbers N in MoC phase and T in BiA phase.
Ensure: The optimal parameters θ_{rea}^s .

```

1: % The MoC Phase
2: for  $n = 0 : N - 1$  do
3:   Calculate the gradient  $\mathbf{g}$  of Eq. (2).
4:    $\theta_{\text{syn}}^s(n + 1) = \theta_{\text{syn}}^s(n) - \eta_{\text{syn}}^m \mathbf{g}(\theta_{\text{syn}}^s(n), \theta_{\text{syn}}^t)$ .
5: end for
6: Initialize the dehazing model’s parameter in the real domain by  $\theta_{\text{rea}}^s = \theta_{\text{syn}}^s(N)$ .
7: % The BiA Phase
8: for  $t = 0 : T - 1$  do
9:   % Update the lower-level parameter  $\theta_{\text{syn}}^s$  in  $\mathcal{D}_{\text{rea}}$ .
10:   $\theta_{\text{syn}}^s(t + 1) = \theta_{\text{syn}}^s(t) - \eta_{\text{syn}}^b \frac{\partial (\mathcal{L}_{\text{rea}} + \mathcal{L}_1)}{\partial \theta_{\text{syn}}^s}$ .
11:  % Update the upper-level parameter  $\theta_{\text{rea}}^s$  by EMA.
12:   $\theta_{\text{rea}}^s(t + 1) = \alpha \theta_{\text{rea}}^s(t) + (1 - \alpha) \theta_{\text{syn}}^s(t + 1)$ .
13: end for
14: return  $\theta_{\text{rea}}^s(T)$ .

```

relationship. This coupling can be effectively modeled using bilevel programming techniques [9, 29]. From the perspective of hyperparameter optimization, our cross-domain bilevel model is written as:

$$\begin{aligned} & \min_{\theta_{\text{rea}}^s} \mathcal{L}_{\text{rea}}(\theta_{\text{rea}}^s, \theta_{\text{syn}}^s(\theta_{\text{rea}}^s); \mathcal{D}_{\text{rea}}), \\ & \text{s.t., } \theta_{\text{syn}}^s(\theta_{\text{rea}}^s) \in \arg \min_{\theta_{\text{syn}}^s} \Psi(\theta_{\text{rea}}^s, \theta_{\text{syn}}^s), \end{aligned} \quad (3)$$

where

$$\Psi(\theta_{\text{rea}}^s, \theta_{\text{syn}}^s) = \mathcal{L}_{\text{rea}}(\theta_{\text{syn}}^s; \mathcal{D}_{\text{rea}}) + \mathcal{L}_1(\theta_{\text{syn}}^s, \theta_{\text{rea}}^s), \quad (4)$$

where \mathcal{L}_{rea} and \mathcal{L}_1 represent the loss constraints² for real domain and alignment. The datasets \mathcal{D}_{rea} and \mathcal{D}_{syn} are captured in the real and synthetic domains, respectively.

Actually, suppose an effective loss \mathcal{L}_{rea} could be designed to constrain the output in the real domain, the adaptation might be achieved solely through the upper-level model

¹Details of the specific loss functions can be found in Sec. 2.4.

²Their specific forms can be found in Sec. 2.4.

Testing Benchmark	Metrics	Stability across various synthetic domain						Flexibility with different dehazing model					
		RESIDE		Haze4K		THaze		MSBDN		DehazeFormer		DEA	
		Base	CoA	Base	CoA	Base	CoA	Base	CoA	Base	CoA	Base	CoA
RTTS [22]	FADE↓	1.4753	1.2687	1.7368	1.0850	1.2078	0.8594	1.3354	1.1556	1.1866	0.9325	1.2117	0.9332
	PM2.5↓	163.77	139.04	170.03	110.84	131.77	88.871	135.68	95.793	126.77	115.00	121.98	89.813
	Entropy↑	7.2963	7.3458	7.3446	7.4637	7.4319	7.5798	7.3968	7.5382	7.4815	7.5511	7.4623	7.5426
	BIQME↑	0.5562	0.5639	0.5511	0.5748	0.5773	0.5932	0.5696	0.5819	0.5733	0.5928	0.5789	0.5860
URHI [22]	FADE↓	1.4463	1.3013	1.7263	1.0278	1.3146	0.9272	1.3488	0.9549	1.3263	1.1726	1.2796	1.0321
	PM2.5↓	164.36	145.01	172.98	119.97	140.35	98.793	146.43	99.856	141.11	125.87	133.97	108.94
	Entropy↑	7.2867	7.3152	7.3711	7.4657	7.4652	7.5928	7.4144	7.5574	7.5009	7.5817	7.4920	7.5658
	BIQME↑	0.5499	0.5542	0.5481	0.5681	0.5821	0.5961	0.5727	0.5832	0.5757	0.5948	0.5786	0.5863
FATTAL [6]	FADE↓	0.5136	0.4778	0.5828	0.3734	0.4203	0.3142	0.4315	0.3327	0.4271	0.3875	0.4270	0.3976
	PM2.5↓	81.738	80.289	81.401	110.97	62.312	91.639	69.272	104.56	78.093	60.693	81.387	80.723
	Entropy↑	7.3847	7.4575	7.3999	7.5006	7.4482	7.5858	7.4121	7.5460	7.4566	7.5610	7.4315	7.5569
	BIQME↑	0.5645	0.5802	0.5554	0.5919	0.5945	0.6189	0.5807	0.6055	0.5927	0.6117	0.5903	0.6097

Table 1. **Quantitative comparison of algorithmic properties.** In the left portion, our designed architecture serves as the baseline teacher model. In the right portion, the THaze dataset is used as the synthetic domain.

(initialized with θ_{syn}^s). However, as discussed in Sec. 2.1, the challenge of accurately modeling real haze distributions makes this process difficult to implement. To this end, we introduce lower-level sub-optimization to ensure a feasible solution and maintain the correct optimization trajectory. Experimental validation can be seen in Sec. 5.2.

2.3.2. Bilevel Adaptive Learning

Without compromising the model’s dehazing capability in the synthetic domain, narrowing the performance gap between the synthetic and real domains is a central focus. Inspired by the Exponential Moving Average (EMA) method [2], we draw on its smoothing characteristics during the training process and introduce a dynamic weight adjustment mechanism to achieve a balance in learning between the synthetic and real domains, formulated as

$$\begin{cases} \theta_{\text{syn}}^s(t+1) = \theta_{\text{syn}}^s(t) - \eta_{\text{syn}}^b \frac{\partial(\mathcal{L}_{\text{rea}} + \mathcal{L}_1)}{\partial \theta_{\text{syn}}^s}, \\ \theta_{\text{rea}}^s(t+1) = \alpha \theta_{\text{rea}}^s(t) + (1-\alpha) \theta_{\text{syn}}^s(t+1), \end{cases} \quad (5)$$

where α represents the moving average coefficient, and \mathcal{L}_1 denotes the ℓ_1 -norm supervised loss used to reduce the risk of the model overfitting to the real domain. Specifically, during each update, θ_{rea}^s is adjusted by θ_{syn}^s , allowing the model to strike a balance between the synthetic and real domains throughout the learning process. The overall algorithm can be seen in Alg. 1.

2.4. Network Architecture and Training Loss

This section introduces the details about network architecture and training loss adopted in this work.

2.4.1. Network Architecture

We emphasize that our primary focus is on developing an effective learning mechanism to enhance the model’s adaptability to the real domain. Notably, our approach offers

flexibility³ with different dehazing models. To this end, we adopt a hybrid architectural design, incorporating a pre-trained Res2Net encoder [10] with the decoder of multi-scale boosted dehazing network [5].

2.4.2. Loss Functions in MoC Phase

In the MoC process, to enable the student model to learn the features extracted by the teacher model more effectively and to enhance the quality the generated images. We define

$$\mathcal{L}_{\text{syn}} = \lambda_{su} \ell_{su} + \lambda_{ss} \ell_{ss} + \lambda_{pe} \ell_{pe}, \quad (6)$$

where ℓ_{su} , ℓ_{ss} , ℓ_{pe} denote ℓ_1 -norm supervised loss, SSIM loss, and perceptual loss, respectively. The coefficient λ_{su} , λ_{ss} , λ_{pe} are positive balancing parameters.

To more effectively enhance consistency between the student and teacher models in the feature space, the design of \mathcal{L}_{a} incorporates a similarity measure of feature distributions. Its formulation can be written as follows:

$$\mathcal{L}_{\text{a}} = \sum_{i=0}^{L-1} w_i \cdot ((T_i - S_i)^2), \quad (7)$$

where w_i represents the weights of i -th layer in the teacher feature map T_i and the student feature map S_i .

2.4.3. Loss Functions in BiA Phase

Here we incorporate the CLIP model to construct the training loss \mathcal{L}_{rea} to constrain the output in the real domain. We select haze images $I_{\mathcal{H}}$ and clear images $I_{\mathcal{C}}$ as references, initialize corresponding prompts $T_{\mathcal{H}}$ and $T_{\mathcal{C}}$, and input them into the CLIP text encoder, while feeding the images into the image encoder. The sample categories are predicted using the text-image similarity in the CLIP space. To minimize the classification error between positive and negative

³For further details, please refer to Sec. 3.

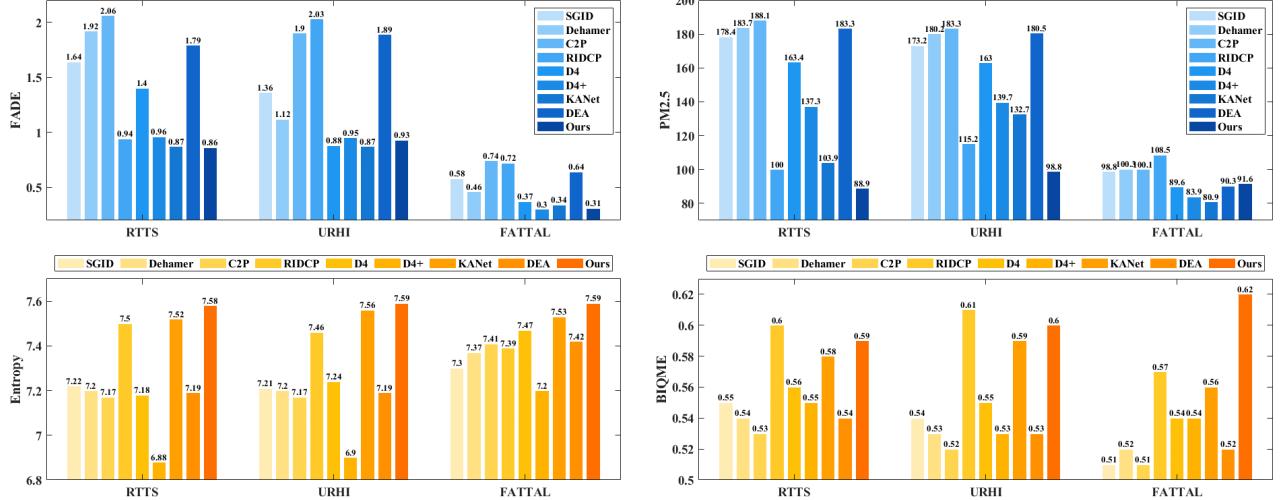


Figure 3. **Quantitative comparison on three real-world datasets.** Four no-reference image quality assessments were calculated.



Figure 4. **Qualitative comparisons on daytime haze and dusty scenes.** All these observations come from RTTS and URHI datasets.

samples, we employ the binary cross entropy loss to optimize the text prompt pair. After training the text prompts to effectively distinguish between real haze and clear images, we can derive \mathcal{L}_{rea} based on the contrastive similarity loss:

$$\mathcal{L}_{\text{rea}} = \frac{e^{\cos(\Phi_{\text{image}}(I_{\mathcal{R}}), \Phi_{\text{text}}(T_{\mathcal{H}}))}}{\sum_{i \in \{\mathcal{H}, \mathcal{C}\}} e^{\cos(\Phi_{\text{image}}(I_{\mathcal{R}}), \Phi_{\text{text}}(T_i))}}, \quad (8)$$

where $I_{\mathcal{R}}$ corresponds to the real haze image processed by the model obtained through the model compression process.

3. Exploring Algorithmic Property

CoA presents an effective learning strategy to improve efficiency and adaptability for real-world scenes, without assumptions on the synthetic domain or teacher model. We analyze CoA’s properties, highlighting its stability across synthetic domains and flexibility with different models.

3.1. Stability across Various Synthetic Domains

We perform three synthetic benchmarks (RESIDE [21] (the outdoor subset), Haze4K [30], and THaze [8]) and

three real-world datasets (RTTS [21], URHI [21], and FATTAL [6]) to assess our stability. As shown in the left section of Table 1, CoA consistently improves scores across four no-reference image quality metrics. Notably, due to the diverse scenes in THaze, CoA achieves near-optimal performance within the synthetic domain. We use THaze as the synthetic domain in our comparative experiments. Fig. 2 (left) compares visual results. The dehazing effects of CoA is significant, especially in distant regions highlighted by arrows. Overall, CoA improves the original teacher model across synthetic domains, demonstrating its stability.

3.2. Flexibility with Different Dehazing Models

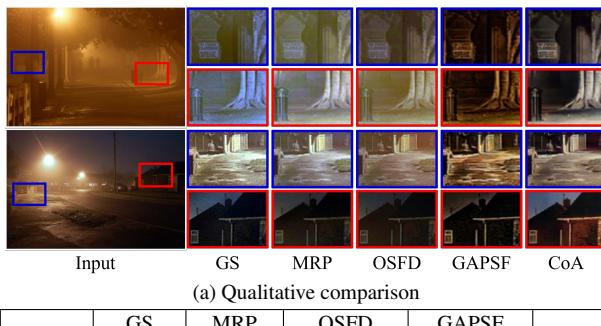
CoA’s teacher model-irrelevant nature allows us to enhance the adaptability of existing dehazing models across various scenes. We applied CoA to three dehazing models (MS-BDN [5], DehazeFormer [34], and DEA [3]) as teacher models. The right section of Table 1 shows significant performance improvements before and after applying CoA across different models and test settings. Fig. 2 (right) compares visual results using DehazeFormer and DEA, highlighting areas where CoA enhances dehazing for more visually appealing outcomes. In summary, CoA demonstrates notable flexibility with various dehazing models. The intro-

Method	RTTS				URHI				FATTAL			
	FADE↓	PM2.5↓	Entropy↑	BIQME↑	FADE↓	PM2.5↓	Entropy↑	BIQME↑	FADE↓	PM2.5↓	Entropy↑	BIQME↑
AirNet	1.3342	155.84	7.2959	0.5471	1.6591	143.02	6.9143	0.5258	0.4063	76.114	7.3545	0.5315
WeatherDiff	2.4102	193.27	7.1443	0.5155	2.1429	180.26	7.1419	0.5102	0.8749	118.79	7.3621	0.4937
DiffUIR	2.1305	189.61	7.1722	0.5303	1.9879	180.56	7.1811	0.5205	0.8405	116.68	7.3609	0.5049
Ours	0.8594	88.871	7.5798	0.5932	0.9272	98.793	7.5928	0.5961	0.3142	91.639	7.5858	0.6189

Table 2. Quantitative comparison on *daytime haze* scenes. All compared methods are designed for multi-weather image restoration.



Figure 5. Qualitative comparisons corresponding to Table 2. All examples are sourced from RTTS and URHI datasets.



(a) Qualitative comparison

Metrics	GS (ICCV'15)	MRP (CVPR'17)	OSFD (ACM MM'20)	GAPSF (ACM MM'23)	CoA
PM2.5↓	<u>102.43</u>	106.20	146.93	110.59	81.141
Entropy↑	6.7025	7.0488	6.9412	6.3769	<u>6.9907</u>
BIQME↑	0.4552	0.4722	<u>0.4684</u>	0.3895	0.5112

(b) Quantitative comparison

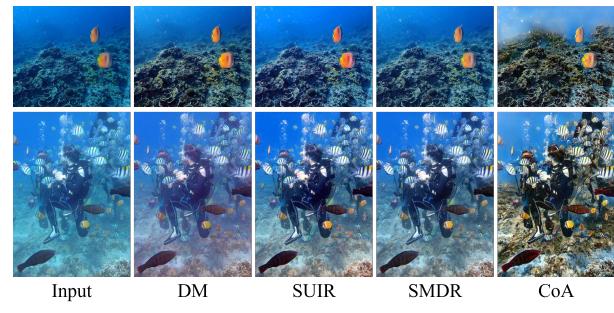
Figure 6. Performance evaluation on *nighttime haze* scenes.

duction of CoA significantly reduces the model’s parameter count, improving computational efficiency. Specifically, the parameters in the three dehazing models decrease by 94.06%, 86.74%, and 73.89%, respectively. This reduction accelerates inference speed and lowers resource requirements, making the models more lightweight and better suited for practical applications.

4. Experimental Results

4.1. Implementation Details

Training settings. The CoA model was implemented using the PyTorch framework on a single NVIDIA RTX 3090 GPU. The Adam optimizer was employed to update each module’s parameters, with β_1 , β_2 , and ε set to 0.9, 0.999, and $1e^{-8}$, respectively. The initial learning rate was set to $1e^{-4}$ and gradually decayed to $1e^{-6}$ using a cosine annealing schedule. The algorithm was trained on RGB channels



(a) Qualitative comparison

Metrics	DM (ACM MM'23)	SUIR (CVPR'23)	SMDR (AAAI'24)	CoA
UIQM↑	<u>4.0965</u>	3.9853	4.0314	4.1284
CCF↑	26.651	31.417	28.439	<u>30.878</u>

(b) Quantitative comparison

Figure 7. Performance evaluation on *underwater* scenes.

with data augmentation via random 90° , 180° , and 270° rotations and horizontal flipping. We cropped random 256×256 sub-regions from images, further expanding the training set. Using a pre-trained Res2Net encoder, the teacher model achieved strong performance after just 20 epochs.

Benchmarks and metrics. Three synthetic datasets (Haze4K, RESIDE (the outdoor subset), and THaze served as the synthetic domain, while three real-world datasets (RTTS, Fattal, and URHI were used for testing in the real domain. We employed four objective metrics for real-world haze scenes: FADE [4], PM2.5 [14], Entropy, and BIQME [13]. For underwater enhancement, we used UIQM and CCF to assess image quality.

Compared methods. We evaluated our CoA method against state-of-the-art techniques across daytime, dusty, nighttime, and underwater conditions. For daytime and dusty conditions, we compared CoA with SGID [1], Dehamer [16], C2P [46], KANet [7], DEA, D4 [38], D4⁺ [39],

Metrics	SGID (TIP'22)	Dehamer (CVPR'22)	C2P (CVPR'23)	RIDCP [†] (CVPR'23)	D4 (CVPR'22)	D4+ (IJCV'24)	KANet (TPAMI'24)	DEA (TIP'24)	Ours
SIZE (M)	13.87	132.40	7.17	28.72	10.70	10.70	55.25	3.65	1.69
FLOPs (G)	108.40	48.91	352.90	144.43	2.82	2.82	4.42	32.20	2.67
Time (ms)	1280×720 1920×1080 2560×1440	240.01 878.94 1514.82	136.36 295.98 519.52	1169.18 2531.43 4409.12	630.64 1588.26 —	51.25 105.31 185.55	51.25 105.31 185.55	22.93 205.40 10631	86.33 190.78 334.78
	† This work requires excessive video memory for high-resolution images, resulting in the failure to produce results.								

† This work requires excessive video memory for high-resolution images, resulting in the failure to produce results.

Table 3. Evaluating the computational efficiency. Note that the size of the testing image for FLOPs calculation is 224×224.

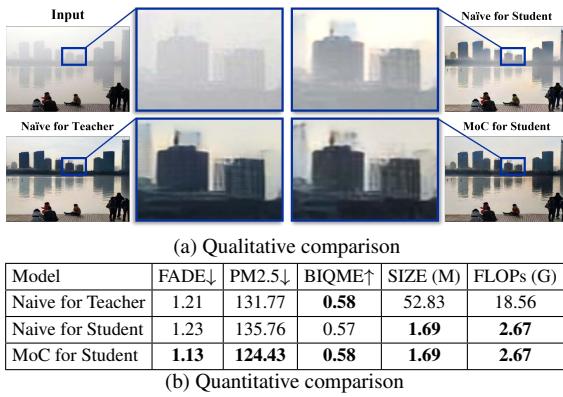


Figure 8. Effects of MoC phase.

and RIDCP [37]. Additionally, we assessed CoA against multi-weather restoration models: AirNet [23], WeatherDiff [33], and DiffUIR [45]. For nighttime dehazing, we benchmarked against GS [25], MRP [42], OSFD [43], and GAPSF [19], and for underwater scenes, we compared with DM [35], SUIR [18], and SMDR [40].

4.2. Evaluation on Regular Scenes

Quantitative comparisons. In Fig. 3, we presented a comparison of eight state-of-the-art dehazing algorithms with our proposed method, using four authoritative unsupervised evaluation metrics. It is evident that our method ranked among the top in almost all metrics, achieving optimal or near-optimal performance in most cases.

Qualitative comparisons. As shown in Fig. 4, the supervised methods (SGID, C2P, Dehamer, DEA) and unsupervised methods (D4, D4⁺) struggle with haze removal and generalization. RIDCP causes color distortion and over-dehazing, especially in sandstorm and colorful haze scenes. KANet has limitations in detail and color restoration. In contrast, CoA excels at haze removal while preserving fine textures, resulting in more natural and realistic images.

Comparisons with multi-weather restorers. Here we compared CoA with state-of-the-art multi-weather degradation methods. As shown in Table 2, CoA ranks second in one metric, outperforming others in all others, highlighting its superiority. Fig. 5 reveals that while existing methods excel in specific scenarios, they show limitations in handling

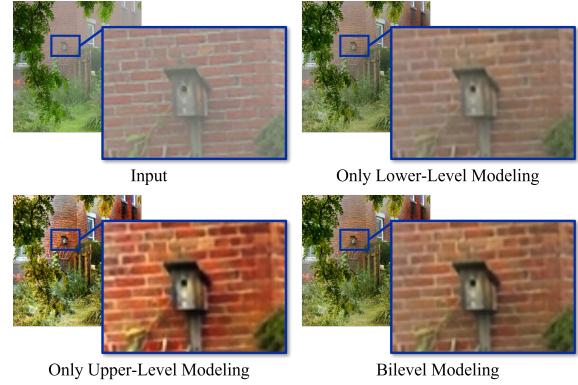


Figure 9. Necessity of bilevel modeling.

the full spectrum of haze conditions.

Efficiency comparisons. Table 3 presents a comprehensive comparison of our proposed CoA method with existing approaches, evaluating parameters, floating-point operations per second (FLOPs), and processing time across various image resolutions. The results clearly show that our method has the smallest number of parameters and FLOPs among all the evaluated methods. Notably, as the image resolution increases, our method demonstrates a significant advantage in processing time, indicating that our model achieves a favorable balance between performance and complexity.

4.3. Adaptability Verification

Nighttime haze scene. Compared to state-of-the-art nighttime dehazing algorithms on the NHRW dataset (Fig. 6 (a)), many existing methods fail to handle overexposure or low-light conditions, resulting in overly dark or distorted images due to noise and artificial light scattering. In contrast, CoA excels in real-world nighttime haze scenes, effectively removing haze, suppressing noise, and preserving details. As shown in Fig. 6 (b), our method outperforms others across three authoritative objective metrics

Underwater scene. Compared to state-of-the-art underwater dehazing algorithms, DM results in reddish and blurry images, failing to correct color distortion effectively, while SUIR and SMDR struggle to fully eliminate haze. As shown in Fig. 7, our method significantly reduces haze impact and enhances clarity and detail visibility, closely align-

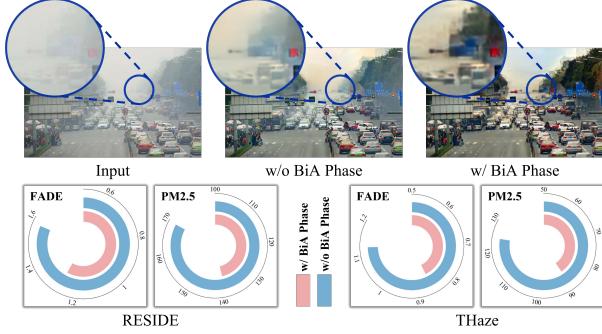


Figure 10. Effects of BiA phase.

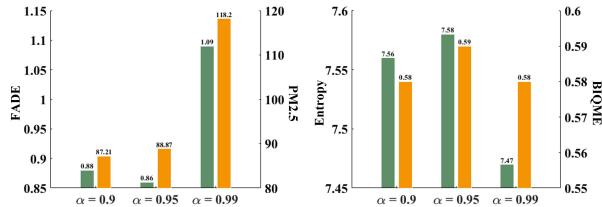


Figure 11. Parameters Analysis. The RTTS dataset is adopted.

ing with human visual perception standards.

5. Analytical Experiments

5.1. Effects of MoC Phase

Without feature transfer from the “naive for teacher” (*naive: end-to-end training, the same below*) model via compression, the “naive for student” model struggles to capture core data patterns, particularly in complex or extreme weather scenarios, resulting in limited dehazing. As illustrated in Fig. 8, while “MoC for student” may not restore details as effectively as “naive for teacher”, it successfully inherits high-quality features. Even under resource constraints, it surpasses “naive for student” in dehazing and performs well on unsupervised metrics.

5.2. Necessity of Bilevel Modeling

Here we highlight the critical role of bilevel modeling by training the same network model on the THaze dataset. As shown in Fig. 9, the results on the real test set show marked contrasts: employing only the lower-level modeling strategy results in suboptimal dehazing, while relying solely on the upper-level strategy with CLIP fine-tuning produces unnatural effects with notable color distortion. Different from them, the bilevel modeling strategy enables both effective dehazing and natural visual quality, achieving a balance between dehazing performance and perceptual realism.

5.3. Effects of BiA Phase

In Fig. 10, we evaluate effects of the BiA phase. The results clearly show significant improvements in both FADE and



Figure 12. Limitations. The example is from RS-Haze [34].

PM2.5 metrics following the BiA phase. Notably, in images with complex backgrounds or high noise levels, our model outperforms the baseline, demonstrating the BiA phase’s effectiveness in capturing key features and improving robustness in challenging scenarios.

5.4. Parameters Analyses

We test different values of the parameter α during the BiA phase, as shown in Fig. 11. The results indicate that $\alpha = 0.95$ yields almost optimal performance. Notably, this setting achieves the fastest convergence and exhibits superior stability and accuracy across multiple evaluation metrics, clearly outperforming other parameter settings.

5.5. Limitations

While effective in many scenarios, our method struggles with remote sensing haze images due to additional imaging constraints. As depicted in Fig. 12, although it improves dehazing (outperforming DEA), non-uniform fog occlusion remains. This stems from varying haze effects across spectral bands in hyperspectral images, requiring multi-band integration for effective dehazing.

6. Conclusion

This work proposes a novel compression-and-adaptation scheme to tackle the problem of unbalance between dehazing quality and computational efficiency in real image dehazing. Following the divide-and-conquer paradigm, we design a new CoA learning strategy that first compresses the predefined model and then adapts it to unlabeled, diverse real-world scenes. We not only make a thorough exploration to take on the excellent properties of CoA, but also we perform extensive experiments to indicate our CoA method’s superiority in terms of image dehazing quality and computational efficiency.

Acknowledgments

This work was supported by the Postdoctoral Fellowship Program of CPSF (No. GZB20240098), China Postdoctoral Science Foundation (No. 2023M740491), Dalian Science and Technology Innovation Fund-Young Tech Star (No. 2023RQ017), Fundamental Research Funds for the Central Universities (No. DUT24BS011).

References

- [1] Haoran Bai, Jinshan Pan, Xinguang Xiang, and Jinhui Tang. Self-guided image dehazing using progressive feature fusion. *IEEE Transactions on Image Processing*, 31:1217–1229, 2022. 6
- [2] Zhaowei Cai, Avinash Ravichandran, Subhransu Maji, Charles Fowlkes, Zhuowen Tu, and Stefano Soatto. Exponential moving average normalization for self-supervised and semi-supervised learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 194–203, 2021. 4
- [3] Zixuan Chen, Zewei He, and Zhe-Ming Lu. Dea-net: Single image dehazing based on detail-enhanced convolution and content-guided attention. *IEEE Transactions on Image Processing*, 2024. 5
- [4] Lark Kwon Choi, Jaehye You, and Alan C. Bovik. Referenceless prediction of perceptual fog density and perceptual image defogging. *IEEE Transactions on Image Processing*, 24(11):3888–3901, 2015. 6
- [5] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. Multi-scale boosted dehazing network with dense feature fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2157–2167, 2020. 4, 5
- [6] Raanan Fattal. Dehazing using color-lines. *ACM transactions on graphics*, 34(1):1–14, 2014. 4, 5
- [7] Yuxin Feng, Long Ma, Xiaozhe Meng, Fan Zhou, Risheng Liu, and Zhuo Su. Advancing real-world image dehazing: Perspective, modules, and training. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12):9303–9320, 2024. 6
- [8] Yuxin Feng, Zhuo Su, Long Ma, Xin Li, Risheng Liu, and Fan Zhou. Bridging the gap between haze scenarios: A unified image dehazing model. *IEEE Transactions on Circuits and Systems for Video Technology*, 34(11):11070–11085, 2024. 2, 5
- [9] Luca Franceschi, Paolo Frasconi, Saverio Salzo, Riccardo Grazzi, and Massimiliano Pontil. Bilevel programming for hyperparameter optimization and meta-learning. In *Proceedings of the International Conference on Machine Learning*, pages 1568–1577, 2018. 3
- [10] Minghan Fu, Huan Liu, Yankun Yu, Jun Chen, and Keyan Wang. DW-GAN: A discrete wavelet transform GAN for nonhomogeneous dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 203–212, 2021. 4
- [11] Xueyang Fu, Jie Xiao, Yurui Zhu, Aiping Liu, Feng Wu, and Zheng-Jun Zha. Continual image deraining with hypergraph convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(8):9534–9551, 2023. 1
- [12] Alona Golts, Daniel Freedman, and Michael Elad. Unsupervised single image dehazing using dark channel prior loss. *IEEE Transactions on Image Processing*, 29:2692–2701, 2019. 2
- [13] Ke Gu, Dacheng Tao, Junfei Qiao, and Weisi Lin. Learning a no-reference quality assessment model of enhanced images with big data. *IEEE Transactions on Neural Networks and Learning Systems*, 29(4):1301–1313, 2018. 6
- [14] Ke Gu, Junfei Qiao, and Xiaoli Li. Highly efficient picture-based prediction of PM2.5 concentration. *IEEE Transactions on Industrial Electronics*, 66(4):3176–3184, 2019. 6
- [15] Jie Gui, Xiaofeng Cong, Lei He, Yuan Yan Tang, and James Tin-Yau Kwok. Illumination controllable dehazing network based on unsupervised retinex embedding. *IEEE Transactions on Multimedia*, 26:4819–4830, 2024. 2
- [16] Chunle Guo, Qixin Yan, Saeed Anwar, Runmin Cong, Wenqi Ren, and Chongyi Li. Image dehazing transformer with transmission-aware 3d position embedding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5802–5810, 2022. 6
- [17] Kaiming He, Jian Sun, and Xiaou Tang. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353, 2010. 2
- [18] Shirui Huang, Keyan Wang, Huan Liu, Jun Chen, and Yun-song Li. Contrastive semi-supervised learning for underwater image restoration via reliable bank. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18145–18155, 2023. 7
- [19] Yeying Jin, Beibei Lin, Wending Yan, Yuan Yuan, Wei Ye, and Robby T. Tan. Enhancing visibility in nighttime haze images using guided APSF and gradient adaptive convolution. In *Proceedings of the ACM International Conference on Multimedia*, pages 2446–2457, 2023. 7
- [20] Mingye Ju, Can Ding, Charles A. Guo, Wenqi Ren, and Dacheng Tao. IDRLP: image dehazing using region line prior. *IEEE Transactions on Image Processing*, 30:9043–9057, 2021. 2
- [21] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2018. 2, 5
- [22] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Benchmarking single-image dehazing and beyond. *IEEE Transactions on Image Processing*, 28(1):492–505, 2019. 4
- [23] Boyun Li, Xiao Liu, Peng Hu, Zhongqin Wu, Jiancheng Lv, and Xi Peng. All-in-one image restoration for unknown corruption. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17431–17441, 2022. 7
- [24] Pengyue Li, Jiandong Tian, Yandong Tang, Guolin Wang, and Chengdong Wu. Deep retinex network for single image dehazing. *IEEE Transactions on Image Processing*, 30: 1100–1115, 2021. 2
- [25] Yu Li, Robby T. Tan, and Michael S. Brown. Nighttime haze removal with glow and multiple light colors. In *Proceedings of the International Conference on Computer Vision*, pages 226–234, 2015. 7
- [26] Yudong Liang, Bin Wang, Wangmeng Zuo, Jiaying Liu, and Wenqi Ren. Self-supervised learning and adaptation for single image dehazing. In *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, pages 1137–1143, 2022. 2

- [27] Pengyang Ling, Huaian Chen, Xiao Tan, Yi Jin, and Enhong Chen. Single image dehazing using saturation line prior. *IEEE Transactions on Image Processing*, 32:3238–3253, 2023. 1
- [28] Jun Liu, Ryan Wen Liu, Jianing Sun, and Tieyong Zeng. Rank-one prior: Real-time scene recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(7):8845–8860, 2023. 1, 2
- [29] Risheng Liu, Jiaxin Gao, Jin Zhang, Deyu Meng, and Zhouchen Lin. Investigating bi-level optimization for learning and vision from a unified perspective: A survey and beyond. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12):10045–10067, 2021. 3
- [30] Ye Liu, Lei Zhu, Shunda Pei, Huazhu Fu, Jing Qin, Qing Zhang, Liang Wan, and Wei Feng. From synthetic to real: Image dehazing collaborating with unlabeled real data. In *Proceedings of the ACM International Conference on Multimedia*, pages 50–58, 2021. 2, 5
- [31] Yidi Liu, Xingbo Wang, Yurui Zhu, Xueyang Fu, and Zheng-Jun Zha. Sdcnet: Spatially-adaptive deformable convolution networks for hr nonhomogeneous dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6682–6691, 2024. 2
- [32] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5637–5646, 2022. 1
- [33] Ozan Özdenizci and Robert Legenstein. Restoring vision in adverse weather conditions with patch-based denoising diffusion models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–12, 2023. 7
- [34] Yuda Song, Zhuqing He, Hui Qian, and Xin Du. Vision transformers for single image dehazing. *IEEE Transactions on Image Processing*, 32:1927–1941, 2023. 5, 8
- [35] Yi Tang, Hiroshi Kawasaki, and Takafumi Iwaguchi. Underwater image enhancement by transformer-based diffusion model with non-uniform sampling for skip strategy. In *Proceedings of the ACM International Conference on Multimedia*, pages 5419–5427, 2023. 7
- [36] Zhongze Wang, Haitao Zhao, Jingchao Peng, Lujian Yao, and Kaijie Zhao. Oder: Orthogonal decoupling contrastive regularization for unpaired image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 25479–25489, 2024. 2
- [37] Ruiqi Wu, Zheng-Peng Duan, Chun-Le Guo, Zhi Chai, and Chongyi Li. RIDCP: revitalizing real image dehazing via high-quality codebook priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22282–22291, 2023. 2, 7
- [38] Yang Yang, Chaoyue Wang, Risheng Liu, Lin Zhang, Xiaojie Guo, and Dacheng Tao. Self-augmented unpaired image dehazing via density and depth decomposition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2037–2046, 2022. 6
- [39] Yang Yang, Chaoyue Wang, Xiaojie Guo, and Dacheng Tao. Robust unpaired image dehazing via density and depth de- composition. *International Journal of Computer Vision*, 132(5):1557–1577, 2024. 6
- [40] Dehuan Zhang, Jingchun Zhou, Chunle Guo, Weishi Zhang, and Chongyi Li. Synergistic multiscale detail refinement via intrinsic supervision for underwater image enhancement. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 7033–7041, 2024. 7
- [41] He Zhang and Vishal M Patel. Densely connected pyramid dehazing network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3194–3203, 2018. 2
- [42] Jing Zhang, Yang Cao, and et al. Fast haze removal for nighttime image using maximum reflectance prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7016–7024, 2017. 7
- [43] Jing Zhang, Yang Cao, and et al. Nighttime dehazing with a synthetic benchmark. In *Proceedings of the 28th ACM International Conference on Multimedia*, pages 2355–2363, 2020. 7
- [44] Yafei Zhang, Shen Zhou, and Huafeng Li. Depth information assisted collaborative mutual promotion network for single image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2846–2855, 2024. 2
- [45] Dian Zheng, Xiao-Ming Wu, Shuzhou Yang, Jian Zhang, Jian-Fang Hu, and Wei-Shi Zheng. Selective hourglass mapping for universal image restoration based on diffusion model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024. 7
- [46] Yu Zheng, Jiahui Zhan, and et al. Curricular contrastive regularization for physics-aware single image dehazing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, . 1, 6