LPATR-Net: Learnable Piecewise Affine Transformation Regression Assisted Data-Driven Dehazing Framework

Yuelong Li, Fei Chen, Zhenwei Liu, Tianyu Zang, and Jianming Wang

Abstract*—Nowadays, data-driven learning based deep neural network (DNN) is the most dominant SOTA image dehazing framework. Here, learning to perfectly simulate the underlying mapping rules (from hazy to clear) told by massive paired training data is its core driving force. However, under genuine scenarios, it is extremely hard to guarantee the 100% qualification of all collected ground truth (GT) haze-free data. That's because natural weather is hardly controlled, and many weathers are actually in a chaotic status existing between foggy and fog-free. Thus, unlike most supervised learning issues, the image dehazing society is born with the torture of part of faulty ground truth no-haze samples. Therefore, totally trusting training data and solely pursuing more fitting powerful data-driven model may not be a wise solution. To cope with this thorny challenge, in this paper, instead of faithfully pursuing for fitting capacity promotion, we on the contrary choose to intentionally cut down the fitting flexibility to achieve higher-level robustness. That is the LPATR-Net, a novel dehazing framework specially armed with fitting power suppression mechanism to resist intrinsic annoying faulty GT. This solution does not involve any extra manually labeling. Specifically, the LPATR-Net architecture is created completely around elaborately designed fitting-restrained learnable piecewise affine transformation regression. Since such low-order linear regression structure genetically can only fit for majority of data, the interference of minority of unqualified GT samples is expected to be effectively suppressed. Through further coupled with a highly customized multi-concerns high-accuracy dehazing fitting companion component, All-Mattering, proposed LPATR-Net elegantly achieves the seamless integration of traditional majority determining fixed-form regression and modern all freedom data-driven deep learning. Extensive experiments have been conducted on five commonly utilized public datasets to verify its effectiveness. In addition, the wide-range transplantability of the proposed core regression structure has also been experimentally confirmed. Source code is available at *https://github.com/FeiChen829/LPATR-Net

***Index Terms*—Dehazing, faulty GT, fitting freedom decreasing, merging regression and DNN, affine transformation regression.**

I. Introduction

HAZE is a common natural weather phenomenon, which may be occasionally romantic but definitely all the time dangerous. It is well-known that haze may seriously degrade human eyes' visibility, which is the exclusively dominant sensing organ of mankind. Every year, plenty of traffic accidents occur due to this weather curse. On the other hand, haze is also a typical trouble to most vision-based automatic understanding researches, such as remote sensing, object detection, and motion tracking. Thus, image dehazing is a vital content enhancing research topic both for human sensing and intelligent machine understanding.

Due to its significance, image dehazing related researches began quite early and have been going on for a long time. At the beginning, since dehazing is a well-defined low-level image quality promotion mission, a great many of generic image enhancing techniques were transplanted into this topic, such as Histogram Equalization [1], [2], Homomorphic Filtering [3], [4], Gamma Correction [5], [6], [7], and Laplacian Filtering [8]. Besides general enhancing techniques, digging into the underlying mechanism of haze's composition and then accordingly looking for dehazing solution is also a popular feasible route. The approaches based on the dark channel prior [9], linear color attenuation prior [10], statistical color ellipsoid prior [11], and depth-edge aware prior [12], all belong to this category.

When the era of deep learning comes, this breakthrough data-driven modeling technique also fundamentally shakes the whole dehazing society. Owing to record-breaking accommodable parameter amount and various trainable deep architecture, DNNs could easily achieve accurate free-form function approximation, and hence have the potential capacity to accurately simulate any form of complex data mapping. Thus, once sufficient reliable training samples are collected, it is widely believed that deep learning is always a credible modeling tool deserved to be fully trusted. In the past few years, a large number of successful deep learning centered dehazing approaches have been worked out and achieved SOTA performance [14],

Received 20 December 2024; revised 27 August 2025 and 19 November 2025; accepted 21 November 2025. Date of publication 5 December 2025; date of current version 11 December 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 61771340 and Grant 62471333, in part by Tianjin Natural Science Foundation under Grant 24JCYBJC00310 and Grant 25JCYBJC00160, in part by the National Social Science Fund of China under Grant 25BC057, in part by the Research Project of Humanities and Social Science of Ministry of Education under Grant 23YJCZH258, and in part by China Scholarship Council under Grant 202408120075. The associate editor coordinating the review of this article and approving it for publication was Dr. Yan-Tsung Peng. *(Corresponding author: Jianming Wang.)*

Yuelong Li is with the School of Artificial Intelligence, Tiangong University, Tianjin 300387, China.

Fei Chen is with the School of Innovation, Tiangong University, Tianjin 300387, China.

Zhenwei Liu is with Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen 518000, China.

Tianyu Zang is with School of Software, Tiangong University, Tianjin 300387, China.

Jianming Wang is with Tianjin Key Laboratory of Autonomous Intelligence Technology and Systems, Tiangong University, Tianjin 300387, China (e-mail: wangjianming@tiangong.edu.cn).

This article has supplementary downloadable material available at https://doi.org/10.1109/TIP.2025.3637687, provided by the authors.

Digital Object Identifier 10.1109/TIP.2025.3637687

Fig. 1. A number of so called "ground truth (GT) no-haze images" in one of the most famous dehazing benchmark SOTS-outdoor of RESIDE [13]. Here, the left column demonstrates relatively clear GT image as reference, the middle column shows partially suspicious GT which may contain certain haze, and the right column illustrates highly unreliable GT, where evidently hazy parts could be easily observed. It should be emphasized that all above images come from "ground truth no-haze subsets" defined in the RESIDE dataset.

[15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25]. However, nothing is perfect. It should be emphasized that just as all other classical data-driven techniques, the performance of DNNs is directly restricted by the quality of training data as well. Furthermore, due to the inherent complicacy of natural weather conditions, it is almost an inescapable curse to image dehazing society.

It is well-known that haze is a natural meteorological condition that is out of human control. Furthermore, unlike rain, it is almost impossible to always exactly discriminate hazy and clear weather, because the qualitative visibility is a relatively subjective concept to which different people may have different understandings. Therefore, practically, it is extremely hard to always guarantee only pure haze-free images are collected when constructing the ground truth clear training dataset. In fact, we could easily locate a number of suspicious ground truth samples that contain certain extent of haze, even in most famous dehazing benchmark datasets, as shown in Fig. 1. Here, it could be observed that compared with the 1st column from left, images of the 2nd column may contain certain light haze. Thus we assert their reliability as haze-free ground truth data may be suspicious. While, samples of the 3rd column are hazy even more, and thus they are more unreliable.

Based on above principal analysis and sample demonstration, it could be concluded that the unreliability of at least a minority of ground truth (GT) no-haze training data is an inevitable precondition of the practical image dehazing mission.[1] In other words, dehazing is an inherent partially faulty GT research topic. Thus, within this society, it is unwise to always assume all training data are perfect.

Focusing on this challenge, in this paper, we choose to fully respect this intrinsic deficiency of dehazing GT data, and endeavour to work out a specialized counter-strategy to depress their negative affects. Following current SOTA, this work will be conducted under deep neural network framework.

[1]Though this deficiency seemingly could be relieved to certain extent through enrolling further intensive expert labeling, such delicate high-standard calibration must be time- and labor- consuming. Furthermore, appealing to human labor is definitely the last choice we want to in the era of artificial intelligence. Thus, under the practical image dehazing scenario, we did not presuppose further human intervention as a plausible solution.

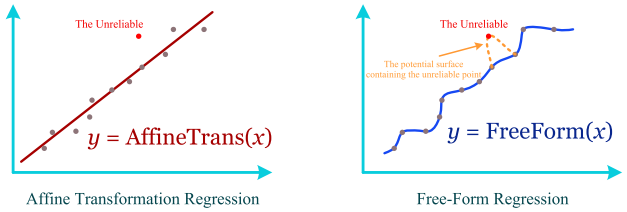

Fig. 2. A toy example illustration of the tolerance of affine transformation (left) and high flexibility free-form (right) regression to an unreliable outlier. Here, it could be observed that an isolated unreliable point may evidently alter the topology of the fitting hypersurface of free-form regression, while the influence to the majority voting affine transformation regression is nearly ignorable.

Within deep learning society, since structure-bulky DNNs theoretically have the potential to approximate arbitrary-form hypersurface, SOTA approaches get used to pursue higher-level fitting capability to each involved data to further boost model accuracy. But considering the immutable existence of faulty GT under dehazing context, this classical solution route may be questionable. A 2D toy example is shown in Fig. 2. From this figure, it could be observed that though it seems the flexible free-from regression could achieve higher accuracy global fitting, it is also more easily ruined by isolated unreliable data. That's because structurally it has sufficient flexibility to pay exact attention to each individual point. While, on the contrary, subject to the intrinsic restriction of low-order fixed-form fitting hypersurface, affine transformation regression could only focus on the majority of training data, and hence it is less effected by few unreliable outliers. Therefore, moderately restraining excessive modeling flexibility should be preferable for the practical GT-impure dehazing mission.

Inspired by above observation and analysis, in this paper, we propose a general robust image dehazing network framework, LPATR-Net, which could effectively suppresses the interference of intrinsic faulty GT through enrolling fixed-form affine transformation regression. The core underlying idea of LPATR-Net is that opposite to traditional flexibility promotion strategy, we intentionally decrease the freedom of training data fitting to counteract the traps of minority unreliable ground truth data. As to the specific implementation structure, since the low-level image dehazing task is microscopic-various, we introduce all-round elaborate piecewise linear affine transformation regression with all parameters automatically learnable, so as to achieve the subtle balance between representation accuracy and outlier tolerance.

Here, in order to preliminarily verify our core assumption, and meanwhile test the effectiveness of proposed learnable piecewise linear structure, we establish a structurally straightforward primary network, Naive LPATR.[2] It is constituted only of an isolated affine transformation regression component (namely the proposed core component, Learnable Piecewise Affine Transformation Regression, LPATR) accompanied by minimal auxiliary layers, so that all earnings can be directly attributed to proposed regression structure. We implemented the Naive LPATR[3] on two well-known dehazing landmarks SOTS-outdoor [13]

[2]The specific architecture of Naive LPATR is demonstrated in Section I of the attached Supplementary.

[3]Specific experimental configurations are provided in the experiment section.

TABLE I
THE DEHAZING PERFORMANCE OF NAIVE LPATR UNDER NO-REFERENCE CRITERION

Dataset	SOTS-outdoor [13]			Haze4K [20]		
Criterion	NIQE(↓)	deepIQA(↑)	BRISQUE(↑)	NIQE(↓)	deepIQA(↑)	BRISQUE(↑)
Ground Truth	2.9588	18.9359	0.5040	4.0352	**22.3500**	0.4980
Naive LPATR	**2.9009**	**21.3931**	**0.5160**	**3.9692**	18.8691	**0.5082**

and Haze4K [20], which of course unavoidably contain partial faulty GT data. It could be observed from Table I[4] that according to several no-reference criteria,[5] the overall quality of the dehazed acquired through Naive LPATR is even better than that of corresponding ground truth. This is an objective evidence that proposed learnable affine regression structure could effectively overcome the disturbance of faulty GT and achieve higher genuine accuracy, rather than just simply learning to blindly approximate all GT. Thus, based on this structure, we could imagine better performance than what raw GT could tell us.

Meanwhile, it should also be emphasized that in order to achieve overall high-quality robust dehazing, only a sole faulty GT resisting component is clearly not enough. High-accuracy overall fitting is also an essential precondition for fine-grained (ultra-high dimensional data oriented) dehazing mission, while it is hardly achievable through fitting restrained affine regression. Therefore, proposed LPATR-Net is actually a two-component based hybrid dehazing framework. Specifically, one is the LFFT (Learnable Fitting Flexibility Tightening component), a novel substructure constructed mainly around the proposed LPATR, responsible for faulty GT resisting; the other is All-Mattering, a Multi-Concerns Toward High-Accuracy Dehazing Fitting component, devoting to visual-level precise overall data fitting, which is specially customized to compensate for what LPATR is not good at. Through elaborately created overview architecture, LPATR-Net effectively achieves the organic merging of fixed-form low-order regression centered faulty GT resistance and high-accuracy data-driven free-form dehazing deep learning.

The main contributions of this paper could be summarized as follows,

- Formally cope with the habitually-overlooked but inherently existing imperfect GT data issue under practical image dehazing scenario.
- Propose to intentionally decrease DNN model's fitting freedom to counteract the interruption of defective ground truth dehazing data.
- Work out a novel hybird dehazing framework which organically merges traditional fixed-form regression intuition into modern big data motivated high-accuracy fitting DNN system.
- Create a generic faulty GT resisting component LFFT, which could be directly transplanted into other deep dehazing fitting architectures to promote general performance.
- Evidently upgrade dehazing robustness without incurring any extra manual labeling burden.

[4]Refer to the Supplementary (Figure 1) for corresponding visual qualitative illustration.

[5]Since there is no genuine GT corresponding to these faulty GT in the dataset, the commonly adopted referenced criteria could not be used here.

The following content of this paper is organized as follows: Section II introduces related work within dehazing society; the detailed methodology of proposed LPATR-Net is discussed in Section III; systematically conducted verification experiments are presented in Section IV; Section V summarizes our work.

II. RELATED WORK

Image dehazing is a long-standing research topic of content enhancing field, within image processing and computer vision society. In the past few decades, related researches developed fast and a large number of talent image dehazing approaches have been worked out. Generally, they could be mainly divided into three categories, namely, the extensions of general image enhancing techniques, the approaches motivated by physical haze composition model, and the rising series centered around powerful data-driven DNNs. In this section, we will give a brief introduction to each of them.

A. Extensions of General Image Enhancing

In essence, visual dehazing belongs to the broad category of low-level image enhancing, and hence transplanting the idea of general enhancing techniques is the most straightforward solution to realize haze removal. Such approaches used to be quite popular. Histogram Equalization [1], [2] suppresses the hazy effects by redistributing the overall intensity. Through separating the illumination and reflectance components in the frequency domain, Homomorphic Filtering [3], [4] improves the quality of hazy images. Bilateral Filter [26], [27] realizes hazy image enhancing based on estimating the clear pixel value in view of both spatial and intensity similarities. The intrinsic connection between gamma correction and the traditional atmospheric scattering model is highlighted in [5], where a key parameter in the dehazing process is sought by fitting a one-dimensional function. Fan et al. proposed a dehazing method based on the Retinex algorithm [28], where both the HSV color and RGB spatial details are improved.

B. Physical Haze Composition Model Based Dehazing

Haze is a common meteorological condition, and hence it is always meaningful to explore its underling physical composition mechanism. Accordingly, hand crafted physical haze constitution model also becomes a critical foundation for many effective dehazing approaches. DCP [9] proposes to combine the low-intensity color channels with the atmospheric scattering model to restore high-quality haze-free images. CAP [10] believes the combination between depth map and the atmospheric scattering model is helpful to remove haze from a single image. Berman et al. introduced a non-local prior based dehazing method [29]. DCPDN [30] effectively integrates the atmospheric scattering model into a deep learning network. In [31], a random forest is used with various prior features as the input to estimate the transmission map. A polarization-based low-illumination enhancement model is constructed to remove haze and improve image brightness [32].

C. Data-Driven DNN Centered Dehazing

In recent years, the DNN centered data-driven learning based dehazing techniques attract great attention and significantly promote the SOTA level. Thus, currently, this is the

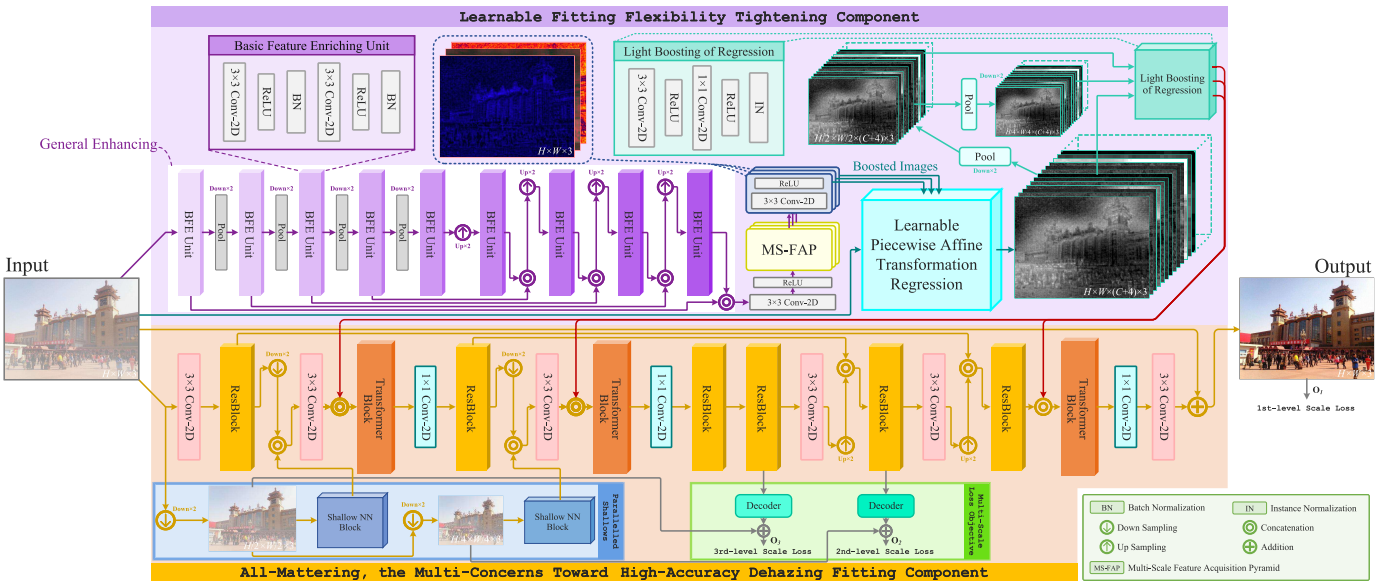

Fig. 3. The overview of proposed LPATR-Net, which is mainly composed of two modules: the learnable fitting flexibility tightening component (Top), and the multi-concerns toward high-accuracy dehazing fitting component, All-Mattering (Bottom). Here, the former is mainly in charge of depressing model's flexibility to resist the disturbance of inherent faulty GT data, while the latter compensatively focuses on high-accuracy data-driven visual dehazing fitting to boost the entire model's overall fitting capacity. In addition, as to the detailed structure of the mentioned MS-FAP (Multi-Scale Feature Acquisition Pyramid) block, which locates in the middle of the top component part, please refers to Fig. 4.

most popular research series, as to general image dehazing. GridDehazeNet [33] introduces a novel attention-based multi-scale estimation method. FFANet [34] designs an innovative feature attention module that integrates both channel attention and pixel attention mechanisms. Zheng et al. proposed a multi-guided bilateral learning based ultra-high-definition image dehazing method [35]. DeHamer [36] fully integrates both CNN and Transformer for image dehazing. Song et al. imported the Swin Transformer into image dehazing task [15]. Qiu et al. worked out the Transformer-based multi-branch dehazing network MB-TaylorFormer [16]. DCMPNet [14] enrolls depth estimation for dehazing through a dual-task interaction mechanism. DEA-Net [37] includes a detail-enhanced attention block composed of detail-enhanced convolutions and content-guided attention. Besides the networks specifically designed for dehazing task, a number of general data-driven image restoration models are also able to be utilized to remove haze, such as FocalNet [38], FSNet [39], and DSANet [40].

D. Works Directly Against Faulty GT

According to our literature investigation, though in the past few decades a great many of effective dehazing approaches have been worked out, those intentionally focusing on and committing to the proper handling of flawed ground truth samples are relatively rare. But as we expounded in Section I, this intrinsic challenge indeed deserves attention. Thus, in this paper, we choose to formally face this issue and endeavour to work out plausible robust coping solution under the prerequisite of not incurring extra human labeling burden.

III. Method

As mentioned above, proposed LPATR-Net is specifically designed for practical dehazing scenario involving unavoidable faulty ground truth samples. Unlike traditional data-driven flexibility keep-expanding solutions, it is a moderately weaker fitting-flexibility oriented fault-resisting dehazing framework. In this section, we will present its detailed network architecture, critical components (including theoretical foundation), and specific learning objective.

A. Overall Architecture

As introduced in the beginning, the core motivation of LPATR-Net is twofold: for one thing, it should be moderately restrained in general fitting freedom to resist imperfect training GT; while for another, the high-performance overall fitting capability needs to be fully pursued to be competent in the fine-grained ultra-high-dimensional visual dehazing mission. However, clearly, above both sides are actually in conflict with each other to certain extent. Thus, to guarantee their mutually compatible for better co-prosperity, we design a two components characterized hybird topology architecture, highlighted by independent function modeling plus dense-communication, as shown in Fig. 3. Specifically, the two core functional components are respectively: the Learnable Fitting Flexibility Tightening Component (Top), and All-Mattering, the Multi-Concerns Toward High-Accuracy Dehazing Fitting Component (Bottom). Here, the former is armed with specialized model flexibility restriction mechanism to suppress the disturbance of faulty GT; the latter pursues for higher-level multi-concerns general visual fitting capacity promotion (in nature it is a component specially customized to compensate for what the restrained fitting component not good at, so as to achieve evident overall performance improvement). Their specific introductions are as follows:

1) Learnable Fitting Flexibility Tightening Component: The proposed Learnable Fitting Flexibility Tightening (LFFT) is a data-driven parameter learnable component specifically created to moderately restrain the overall model's fitting flexibility. As shown in the upper part of Fig. 3, it is a

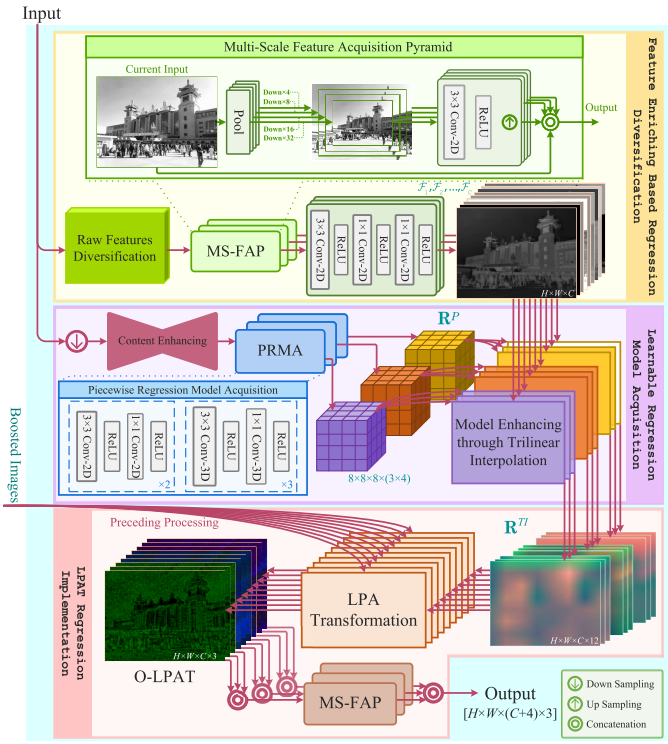

Fig. 4. The detailed structure of the critical Learnable Piecewise Affine Transformation Regression (LPATR) component.

subarchitecture relatively independent of the below high-accuracy dehazing fitting counterpart. But both components are densely connected to realize sufficient communication and co-prosperity.

Structurally, this LFFT component is mainly comprised by four modules: General Enhancing (GE), Multi-Scale Feature Acquisition Pyramid (MS-FAP), Learnable Piecewise Affine Transformation Regression (LPATR), and Light Boosting of Regression (LBR). Among them, LPATR is the core functional composition devoted to fixed-form affine regression centered fitting flexibility restriction (refer to Section III-B for detailed introduction). While, the left three are essential auxiliaries.

General Enhancing is a multi-level hierarchical source input enhancing module. It is worked out and deployed because fixed-form regression is a relatively junior fitting model and may easily be degraded by the fluctuation of input image's quality. While, attaching a front-placed content enhancing companion could beneficially relieve this vulnerability. Structurally, GE is a four scale levels comprehensive feature enhancing system. This refined hierarchy effectively ensures the full cover to all visual knowledge levels, which is critical to all-round visual knowledge boosting. In more detail, the enhancing process is specifically implemented through the Basic Feature Enriching Unit (BFE Unit). It is a two small-size convolution layers centered shallow net structure, as shown in the top left of Fig. 3. Compared with single layer counterpart, BFE Unit efficiently achieves wider receptive field.

Following GE, the feature pyramid exploration module MS-FAP is introduced to physically construct a multi-level knowledge combination complex, in order to guarantee the even emphasis to each information level. As shown in the top of Fig. 4, MS-FAP systematically integrates five knowledge levels (down to minimum $\times 32$). Here all downsampled knowledge are slightly strengthened by following 3×3 convolution.

Then, as shown in the top right corner of Fig. 3, the last LBR module is mainly charge of the preliminary boosting of acquired regression knowledge. It is a shallow two consecutive convolution layers structure, which respectively corresponds to the kernel size 3×3 and 1×1.

Finally, as demonstrated through the purple color route lines (middle of Fig. 3), the acquired affine regression based dehazing knowledge are densely merged into the high-precision fitting purpose All-Mattering component, under various scales and stages. Specifically, this knowledge injection includes three scale-levels, and involves both the encoding and decoding stage. We encourage such diversified dense communication to sufficiently fulfill affine regression's flexibility restraining role.

2) All-Mattering, the Multi-Concerns Toward High-Accuracy Dehazing Fitting Component: Image dehazing is a pixel-level microscopic content restoration mission facing complex high-dimensional image data, and hence theoretically requires a powerful high-accuracy overall dehazing data mapping modeling mechanism. What's more, as mentioned above, proposed LPATR will intentionally downgrade the overall model's fitting capacity. Therefore, a more powerful dehazing fitting framework must be established as a solid compensation and key functional partner of the flexibility-restrained LPATR.

Towards this goal, we customize a multi-aspects all mattering high-accuracy dehazing fitting subarchitecture, All-Mattering, as shown in the lower part of Fig. 3. It is well-known that modeling the mapping of sophisticated visual data is tough, because besides plain ultra-high-dimensional data matrix, a number of various multi-view quality evaluation criteria must be simultaneously satisfied either. While it is clearly extremely hard for the intrinsically fixed-form affine regression component to accommodate these requirements. Thus, as a competent companion compensator, proposed All-Mattering is genetically designed with physical multi-concerns to these diversified key quality evaluation aspects,[6] so as to achieve the full-range high-precision fitting within the complicated visual dehazing scenario. Compared with traditional single or few concerns approaches, this comprehensive multi-aspect concerns framework cares much more quality evaluation views, such as how the image knowledge levels are covered (*Knowledge-Hierarchy*), besides space how the effects are on frequency domain (*Frequency-Domain*), the robustness to various mission difficulty levels: from tough to easy handled candidates (*Difficulty-Robustness*), the full coverage capacity to both global and local scope features (*Feature-Scope*), and so forth. As to the specific implementation, all of above concerns are intrinsically embedded into All-Mattering's framework through the ingenious integration of a series of elaborately designed functional-substructures, *i.e.,*

i. *Knowledge-Hierarchy*: Constructing a multiple resolution hierarchy system based overall analysis and modeling global framework (for multi-level knowledge restoration);

[6]It is named All-Mattering, because under this framework, it is believed that all aspects of dehazing quality evaluation matter to the final effects, and hence we pursue the simultaneous satisfaction to all of them.

ii. *Frequency-Domain*: Equipped with specialized frequency domain feature analysis and recovery module (for the equal attention to both the spatial and frequency domain);

iii. *Difficulty-Robustness*: Possessing parallelled shallow and deep exploration pipelines (for effectively improving the robustness to diverse mission difficulty levels, namely the full range cover from tough to easy candidates);

iv. *Feature-Scope*: The co-inclusion of both local-info focused convolution operator and global-info preferred long-range Transformer block (for sufficiently covering various-scope features).

Corresponding detailed introductions are as follows:

i. Though image dehazing is a low-level microscopic content restoration mission, clearly, the simultaneous maintenance of higher-level semantics and knowledge is also critical to a fully competent approach. Therefore, as shown in the lower half of Fig. 3, we establish an overall scale shrinking and then expanding multi-scale framework to achieve multiple resolution knowledge analysis and dehazing restoration. Under this fundamental framework, several critical components (convolution, ResBlock, and Transformer) are deliberately deployed to systematically realize the coinstantaneous concern to multiple various quality evaluation aspects. Here, various conv-kernel size (3×3 and 1×1) is enrolled to match and assist the resolution changes. In addition, skip connections are densely arranged to avoid high-resolution info loss and facilitate data propagation.

ii. The theoretical foundation sound frequency domain analysis is a critical view to evaluate and understand visual image. Compared with pixel-level spatial domain, it provides massive global-range visual characteristics parsing. Thus besides intuitive spatial domain, a qualified dehazing image should also be excellent in counterpart frequency domain, namely

$$Qualified\_Domain = \text{Combine}(Spatial, Frequency). \quad (1)$$

According to this motivation, we intentionally deploy a series (six in all) of frequency domain specialized info extraction and restoration blocks, ResBlock. Inspired by the works of Cui et al. [39], this block explicitly extracts and processes various frequency subbands to achieve more delicate frequency feature rectification.

iii. It is well-known that deep neural networks are easily "over-armed", *i.e.,* the bulky structure is too complicated to some relatively simple tasks, for example under dehazing scenario, the light mist. Therefore, besides the mainstream deep exploration pipeline, we also assign a paralleled shallow counterpart as a light mission supplement, as shown in the left bottom of Fig. 3. Specifically, it is composed of two consecutive Shallow NN blocks, respectively matching 1/2 and 1/4 resolution, each of which contains two various field convolution and ReLU layers. From a global perspective, this light pipeline is in fact an intermediate route between full deep exploration and direct short cut (skip connections), namely,

$$Pipeline = \text{Integrate}(Deep\ NN // ShallowNN // Direct). \quad (2)$$

Clearly, this comprehensive three-in-one fitting pipeline structure is more beneficial to fully cover the various unpredictable mission difficulty levels.

iv. In general, based on the involved receptive range, image features could be roughly divided into two categories, the global and the local. Clearly, a high-quality image should be excellent on the features of both scope types, namely,

$$Covered\_Feature\_Scopes = Global + Local. \quad (3)$$

Hence, to fully respect each feature category, we arrange an independent description plus alternating deployment based hybrid feature extraction and analysis structure. Here both feature classes are respectively modeled through specialized physical structure units according to their own characteristics. In detail, the small-size convolution operator is deployed for modeling local range detail features, while a customized long-range feature description unit, the Transformer Block introduced by Zhao et al. [41] is enrolled for that of global. As shown in the bottom of Fig. 3, both units are alternating arrayed. Since the modeling of various refined local details is relatively more complex than that of long-range global, both units are unbalanced deployed with the quantity ratio 9:3.

B. Learnable Piecewise Affine Transformation Regression

In order to intentionally suppress data-driven dehazing model's fitting flexibility for better faulty GT resistance, we choose to specifically create and structurally embed (through dense communication) a novel learnable linear regression centered fitting degenerating mechanism/substructure, LPATR.

1) Theoretical Foundation and Methodology: It is well-known that linear regression may be the weakest fitting model due to the fixed-form straight line fitting shape (hyperplane). Hence, it is an ideal tool to degrade a comprehensive model's overall fitting flexibility.

While, on the other hand, in McCartney's well-known atmospheric scattering model [42] for haze decomposition, it is defined

$$\mathbf{I}(x) = J(x) \cdot t(x) + A \cdot [1 - t(x)], \quad (4)$$

where $\mathbf{x} = [x_h, x_w]^T$ denotes pixel location, $\mathbf{I}(\mathbf{x})$ represents hazy image, $\mathbf{J}(\mathbf{x})$ is the corresponding haze-free image, $\mathbf{t}(\mathbf{x})$ indicates scene transmission, and $\mathbf{A}$ is the global atmospheric light. In this classical model, there is a clear pixel-level linear relationship between the hazy and clear image. Therefore, linear regression should be effective in modeling image dehazing, as least under above mentioned ideal circumstance.

Therefore, according to the above twofold reasons, linear form regression fundamentally satisfies our intention for flexibility restrained dehazing. Hence, we specifically choose affine transformation (AT), the most typical linear regression modality, as our core foundation of restricted dehazing fitting, namely,

$$\mathbf{I}^{AT}(x) = \Omega(x) \cdot I(x) + b(x) = R^0(x) \cdot [I^T(x), 1]^T, \quad (5)$$

where under the common true color image configurations, there is RGB-channel image pixels $\mathbf{I}^{AT}(\mathbf{x}), \mathbf{I}(\mathbf{x}) \in \mathbb{R}^{3\times 1}$, and $\mathbf{R}^0(\mathbf{x}) = [\mathbf{\Omega}(\mathbf{x}), \mathbf{b}(\mathbf{x})] \in \mathbb{R}^{3\times 4}$ indicates the full parameters of affine transformation.

On the other hand, considering the widely-exsiting local similarity on visual image, the extremely inefficient pixel-level

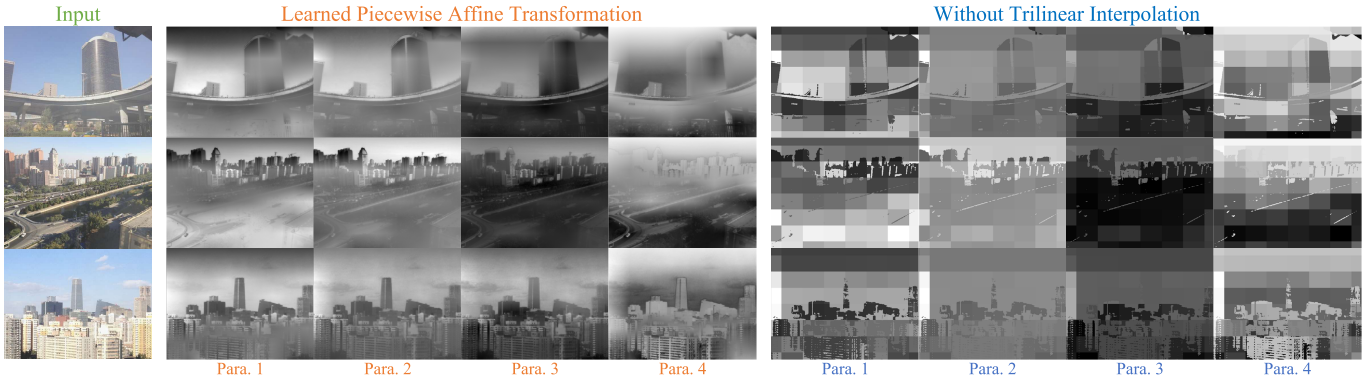

Fig. 5. The demonstration of a group of learned piecewise affine transformations (middle column group), where each column respectively corresponds to one of the four affine transformation parameters (Para.). Since here the visible appearance difference among each of the three RGB channel is trivial, only one channel is presented for conciseness. In addition, the feature number C is set as 1. Thus, the demonstrated transformation $\mathbf{R}^{TI}$ is of size $H \times W \times 4$ and could be illustrated in four images. Besides, we also exhibit the corresponding affine transformation acquired without the boundary artifacts resisting trilinear interpolation in the right side column group as comparison.

ultra-high-density regression modeling is evidently unnecessary. Thus, we further evolve above individual pixel regression model into that of three-dimensional piecewise alternate, *i.e.,*

$$\mathbf{I}^{PAT}(x) = \underbrace{\mathbf{R}^{P}\left[\tau_h(\mathbf{x}), \tau_w(\mathbf{x}), \nu(\mathcal{F}(\mathbf{I}(\mathbf{x})))\right]}_{\text{3D Piecewise Regression}} \cdot \left[I^{T}(x), 1\right]^{T}, \quad (6)$$

where $\tau(\cdot) = [\tau_h, \tau_w]$ is the piece locating function that maps 2D pixel location into their corresponding quantified piece, $\left\{\mathcal{F}\middle|\mathbb{R}^{3\times 1} \mapsto \mathbb{R}^{1}\right\}$ acts as a guiding image feature acquisition function normalized to the range $[0 \sim 1]$, and function $\nu(\cdot)$ is charge of the piecewise quantization mapping on the feature value range dimension [43]. In our specific settings, we let piecewise regression parameter matrix $\mathbf{R}^{P} \in \mathbb{R}^{\mathbf{8}\times\mathbf{8}\times\mathbf{8}\times(\mathbf{3}\times\mathbf{4})}$, the image size is $[H, W]$, and hence there is

$$\tau\middle|[1:H, 1:W] \mapsto [1:8, 1:8], \quad (7)$$

$$\nu\middle|[0 \sim 1] \mapsto [1:8]. \quad (8)$$

Here, compared with classical sole location 2D quantization, the enrollment of an extra feature value piecewise quantization evidently promotes the general fitting performance with quite limited computing overheads. In addition, it is worth mentioning that through diversifying the feature extractor $\mathcal{F}(\cdot)$, various regression equations/models could be obtained from the parameter matrix $\mathbf{R}^{P}\left[\tau_h(\mathbf{x}), \tau_w(\mathbf{x}), \nu(\mathcal{F}(\mathbf{I}(\mathbf{x})))\right]$. This kind of regression mapping extension could directly increase the fitting diversity and thus effectively enrich the overall model's representation ability.

But it should also be emphasized that, under image processing scenario, piecewise model may unavoidably incur certain visual gaps at piece boundaries, and hence degrade the overall processing quality. Even though the mentioned extra value-domain quantization may partly relieve this negative effect through simultaneously focusing on both space and value dimension, generally the grid boundary artifacts are hardly totally eliminated. The right column group of Fig. 5 in the experiment section (Section IV-C) gives a straightforward visual demonstration.

Therefore, to suppress such artifacts, instead of directly employing Equation (6), we choose to further evolve it to its trilinearly interpolated (*TI*) version, namely,

$$\mathbf{R}^{TI}(\mathbf{x}, \mathcal{F}) = \underset{\mathbf{x}, \mathcal{F}\to\tau_h, \tau_w, \nu}{\text{TriInterpolate}}\left[\mathbf{R}^{P}(\cdot,\cdot,\cdot)\right]. \quad (9)$$

Here, since linear interpolation is naturally equipped with the weighted sum operator among neighboring area grids, in theory the boundary artifacts could be effectively alleviated. The illustration shown in the middle column group of Fig. 5 verified this assertion (through comparing with above-mentioned that of the right side column group).

At last, in order to conveniently obtain these regression parameters and at the same time seamlessly merge into the entire dehazing fitting framework, above mentioned piecewise regression model will be totally learned and implemented through data-driven deep neural network structure, as introduced follows.

2) Network Implementation: As mentioned in the beginning, DNN architecture is currently the overwhelming SOTA within dehazing society. Following this track, what we want in this paper is a highly-integrated end-to-end robust deep neural network dehazing architecture. Thus, all above-mentioned linear regression dehazing methodology will be specifically implemented through transforming into corresponding network structure, the Learnable Piecewise Affine Transformation Regression (LPATR) component, as shown in Fig. 4. As shown in the figure, in structure, LPATR could be roughly divided into three parts (from top to bottom): the Feature Enriching Based Regression Diversification, Learnable Regression Model Acquisition, and LPAT Regression Implementation.

Specifically, the first regression diversification block is intentionally constructed to sufficiently extend the candidate regression pool. As mentioned above, through continuously alternating the feature extractor $\mathcal{F}(\cdot)$, a series of various available piecewise regressions could be obtained. Benefited from this diversity augmentation, the overall model's fitting ability could be evidently promoted. Structurally, the first composition of this block is the Raw Feature Diversification unit. Here, the input raw image is unfolded in channel dimension to facilitate the exploration to more visual fine details. This unit is mainly realized through efficient pointwise depth-wise separable convolution. The second processing unit is

above-mentioned (Paragraph IV of Page 5, Section III-A.1) MS-FAP, as shown in the top of Fig. 4. It is mainly responsible for scale-level feature extension (in pyramid hierarchy). At last, we also attach a shallow feature analysis unit composed of three sequential convolution-ReLU layers. The final output of this block is a set of diversified features, denoted as $\Gamma = \{\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_C\}$.

The middle regression acquisition block of Fig. 4 is the core of LPATR component, where all pre-designed dehazing regressions are totally settled down through data-driven automatic learning. The main learning course is conducted on down-sampled images to cut down computing burden. In structure, as shown in the figure, this block mainly consists of four compositions: Downsampling, Content Enhancing, PRMA, and Model Enhancing through Trilinearly Interpolation. The downsampling process is placed ahead of all to shrink the overall data scale. The following Content Enhancing structure is enrolled to preliminarily enhance raw data and moderately alleviate the quality degradation incurred by former downsampling. This part is implemented based on a U-Net [44] like encoding-decoding architecture. Then, as the functional core of this block, PRMA (Piecewise Regression Model Acquisition) is in charge of determining the regression model, namely figuring out the specific regression parameters. Since linear regression is a kind of relatively concise model form, here as shown in the figure, we work out a single route deep convolution based efficient model prediction architecture. Structurally, it is comprised of two sections of stacked ($\times$ 2 and $\times$ 3) convolution blocks, respectively corresponding to 2D and 3D convolution. This is an elaborate design for the subtle balance among dimension cover, generalization, and efficiency. Generally speaking, 3D convolution is powerful in both space and channel dimension cover, but relatively slow and easily incurs overfitting. While, the systematical combination of 2D and 3D convolution could evidently overcome these defects. That's because the import of space dimenson only 2D convolution is beneficial to both computing saving and generalization promotion. Meanwhile the remaining 3D convolution could keep sufficient attention to channel dimension analysis, and the fusion of space and channel dimension. Thus, this is the optimal choice considering all factors. In addition, similar with the multi-heading mechanism [45], we also stack a number of PRMAs to further diversify and boost the acquired regression model, $\mathbf{R}^P$. At last, in the trilinear interpolation based model enhancing unit, through importing diversified features Γ and executing the interpolation conducting Equation (9), the evolved fully functional regression model $\mathbf{R}^{TI}$ is acquired. It is the key to carry out proposed LPATR to boost the overall dehazing framework.

In the end, the bottom row of Fig. 4 demonstrates the execution of acquired piecewise affine regression model. This block includes two main implementation constituents. First, in the LPA Transformation unit, the collected regression $\mathbf{R}^{TI}$ is directly imposed on current data, namely the boosted images obtained through preceding processing (shown in Fig. 3). Here, this implementation process is the straightforward execution of pixel-level affine transformation (Equation (5)). Second, another stack of multi-scale feature extension unit MS-FAP is enrolled to realize scale-level knowledge enhancement. Then we could acquire the ultimate output of LPATR component in size $H \times W \times (C+4) \times 3$. As shown in the middle of Fig. 3, this output will be densely merged into the below high-accuracy dehazing fitting component, All-Mattering, to collaboratively realize desired flexibility restrained robust image dehazing.

C. Learning Objective

Image dehazing is a high-standard comprehensive content restoration mission, which pursues for perfect visual appearance under each plausible aspect of consideration. Thus, a fully qualified dehazing approach should seek to achieve full-range completely competent.

According to this goal, as to the training of proposed robust high-accuracy dehazing framework LPATR-Net, we specially construct an end-to-end multi-view plus multi-scale oriented comprehensive training objective guiding system. It is realized through the composite integration of a number of single aspect supervising objectives.

In detail, the view and scale single loss is defined as,

$$\mathcal{L}_{single}(v, s) = \mathbb{E}_{\mathbf{I}^{GT}, \mathbf{O}_s} \left\| \Psi_v \left[DS\left(\mathbf{I}^{GT}, \theta_s\right) \right] - \Psi_v(\mathbf{O}_s) \right\|_1, \quad (10)$$

where v denotes the observation view (namely feature category) ranging from microscopic pixel to macroscopic global frequency, s indicates the feature scale-level, $\Psi_v(\cdot)$ represents feature extraction operator, $DS(\cdot,\cdot)$ stands for image downsampling at designated ratio, $\mathbf{I}^{GT}$ is ground truth image, $\theta_s = 1/2^{s-1}$ denotes downsampling ratio, and $\mathbf{O}_s$ is the multi-scale phase output as annotated in the Multi-Scale Loss Objective block (plus the final output) in lower right corner of Fig. 3.

Based on each single aspect loss, the integrated overall composite objective is constructed as,

$$\mathcal{L}_{com} = \sum_{v \sim micro}^{macro} \alpha_v \sum_{s=1}^{\mathrm{S}} \beta_s \cdot \mathcal{L}_{single}(v, s), \quad (11)$$

where S indicates the number of scale levels, and α_v and β_s are the relative weights respectively specified to each view and scale level.

IV. Experiments

In this section, proposed LPATR-Net will be extensively and multi-perspectively evaluated to verify its effectiveness. The specific experiment schedule is as follows:

First, detailed experimental exploration on the nature and function of the core part of proposed approach will be conducted in Subsection C. Second, the specific quantitative and qualitative experiment performance on multiple public datasets will be discussed in Subsection D and E, respectively. Third, the constituents' necessity exploration namely ablation study will be systematically conducted in Subsection F. Fourth, since faulty GT is also inevitable in test set's construction, current experimental performance may be inaccurate due to such disturbance. Thus, in pursuit of genuine performance evaluation, in Subsection G, we will inspect LPATR-Net's achievement on a manually constructed fault-free test dataset. At last, since proposed fitting restriction core component LFFT is created relatively independent with other remaining structure, we are curious about and will explore its transplanting performance, namely when it is injected into and collaborates with other general dehazing fitting frameworks, in Subsection I.

A. Datasets

Our experiments are systematically organized and conducted on a group of publicly available datasets, including RESIDE [13], O-HAZE [46], Dense-Haze [47], I-HAZE [48], and NH-HAZE [49]. The RESIDE is a famous commonly adopted large-scale dehazing benchmark dataset comprised by five subsets. In our experiments, we specifically use the Outdoor Training Set (OTS), Indoor Training Set (ITS), and Synthetic Objective Testing Set (SOTS). The OTS is composed by 313,740 outdoor hazy images generated from 8,970 outdoor haze-free images. The ITS contains 13,990 indoor hazy images generated from 1,399 indoor images. The SOTS is split into the SOTS-outdoor and SOTS-indoor, each of which contains 500 pairs of hazy/clear images. The model is trained on OTS and ITS, and then tested correspondingly on the SOTS-ourdoor and SOTS-indoor respectively. Meanwhile, the O-HAZE [46], Dense-Haze [47], I-HAZE [48], and NH-HAZE [49] are all small-scale real haze datasets where the haze is generated by professional machine.

B. Implementation Details

Our verification experiments are conducted based on Torch 1.9.1 on NVIDIA GTX 3090 Ti, A100, and L40 GPUs. The Adam optimizer [50] is specifically adopted for training, with the value of 0.9 for β_1 and 0.999 for β_2 respectively. Our initial learning rate is set as 0.0001 with gradually descending strategy applied [51]. Proposed LPATR-Net is accompanied and will be compared with 28 newly published approaches, including AOD-Net [52], GFN [53], DM2F-Net [54], GCANet [55], GridDehazeNet [33], PFDN [56], FFA-Net [34], DW-GAN [57], DeHamer [36], PMNet [58], MAXIM-2S [59], UDN [60], D4 [61], SGID-PFF [23], SANet [62], MITNet [63], FocalNet [38], SFNet [64], DehazeFormer-B [15], FSNet [39], IRNeXt [65], MB-TaylorFormer-L [16], C^2PNet [66], DEA-Net-CR [37], DCMPNet [14], OKNet [67], EENet [68], and PoolNet-B [69].

C. Explore the Nature and Function of the Learned Piecewise Affine Transformation Regression

As mentioned in the beginning, the newly-designed piecewise affine regression component is a critical key ingredient of the proposed robust dehazing architecture. Theoretically, subject to the fixed-form fitting hyperplane and strictly restricted parameter amount, this regression module can only rely on the majority of training data and hence could inherently resist the misleading of minority of faulty GT. In order to actually unpack this core structure, in this subsection, we will in-depth explore its nature characteristics and specific function through visualized practical experimental data.

A series of practically learned piecewise affine transformations, namely the $\mathbf{R}^{TI}$ in Equation (9), are shown in Fig. 5 (middle column group). Here, each column respectively corresponds to one of the four affine parameters as presented in the row of Equation (5). From this figure, it could be observed that the acquired transformations are full of information derived from the input image, but each parameter image has its special focus. For example, the 2nd row image of Para. I focuses on remote buildings, while the corresponding one of Para. IV pays more attention in nearby scene. Through the pixel-level affine combination of these highly-relative meanwhile focus various information, theoretically it is plausible to achieve moderate image content enhancing (dehazing). Especially when these transformation parameters are all automatically obtained through high-precision data-driven based deep learning. But on the other hand, it is also worth noting that this dehazing fitting mechanism is relatively weak, due to the strictly limited parameter amount (only four) and the fixed-form low-order fitting modality (affine). Thus, in a word, this acquired system is in line with our original purpose, *i.e.,* it could achieve moderately dehazing fitting but is genetically weak in general fitting flexibility.

In addition, the right side column group of Fig. 5 provides a straightforward visualized parameter comparison, where the grid based trilinear interpolation operator introduced in Equation (9) is discarded. According to the figure, it could be found that without this deliberately deployed grid-level smoothing mechanism, the overall quality of learned transformations drops dramatically, especially in detail expression and boundary artifacts resistance. Thus, it could be asserted that the introduced extra trilinear interpolation mechanism is valuable to the acquisition of high quality regression model.

Besides the regression model's visualization, in Fig. 6, we also demonstrate the direct outcomes of applying the acquired Learnable Piecewise Affine Transformation, namely the output of the LPA Transformation unit shown in the bottom of Fig. 4, the O-LPAT matrix. In Fig. 6, it could be observed that through executing the learned affine transformation, a series of content enhancement could be achieved, *i.e.,*

- As for O-LPAT 1 (2nd Row), the global noise is effectively depressed without evidently losing content details. Meanwhile, the overall frequency is simultaneously diminished, which is beneficial for dehazing because generally the widely-distributed haze will elevate image's holistic frequency.
- As for O-LPAT 2 (3rd Row), the contrast between low-frequency composition (central zone) and that of high-frequency (surrounding zone) is moderately eased. This is a positive effect action, because global-range gradual hazy image generally has stronger low-frequency which means the low-high visual contrast will be more evident.
- As for O-LPAT 3 (bottom Row), the specific enhancing processing on each column seems prominent but diverse (in both spacial and frequency domain). For example, the 1st column evidently promotes the overall clarity, the 2nd focuses on general smoothing, and the 3rd specializes in detail restoration. This is plausible and profitable, after all, the haze removing calculation to each specific image should be exclusive and highly-customized based on its own characteristics.

According to these visualized outcomes, it could be asserted that proposed LPAT regression is effective in diversified content enhancement. This is positive for final dehazing performance promotion.

D. Quantitative Evaluation

The quantitative performance on both the SOTS-indoor and SOTS-outdoor dataset is summarized in Table II. According to the table, generally proposed LPATR-Net performs well on both indoor and outdoor hazy images. Specifically, LPATR-Net

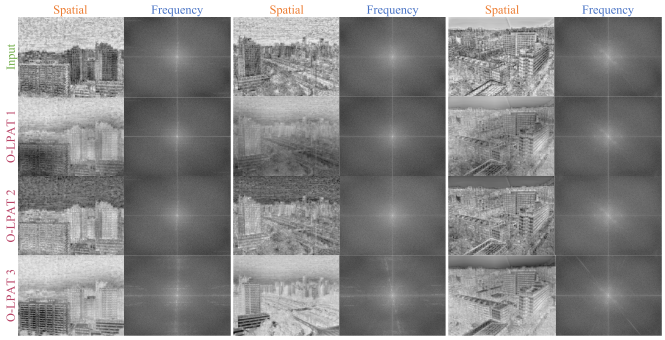

Fig. 6. The demonstration of practical outcomes of applying the acquired Learnable Piecewise Affine Transformation (LPAT). In this experiment, feature number $C = 1$, and hence the size of LPAT's output matrix (O-LPAT) is $H \times W \times 3$, which are specifically exhibited through the three image rows, from O-LPAT 1 to O-LPAT 3. In addition, the corresponding frequency domain images are also demonstrated aside for thorough analysis.

TABLE II

THE QUANTITATIVE DEHAZING PERFORMANCE ON THE SOTS-INDOOR AND SOTS-OUTDOOR DATASET

Methods	Venues	SOTS-indoor		SOTS-outdoor		FLOPs
		PSNR(↑)	SSIM(↑)	PSNR(↑)	SSIM(↑)	
AOD-Net	ICCV'17	20.51	0.816	24.14	0.920	**0.54 G**
GFN	CVPR'18	22.30	0.880	21.55	0.844	-
DM2F-Net	ICCV'19	34.29	0.984	-	-	112.19 G
GCANet	WACV'19	30.23	0.980	-	-	68.34 G
GridDehazeNet	ICCV'19	32.16	0.984	30.86	0.982	81.43 G
PFDN	ECCV'20	32.68	0.976	-	-	-
FFA-Net	AAAI'20	36.39	0.989	33.38	0.980	1251.29 G
DW-GAN	CVPRW'21	35.94	0.986	-	-	142.47 G
DeHamer	CVPR'22	36.63	0.988	35.18	0.986	214.88 G
PMNet	ECCV'22	38.41	0.990	34.74	0.985	-
MAXIM-2S	CVPR'22	38.11	0.991	34.19	0.985	-
UDN	AAAI'22	38.62	0.991	34.92	0.987	-
D4	CVPR'22	25.42	0.932	25.83	0.956	-
SGID-PFF	TIP'22	38.52	0.991	-	-	716.15 G
SANet	IJCAI'23	40.40	0.996	38.01	0.995	147.65 G
MITNet	MM'23	40.23	0.992	35.18	0.988	72.25 G
FocalNet	ICCV'23	40.82	0.996	37.71	0.995	118.91 G
SFNet	ICLR'23	41.24	0.996	40.05	0.996	538.11 G
DehazeFormer-B	TIP'23	37.84	0.994	34.95	0.984	111.15 G
FSNet	TPAMI'23	42.45	**0.997**	40.40	**0.997**	475.34 G
IRNeXt	ICML'23	41.21	0.996	39.18	0.996	182.01 G
MB-TaylorFormer-L	ICCV'23	**42.64**	0.994	38.09	0.991	368.59 G
C^{2}PNet	CVPR'23	42.56	0.995	36.68	0.990	2005.98 G
DEA-Net-CR	TIP'24	41.31	0.995	36.59	0.990	140.26 G
DCMPNet	CVPR'24	42.18	**0.997**	36.56	0.993	289.63 G
OKNet	AAAI'24	40.79	0.996	37.68	0.995	158.10 G
EENet	PR'25	42.08	**0.997**	39.10	0.996	202.82 G
PoolNet-B	TIP'25	42.01	**0.997**	-	-	273.84 G
LPATR-Net	-	42.58	**0.997**	**41.24**	**0.997**	494.97 G

achieves optimal accuracy under three of all four indicators. The only exception is the PSNR on SOTS-indoor, where LPATR-Net is slightly lower than that of MB-TaylorFormer-L (by 0.06), but still better than all others. While, the most evident promotion appears on the PSNR of SOTS-outdoor, where LPATR-Net reaches 41.24 and attains 0.84 advantage compared with the second optimal approach (FSNet). Considering the overall difficulty of SOTS-outdoor, this improvement is promising, especially when here the involved outdoor scenario is doomed to be haunted by faulty GT. On the other hand, it is sad to see LPATR-Net doesn't achieve efficiency promotion as to runtime FLOPs, but it is still comparable with several SOTA approaches, such as C^{2}PNet, FSNet, SFNet, SGID-PFF, FFA-Net, and so forth. Meanwhile, as will be demonstrated in Section IV-F (Ablation) and IV-I (Transplantability), proposed core structure could be aggressively pruned or transplanted into other general dehazing architectures to evidently depress computing burden while maintain moderate dehazing accuracy. Those will be more efficient options.

TABLE III

THE QUANTITATIVE DEHAZING PERFORMANCE ON THE O-HAZE, DENSE-HAZE, I-HAZE, AND NH-HAZE DATASET

Methods	Venues	O-HAZE		Dense-Haze		I-HAZE		NH-HAZE	
		PSNR(↑)	SSIM(↑)	PSNR(↑)	SSIM(↑)	PSNR(↑)	SSIM(↑)	PSNR(↑)	SSIM(↑)
SANet	IJCAI'23	27.55	0.951	17.25	0.703	20.81	0.867	19.73	0.792
FocalNet	ICCV'23	27.61	0.955	16.91	0.695	20.43	0.858	19.83	0.792
SFNet	ICLR'23	27.63	0.955	17.29	0.698	20.39	0.858	20.32	0.803
DehazeFormer-B	TIP'23	26.88	0.944	16.92	0.668	20.33	0.835	19.66	0.771
FSNet	TPAMI'23	27.71	0.953	17.50	**0.719**	20.15	0.864	20.25	0.806
DEA-Net-CR	TIP'24	27.52	0.953	17.10	0.699	20.68	0.867	19.90	0.789
OKNet	AAAI'24	27.30	0.951	16.75	0.673	20.48	0.866	19.21	0.781
EENet	PR'25	27.67	0.955	17.45	0.701	**21.18**	0.863	19.76	0.794
LPATR-Net	-	**27.72**	**0.956**	**17.62**	0.704	21.01	**0.876**	**20.33**	**0.809**

TABLE IV

AVERAGE RUNTIME PERFORMANCE (MS) ON THE RESIDE SOTS-OUTDOOR DATASET

SGID-PFF (TIP'22)	SFNet (ICLR'23)	FSNet (TPAMI'23)	DehazeFormer-B (TIP'23)	MB-TaylorFormer-B (ICCV'23)	C^{2}PNet (CVPR'23)	LPATR-Net
167.99	162.22	148.45	**115.20**	652.81	323.82	200.36

In addition, the general runtime performance on the RESIDE SOTS-outdoor dataset is reported in Table IV. All this series of experiments are conducted on a same hardware platform, with Intel Xeon 3204 CPU, 32 GB Memory, and single NVIDIA GTX 3090Ti with 24GB GPU. There are all 500 images involved and we report the mean runtime elapse. The image size on this dataset is not even, and the average size is 549.95×477.62. According to the table, it could be observed that though as a complicated multi-components dense parameters network structure, LPATR-Net is not good at running faster, generally its runtime cost is in the same level with a number of contemporaneous approaches. In other words, we believe the running speed is not LPATR-Net's unacceptable fatal deficiency.

E. Qualitative Evaluation

The qualitative dehazing performance on the SOTS-indoor and SOTS-outdoor dataset is demonstrated in Fig. 7, where besides input hazy and corresponding ground truth, we directly exhibit the difference image (with corresponding GT) for comparison convenience. According to these visual demonstration, it could be observed that proposed LPATR-Net effectively narrows the general difference between the obtained dehazed and corresponding GT. Lots of high-frequency details which are more easily destroyed by global-range densely distributed haze have been successfully restored. Especially on the tougher SOTS-outdoor dataset, the performance promotion compared with other SOTA is clearly more evident. Fig. 8 and 9 respectively exhibits the dehazing outcomes on the NH-HAZE

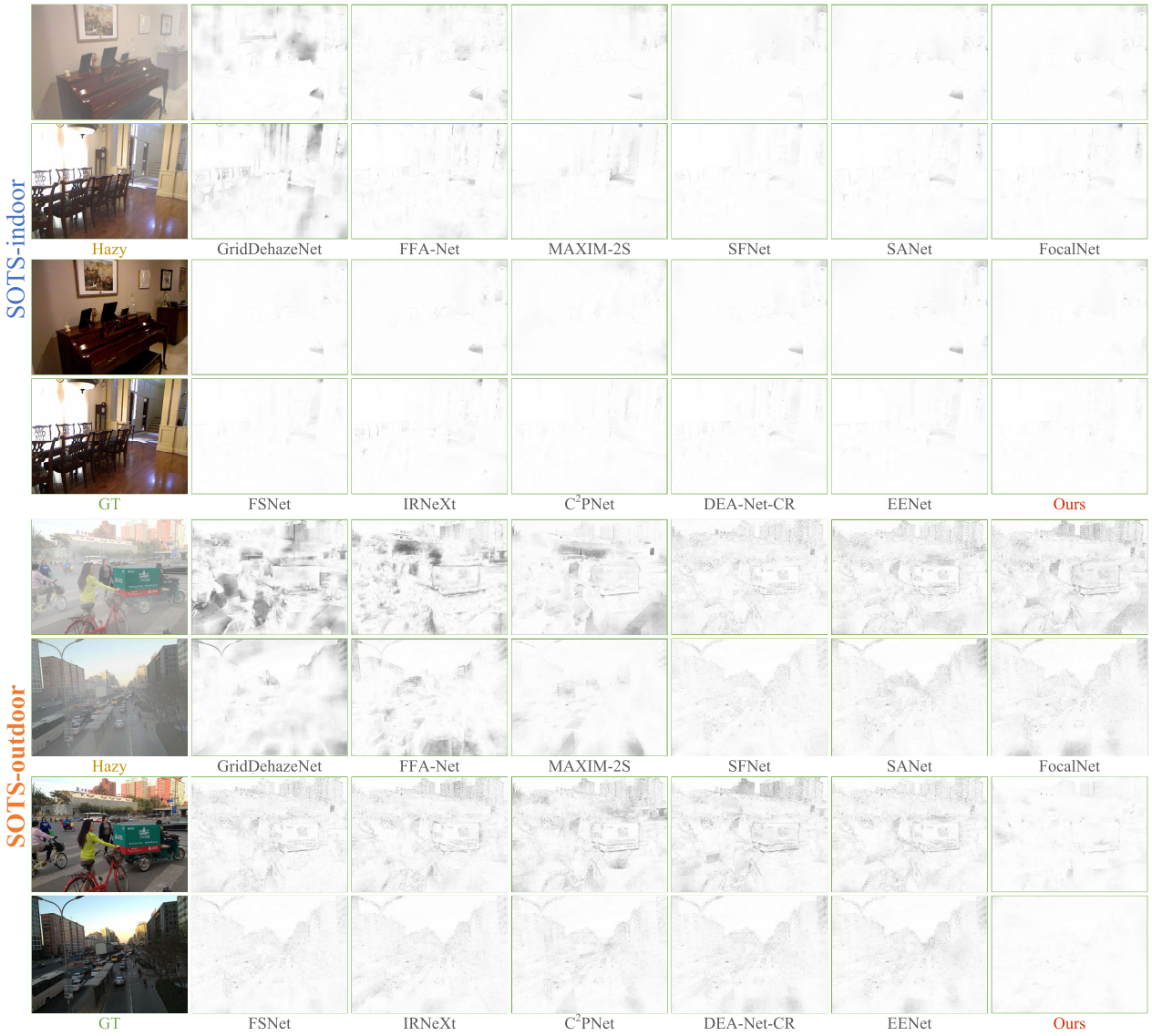

Fig. 7. The qualitative dehazing performance on the SOTS-outdoor and SOTS-indoor dataset. Here the difference image with corresponding ground truth is directly demonstrated for better visual performance evaluation, where black color denotes the difference.

and O-HAZE dataset. According to both figures, generally proposed LPATR-Net achieves promising overall dehazing performance, in details:

(1) More thorough haze removal effects, *e.g.,* the 1st sample of Fig. 8, where certain haze residuals are easily located on output of other approaches, especially in the zoomed-in patches.
(2) Better color restoration, *e.g.,* the 2nd sample of Fig. 8, especially the tile color.
(3) More similar overall or local color tone with corresponding GT, *e.g.,* samples of Fig. 9, especially the wood color tone on sample 1.
(4) Occasionally better than GT (due to the intrinsic faulty GT resisting capacity), *e.g.,* the zoomed-in region on the 1st sample of Fig. 9, where the GT seems more like a light mist image than that dehazed by LPATR-Net.

Due to the strict page amount limit, please refer to the 2nd section of the Supplementary for more qualitative performance demonstration.

F. Ablation Study

LPATR-Net is a relatively complicated structure, and hence it is necessary to carefully check the value of each of its critical elements. The detailed ablation study of proposed LPATR-Net is summarized in Table V. From the first row, it could be observed that the lack of the core LFFT component evidently degrades the overall performance, especially on PSNR of SOTS-outdoor (2.47 ↓). On the other hand, the absence of each of other involved key compositions also incurrs certain general accuracy reduction. Specifically, on the All-Mattering component, the long-range feature descriptor Transformer Block contributes the most; while, on the LFFT component, the outstanding composition reflected by each indicator varies but the general descending trend is clear. In addition, it should be noticed that through removing a certain relatively bulky ingredient, the involved runtime FLOPs could be evidently decreased with limited accuracy loss. For example, cutting down half feature channels could save 202.81G FLOPs at the expense of only 0.88 PSNR decline (to 40.36, outdoor). This accuracy still ranks the 2nd in Table II. Hence, it could

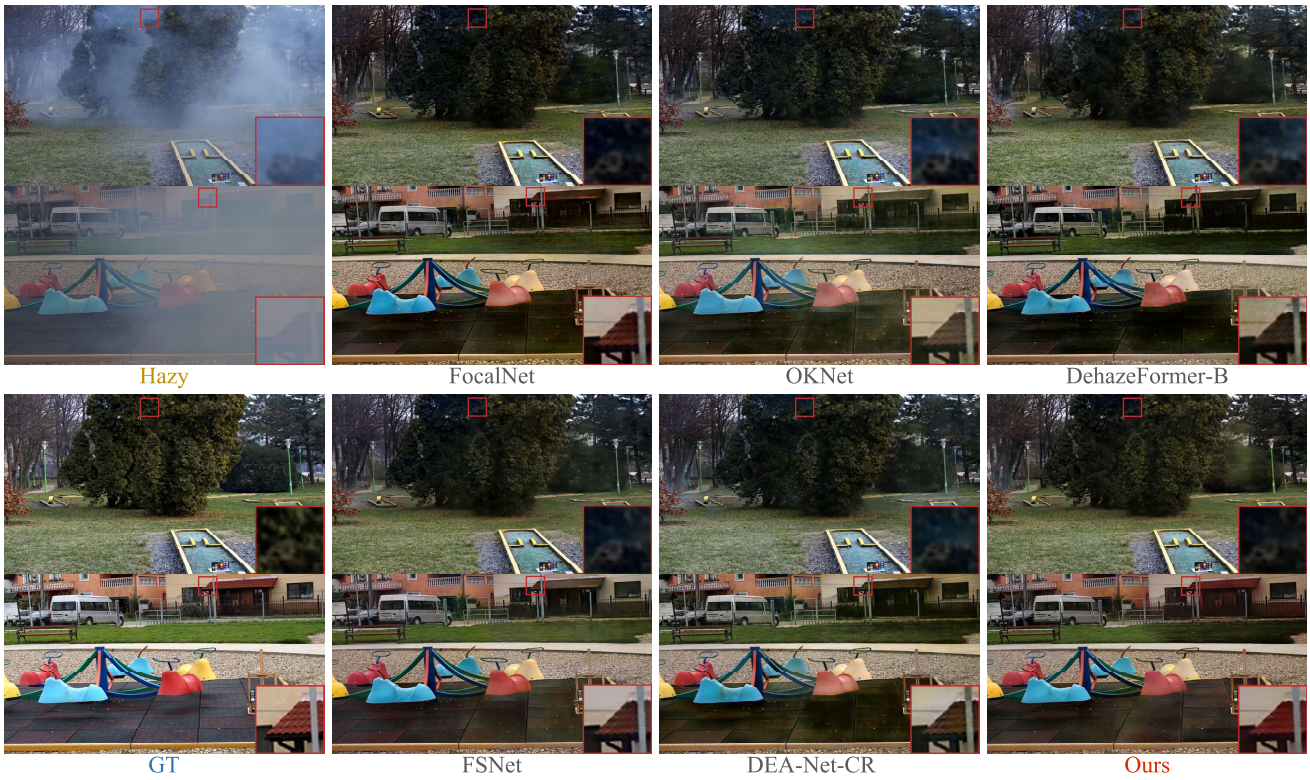

Fig. 8. The qualitative dehazing performance on the NH-HAZE dataset.

Fig. 9. The qualitative dehazing performance on the O-HAZE dataset.

be asserted that proposed LPATR-Net framework actually possesses the potential to pursue higher-level computing efficiency.

Corresponding qualitative performance on SOTS-outdoor is demonstrated in Fig. 10, where we adopt the difference image with ground truth to give more straightforward comparison. Here it could be observed that the full size LPATR-Net has evident visual advantage than any other shrank size counterparts. Thus, the value of each mentioned individual composition of LPATR-Net is verified. Meanwhile, given such evident

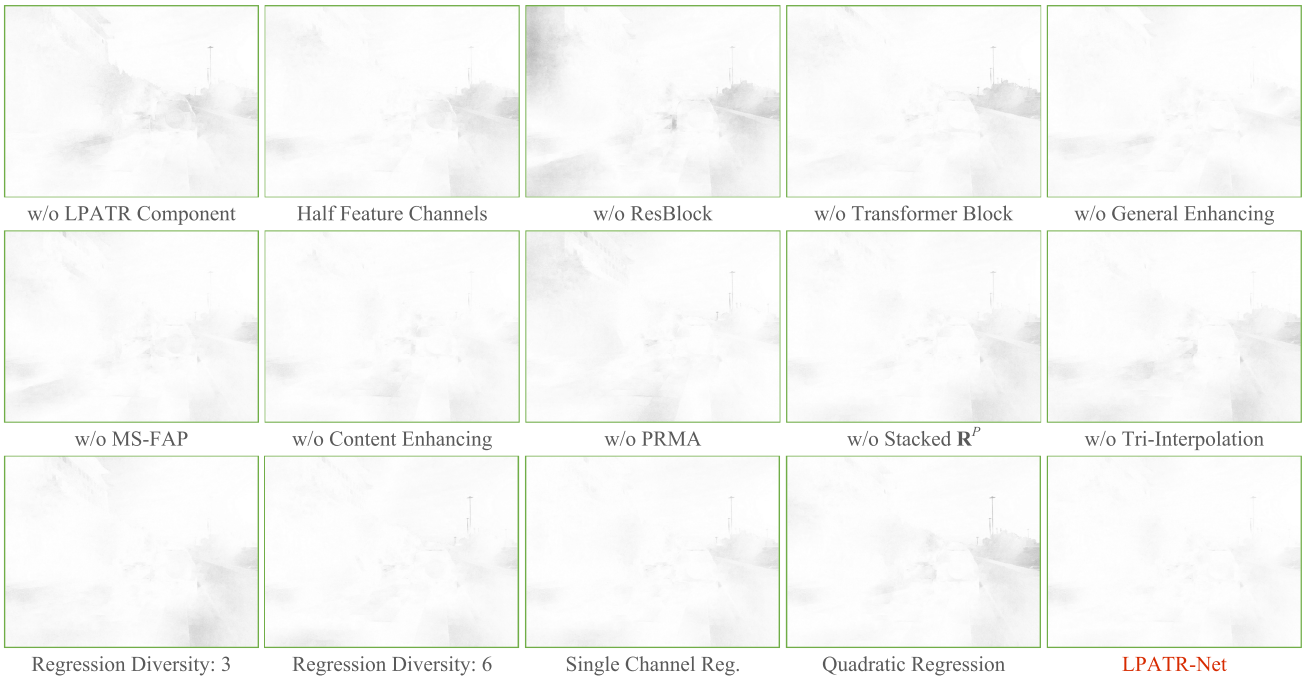

Fig. 10. The qualitative ablation study shown in the difference image with GT.

TABLE V
THE QUANTITATIVE ABLATION STUDIES

Evaluated Component	Models	SOTS-indoor PSNR(↑)	SOTS-indoor SSIM(↑)	SOTS-outdoor PSNR(↑)	SOTS-outdoor SSIM(↑)	FLOPs(↓)
All	w/o LFFT Component	39.79	0.9952	38.77	0.9960	**176.23** G
All-Mattering	Half Feature Channels	40.21	0.9962	40.36	0.9964	292.16 G
	w/o ResBlock	41.29	0.9964	39.93	0.9965	393.59 G
	w/o Transformer Block	39.49	0.9946	39.06	0.9961	363.31 G
Learnable Fitting Flexibility Tightening (LFFT)	w/o General Enhancing	41.55	0.9958	39.79	0.9966	317.45 G
	w/o MS-FAP	41.66	0.9967	39.42	0.9961	495.77 G
	w/o Content Enhancing	41.98	0.9969	40.12	0.9968	454.79 G
	w/o PRMA	41.65	0.9964	39.74	0.9965	494.61 G
	w/o Stacked $\mathbf{R}^P$	41.10	0.9965	40.54	0.9970	494.74 G
	w/o Tri-Interpolation	42.30	0.9970	39.75	0.9966	494.97 G
	Reg. Diversity: $C = 3$	41.60	0.9960	40.72	0.9971	493.92 G
	Reg. Diversity: $C = 6$	41.28	0.9965	40.83	0.9971	494.59 G
	Single Channel Reg.	42.15	0.9971	39.98	0.9966	494.90 G
	Quadratic Regression	42.43	0.9966	40.01	0.9968	495.06 G
LPATR-Net		**42.58**	**0.9972**	**41.24**	**0.9972**	494.97 G

TABLE VI
THE QUANTITATIVE DEHAZING PERFORMANCE ON A GENUINE GT TEST DATASET CONSTRUCTED BASED ON SOTS-OUTDOOR

Methods	Venue	PSNR(↑) on Raw GT	PSNR(↑) on Genuine GT
SFNet	ICLR'23	40.05	39.51 (-0.54)
FSNet	TPAMI'23	40.40	39.71 (-0.69)
MB-TaylorFormer-L	ICCV'23	38.09	38.15 (+0.06)
C^2PNet	CVPR'23	36.68	35.01 (-1.67)
DEA-Net-CR	TIP'24	36.59	36.42 (-0.17)
LPATR-Net	-	**41.24**	**41.58** (+0.34)

appearance superiority, it could be asserted that let efficiency alone, proposed full size LPATR-Net architecture is still the optimal in general dehazing performance. Removing key parts should only be considered when the runtime cost conditions are tough.

G. Performance on Genuine GT Test Dataset

As mentioned in the beginning, since natural weather is hardly controllable and accurately distinguishable, potential faulty GT samples (as shown in Fig. 1) consistently haunt image dehazing society. Here, it should be further emphasized that clearly, this GT impureness uniformly exists in both training and test data, and hence will not only affect the training process, but also the execution of test evaluation. As a consequence, current mainstream testing protocol (as we used above) may be a little bit inaccurate. Because the faulty GT in test set will provide wrong hint and hence impact the fairness of evaluation. Clearly, removing these disturbance will lead to more reasonable evaluation outcomes.

Thus, besides above standardized experiments commonly adopted within current dehazing society, we also would like to conduct an extra fairer evaluation experiment on a genuine GT test dataset as a critical complementation. In detail, here all 156 suspicious faulty GT (out of all 500) within SOTS-outdoor test set are manually filtered out to construct a pure genuine GT test dataset. The specific quantitative performance is listed in Table VI. According to this table, it could be observed that compared with original raw GT, on this genuine GT evaluation, proposed LPATR-Net gains evident performance promotion in PSNR (0.34 ↑), while most other SOTA suffer certain accuracy reduction. This is an indication that our LPATR-Net is more inclined to pursue genuine image dehazing rather than solely labeled data learning.

H. The Tolerance to Faulty Ground Truth

In order to further systematically verify the robustness of proposed approach to endure unqualified flawed ground truth samples, we specially organize a refined validation experiment conducted on gradually increased faulty GT samples.

In detail, we randomly pick up 400 pairs [hazy, clear] data from the OTS of RESIDE dataset to construct the base of training set, while 50 pairs images from the SOTS-outdoor dataset consistently comprise the test set. The series

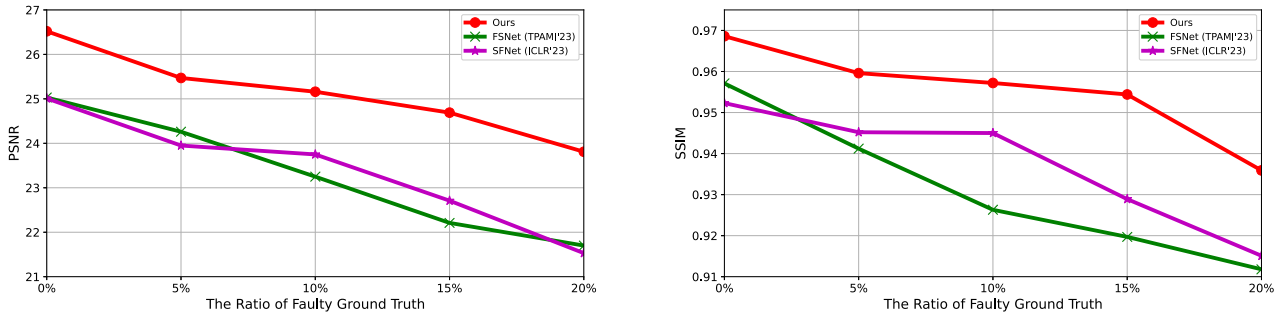

Fig. 11. The performance under incrementally polluted ground truth training datasets.

TABLE VII

THE TRANSPLANTING PERFORMANCE OF THE LFFT COMPONENT INTO FOUR GENERAL DEHAZING ARCHITECTURES

Methods	Venue	SOTS-indoor PSNR(↑)	SOTS-indoor SSIM(↑)	SOTS-outdoor PSNR(↑)	SOTS-outdoor SSIM(↑)	Param(↓)	FLOPs(↓)
SANet	ICLR'23	40.40	0.996	38.01	0.995	3.81 M	147.65 G
+ LFFT		42.66↑	0.997↑	38.21↑	0.995=	42.07 M	497.89 G
+ LFFT (w/o CE)		42.38↑	0.997↑	38.24↑	0.995=	24.80 M	323.91 G
+ LFFT (w/o GE)		42.35↑	0.997↑	38.13↑	0.995=	24.80 M	457.71 G
+ LFFT (w/o CE&GE)		42.46↑	0.997↑	38.40↑	0.996↑	7.53 M	283.73 G
DehazeFormer-B	TIP'23	37.84	0.994	34.95	0.984	2.52 M	111.15 G
+ LFFT		41.62↑	0.996↑	40.31↑	0.997↑	39.95 M	431.81 G
+ LFFT (w/o CE)		39.86↑	0.995↑	38.50↑	0.996↑	22.68 M	257.82 G
+ LFFT (w/o GE)		41.37↑	0.996↑	40.89↑	0.997↑	22.68 M	391.63 G
+ LFFT (w/o CE&GE)		40.22↑	0.996↑	41.00↑	0.997↑	5.41 M	217.64 G
OKNet	AAAI'24	40.79	0.996	37.68	0.995	4.72 M	158.10 G
+ LFFT		42.11↑	0.997↑	38.44↑	0.996↑	42.98 M	508.35 G
+ LFFT (w/o CE)		41.94↑	0.997↑	38.36↑	0.996↑	25.71 M	334.36 G
+ LFFT (w/o GE)		41.96↑	0.997↑	39.11↑	0.996↑	25.71 M	468.17 G
+ LFFT (w/o CE&GE)		42.09↑	0.997↑	38.51↑	0.996↑	8.44 M	294.18 G
EENet	PR'25	42.08	0.997	39.10	0.996	5.44 M	202.82 G
+ LFFT		42.27↑	0.997=	38.16↓	0.995↓	43.69 M	577.88 G
+ LFFT (w/o CE)		42.42↑	0.997=	38.51↓	0.996=	26.42 M	389.54 G
+ LFFT (w/o GE)		42.27↑	0.997=	38.53↓	0.995↓	26.42M	537.70 G
+ LFFT (w/o CE&GE)		42.37↑	0.997=	38.95↓	0.996=	9.15 M	349.36 G
LPATR-Net	-	42.58	0.997	41.24	0.997	42.20 M	494.97 G

of experiments is consisted of multiple independent rounds. In each specific round, we randomly replace a fixed number (gradually increased) of clear GT images in the training set with corresponding hazy images to construct current partially polluted training data. Through these rounds of experiments, we could directly observe how the increasing of faulty GT training samples may decay proposed approach's performance.

The performance under incrementally polluted training sets are demonstrated in Fig. 11. It could be found from this figure that the precision degrading curve of LPATR-Net is generally gentle, especially in the intervals with less than 15% faulty GT. Quantitatively, when the polluted GT are increased from 0% to 25%, LPATR-Net's accuracy reduction is 2.71(PSNR) and 0.0327(SSIM), that of FSNet [39] is 3.33 and 0.0453, and that of SFNet [64] is 3.48 and 0.0372. Their average descending rate ($\Delta y/\Delta x$) is respectively **-13.55**, −16.65, −17.40 (PSNR), and **-0.1635**, −0.2265, −0.1860 (SSIM). These data imply proposed LPATR-Net is more competent to tolerate moderate-level training GT fault, and possesses higher robustness as to sample labeling accuracy compared with SOTA.

I. The Transplantability of LFFT Component

As we have mentioned in Section III-A (also could be observed in Fig. 3), proposed fitting flexibility restriction component, the Learnable Fitting Flexibility Tightening (LFFT), is relatively independent within the introduced overall dehazing architecture, as to both functional purpose and structural topology. Thus, it is interesting to explore its effectiveness when transplanted into other high-accuracy fitting architecture, namely without the specifically designed optimal partner All-Mattering. Accordingly, in this series of experiments, LFFT is transplanted into four SOTA dehazing fitting architectures, and Table VII reports the corresponding performance.

From this table, it could be observed that the enrollment of faulty GT resisting LFFT component generally could improve the overall dehazing accuracy, especially in the DehazeFormer-B and OKNet case. Here, the most evident promotion is ↑6.05 PSNR (to 41.00) in DehazeFormer-B on SOTS-outdoor. This performance means DehazeFormer-B could directly rank the 2nd in Table II, while originally it is only the 14th, an amazing advance. On the other hand, it could also be observed that even the aggressively shrunk versions of LFFT (removing the bulky General Enhancing or/and Content Enhancing) are beneficial. Clearly, these approaches actually realize better balance between accuracy and efficiency. In detail, the minimal model is only 5.41 M (Parameter amount) and 217.64 G (FLOPs), corresponding to DehazeFormer-B + LFFT w/o CE & GE.

Based on these experiments, it could be asserted that proposed LFFT component could be viewed as a promising widely applicable robustness improving plug-in.

V. CONCLUSION

Unlike most supervised image enhancing missions, dehazing is doomed to be bothered by part of unreliable ground truth "haze-free" samples, because natural weather conditions are extremely complicated and hardly correctly categorized. Thus, the capacity of effectively resisting this disturbance should be a requisite skill to any robust dehazing approaches. In this paper, we propose to intentionally restrain fitting freedom to force the model to mainly focus on the majority of training data, so as to effectively diminish the negative effects of faulty ground truth samples. Specifically, we work out a novel robust dehazing framework, LPATR-Net, whose kernel is an elaborately created fitting restricted dehazing regression structure. Through the exceptional cooperation of such flexibility restriction mechanism and the highly-customized optimal compensator, namely the multi-concerns high-accuracy fitting expert All-Mattering, LPATR-Net elegantly achieves evident general performance promotion. It is worth mentioning that this improvement on the robustness to faulty GT does not incur

any extra human labeling burden. Extensive experiments have been conducted on a number of publicly available datasets to verify its effectiveness. Moreover, it is also demonstrated that its core flexibility restriction component LFFT could be solely transplanted into many other general dehazing fitting architectures to widely promote their resistance to faulty GT.

ACKNOWLEDGMENT

The authors would like to thank Professor Mohammed Bennamoun and Farid Boussaid of The University of Western Australia to give valuable suggestions and beneficial discussions.

REFERENCES

[1] W. L. Jun and Z. Rong, "Image defogging algorithm of single color image based on wavelet transform and histogram equalization," *Appl. Math. Sci.*, vol. 7, pp. 3913–3921, May 2013.

[2] A. Y. Azizah, L. Rahadianti, and H. Deborah, "An introductory study on image quality of dehazed images," in *Proc. Int. Conf. Adv. Comput. Sci. Inf. Syst. (ICACSIS)*, Oct. 2020, pp. 301–308.

[3] Y. Lu, X. Liu, and G. Liu, "A new dehazing algorithm based on overlapped sub-block homomorphic filtering," *Proc. SPIE*, vol. 9875, May 2015, Art. no. 987502.

[4] N. Joy and J. Johnson, "Fast image dehazing method based on homomorphism filtering and contrast enhancement," *IJSRET*, vol. 5, pp. 305–309, Mar. 2019.

[5] M. Ju, C. Ding, D. Zhang, and Y. J. Guo, "Gamma-correction-based visibility restoration for single hazy images," *IEEE Signal Process. Lett.*, vol. 25, no. 7, pp. 1084–1088, Jul. 2018.

[6] A. Kumar, R. K. Jha, and N. K. Nishchal, "An improved gamma correction model for image dehazing in a multi-exposure fusion framework," *J. Vis. Commun. Image Represent.*, vol. 78, Jul. 2021, Art. no. 103122.

[7] M. Ju, C. Ding, Y. J. Guo, and D. Zhang, "IDGCP: Image dehazing based on gamma correction prior," *IEEE Trans. Image Process.*, vol. 29, pp. 3104–3118, 2020.

[8] S.-H. Yun, J. H. Kim, and S. Kim, "Image enhancement using a fusion framework of histogram equalization and Laplacian pyramid," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2763–2771, Nov. 2010.

[9] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.

[10] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3522–3533, Nov. 2015.

[11] T. M. Bui and W. Kim, "Single image dehazing using color ellipsoid prior," *IEEE Trans. Image Process.*, vol. 27, no. 2, pp. 999–1009, Feb. 2018.

[12] Y. Gao, H.-M. Hu, B. Li, Q. Guo, and S. Pu, "Detail preserved single image dehazing algorithm based on airlight refinement," *IEEE Trans. Multimedia*, vol. 21, no. 2, pp. 351–362, Feb. 2019.

[13] B. Li et al., "Benchmarking single-image dehazing and beyond," *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, Jan. 2019.

[14] Y. Zhang, S. Zhou, and H. Li, "Depth information assisted collaborative mutual promotion network for single image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2024, pp. 2846–2855.

[15] Y. Song, Z. He, H. Qian, and X. Du, "Vision transformers for single image dehazing," *IEEE Trans. Image Process.*, vol. 32, pp. 1927–1941, 2023.

[16] Y. Qiu, K. Zhang, C. Wang, W. Luo, H. Li, and Z. Jin, "MB-TaylorFormer: Multi-branch efficient transformer expanded by Taylor formula for image dehazing," in *Proc. ICCV*, 2023, pp. 12756–12767.

[17] H. Wu et al., "Contrastive learning for compact single image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2021, pp. 10551–10560.

[18] Y. Li, Y. Liu, Q. Yan, and K. Zhang, "Deep dehazing network with latent ensembling architecture and adversarial learning," *IEEE Trans. Image Process.*, vol. 30, pp. 1354–1368, 2021.

[19] Y. Zheng, J. Su, S. Zhang, M. Tao, and L. Wang, "Dehaze-AGGAN: Unpaired remote sensing image dehazing using enhanced attention-guide generative adversarial networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5630413.

[20] Y. Liu et al., "From synthetic to real: Image dehazing collaborating with unlabeled real data," in *Proc. 29th ACM Int. Conf. Multimedia*, Oct. 2021, pp. 50–58.

[21] Z. Chen, Y. Wang, Y. Yang, and D. Liu, "PSD: Principled synthetic-to-real dehazing guided by physical priors," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 7180–7189.

[22] S. Zhao, L. Zhang, Y. Shen, and Y. Zhou, "RefineDNet: A weakly supervised refinement framework for single image dehazing," *IEEE Trans. Image Process.*, vol. 30, pp. 3391–3404, 2021.

[23] H. Bai, J. Pan, X. Xiang, and J. Tang, "Self-guided image dehazing using progressive feature fusion," *IEEE Trans. Image Process.*, vol. 31, pp. 1217–1229, 2022.

[24] P. Shyam, K.-J. Yoon, and K.-S. Kim, "Towards domain invariant single image dehazing," in *Proc. AAAI Conf. Artif. Intell.*, May 2021, vol. 35, no. 11, pp. 9657–9665.

[25] H. Ullah et al., "Light-DehazeNet: A novel lightweight CNN architecture for single image dehazing," *IEEE Trans. Image Process.*, vol. 30, pp. 8968–8982, 2021.

[26] C. Xiao and J. Gan, "Fast image dehazing using guided joint bilateral filter," *Vis. Comput.*, vol. 28, nos. 6–8, pp. 713–721, Jun. 2012.

[27] J. Chen, A. Adams, N. Wadhwa, and S. W. Hasinoff, "Bilateral guided upsampling," *ACM Trans. Graph.*, vol. 35, no. 6, pp. 1–8, Nov. 2016.

[28] T. Fan, C. Li, X. Ma, Z. Chen, X. Zhang, and L. Chen, "An improved single image defogging method based on retinex," in *Proc. 2nd Int. Conf. Image, Vis. Comput. (ICIVC)*, Jun. 2017, pp. 410–413.

[29] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 1674–1682.

[30] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3194–3203.

[31] K. Tang, J. Yang, and J. Wang, "Investigating haze-relevant features in a learning framework for image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2995–3000.

[32] X. Wang, X. Zhang, C. Yan, Q. Sun, and Y. Li, "A novel simulation for polarization dehazing," *IEEE Signal Process. Lett.*, vol. 31, pp. 341–345, 2023.

[33] X. Liu, Y. Ma, Z. Shi, and J. Chen, "GridDehazeNet: Attention-based multi-scale network for image dehazing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7314–7323.

[34] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Xie, "FFA-Net: Feature fusion attention network for single image dehazing," in *Proc. AAAI Conf. Artif. Intell.*, Feb. 2020, pp. 11908–11915.

[35] Z. Zheng et al., "Ultra-high-definition image dehazing via multi-guided bilateral learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 16180–16189.

[36] C. Guo, Q. Yan, S. Anwar, R. Cong, W. Ren, and C. Li, "Image dehazing transformer with transmission-aware 3D position embedding," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5812–5820.

[37] Z. Chen, Z. He, and Z.-M. Lu, "DEA-Net: Single image dehazing based on detail-enhanced convolution and content-guided attention," *IEEE Trans. Image Process.*, vol. 33, pp. 1002–1015, 2024.

[38] Y. Cui, W. Ren, X. Cao, and A. Knoll, "Focal network for image restoration," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2023, pp. 13001–13011.

[39] Y. Cui, W. Ren, X. Cao, and A. Knoll, "Image restoration via frequency selection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 46, no. 2, pp. 1093–1108, Feb. 2024.

[40] Y. Cui and A. Knoll, "Dual-domain strip attention for image restoration," *Neural Netw.*, vol. 171, pp. 429–439, Mar. 2024.

[41] Z. Zhao et al., "CDDFuse: Correlation-driven dual-branch feature decomposition for multi-modality image fusion," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 5906–5916.

[42] E. J. McCartney, "Optics of the atmosphere: Scattering by molecules and particles," New York, TY, USA, Tech. Rep., 1976.

[43] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, "Deep bilateral learning for real-time image enhancement," *ACM Trans. Graph.*, vol. 36, no. 4, pp. 1–12, Aug. 2017.

[44] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, vol. 9351. Cham, Switzerland: Springer, 2015, pp. 234–241.

[45] A. Vaswani et al., "Attention is all you need," in *Proc. NIPS*, vol. 30, 2017, pp. 1–11.

[46] C. O. Ancuti, C. Ancuti, R. Timofte, and C. De Vleeschouwer, "O-HAZE: A dehazing benchmark with real hazy and haze-free outdoor images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 867–8678.
[47] C. O. Ancuti, C. Ancuti, M. Sbert, and R. Timofte, "Dense-Haze: A benchmark for image dehazing with dense-haze and haze-free images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2019, pp. 1014–1018.
[48] C. Ancuti, C. O. Ancuti, R. Timofte, and C. D. Vleeschouwer, "I-HAZE: A dehazing benchmark with real hazy and haze-free indoor images," in *Proc. Int. Conf. Adv. Concepts Intell. Vis. Syst.* Cham, Switzerland: Springer, 2018, pp. 620–631.
[49] C. O. Ancuti, C. Ancuti, and R. Timofte, "NH-HAZE: An image dehazing benchmark with non-homogeneous hazy and haze-free images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 1798–1805.
[50] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.
[51] I. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," in *Proc. ICLR*, 2016, pp. 1769–1784.
[52] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "AOD-Net: All-in-one dehazing network," in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 2017, pp. 4770–4778.
[53] W. Ren et al., "Gated fusion network for single image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 3253–3261.
[54] Z. Deng et al., "Deep multi-model fusion for single-image dehazing," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 2453–2462.
[55] D. Chen et al., "Gated context aggregation network for image dehazing and deraining," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jun. 2019, pp. 1375–1383.
[56] J. Dong and J. Pan, "Physics-based feature dehazing networks," in *Proc. 16th Eur. Conf. Comput. Vis.*, 2020, pp. 188–204.
[57] M. Fu, H. Liu, Y. Yu, J. Chen, and K. Wang, "DW-GAN: A discrete wavelet transform GAN for nonhomogeneous dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2021, pp. 203–212.
[58] T. Ye et al., "Perceiving and modeling density for image dehazing," in *Proc. ECCV*, 2022, pp. 130–145.
[59] Z. Tu et al., "MAXIM: Multi-axis MLP for image processing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5769–5780.
[60] M. Hong, J. Liu, C. Li, and Y. Qu, "Uncertainty-driven dehazing network," in *Proc. AAAI*, 2022, vol. 36, no. 1, pp. 906–913.
[61] Y. Yang, C. Wang, R. Liu, L. Zhang, X. Guo, and D. Tao, "Self-augmented unpaired image dehazing via density and depth decomposition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 2037–2046.
[62] Y. Cui, Y. Tao, L. Jing, and A. Knoll, "Strip attention for image restoration," in *Proc. 32nd Int. Joint Conf. Artif. Intell.*, Aug. 2023, pp. 645–653.
[63] H. Shen, Z.-Q. Zhao, Y. Zhang, and Z. Zhang, "Mutual information-driven triple interaction network for efficient image dehazing," in *Proc. 31st ACM Int. Conf. Multimedia*, Oct. 2023, pp. 7–16.
[64] Y. Cui et al., "Selective frequency network for image restoration," in *Proc. ICLR*, 2023.
[65] Y. Cui, W. Ren, S. Yang, X. Cao, and A. Knoll, "IRNeXt: Rethinking convolutional network design for image restoration," in *Proc. ICML*, 2023.
[66] Y. Zheng, J. Zhan, S. He, J. Dong, and Y. Du, "Curricular contrastive regularization for physics-aware single image dehazing," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2023, pp. 5785–5794.
[67] Y. Cui, W. Ren, and A. Knoll, "Omni-kernel network for image restoration," in *Proc. AAAI*, 2024, vol. 38, no. 2, pp. 1426–1434.
[68] Y. Cui, Q. Wang, C. Li, W. Ren, and A. Knoll, "EENet: An effective and efficient network for single image dehazing," *Pattern Recognit.*, vol. 158, Feb. 2025, Art. no. 111074.
[69] Y. Cui, W. Ren, and A. Knoll, "Exploring the potential of pooling techniques for universal image restoration," *IEEE Trans. Image Process.*, vol. 34, pp. 3403–3416, 2025.

Yuelong Li received the Ph.D. degree in computer science from Peking University, China. He was a Visitor at the Professor Edwin Hancock's Group, University of York, U.K. He is currently a Professor and the Vice Dean of the School of AI, Tiangong University, Tianjin, China. He is a Visiting Professor of the Professor Mohammed Bennamoun and Farid Boussaid's Group, The University of Western Australia, Australia. He has hosted three projects of the National Natural Science Foundation of China, and two projects of Tianjin Natural Science Foundation. He has published more than 60 academic articles. His research interests focus on low-level enhancing, image synthesis, object detection, object recognition, shape modeling, and data embedding. He is a reviewer for many journals and conferences, such as CVPR, ICCV, ECCV, and ACM MM.

Fei Chen is currently pursuing the bachelor's degree with the School of Innovation, Tiangong University, Tianjin, China. He will join Hunan University to further pursue the master's degree. He has worked on computer vision and artificial intelligence researches for several years in topics of image dehazing and time series analysis.

Zhenwei Liu received the B.S. degree from the School of Software, Tiangong University, Tianjin, China. He is currently pursuing the master's degree with Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen, China. His research interests include image enhancing and network analysis.

Tianyu Zang is currently pursuing the bachelor's degree with the School of Software, Tiangong University, Tianjin, China. He will join the National University of Defense Technology to further pursue the master's degree. His research interests include computer vision and artificial intelligence.

Jianming Wang received the M.S. and Ph.D. degrees from Tianjin University, China, in 2000 and 2003, respectively. From September 2007 to 2008, he was a Visiting Scholar with the Department of Statistics, Carnegie Mellon University, USA. He is currently a Full Professor with the Department of Computer, Tiangong University, China.