PPG-Net: Polarization Physics-Guided Dehazing for Near-Surface Optical Remote Sensing

Wencheng Wang*, *Member, IEEE*, Lun Li, Hongtao Liu, Zhenxue Chen, and Xiaohui Yuan, *Senior Member, IEEE*

***Abstract*—**Atmospheric haze attenuates target radiance and adds path radiance, obscuring land-surface structure in optical imagery. We present PPG-Net, a polarization physics-guided image-dehazing network that converts four analyzer-angle observations into Stokes components, polarization features, a coarse transmission map, and a coarse physical reconstruction. A multiscale stem and dilated residual backbone jointly predict scene radiance and transmission. Inference-time fusion combines the direct restoration, transmission-based inversion, and a dark-channel restoration. Evaluation uses synthetic polarimetric haze generated from four terrestrial RGB-D sources, with splits disjoint at the source RGB-D image level. This short-range benchmark does not reproduce airborne or satellite atmospheric paths. PPG-Net obtains 23.11 dB PSNR and 0.8290 SSIM on the test set. Budget-matched retraining of multiple control variants evaluates the contributions of input representation and transmission supervision before inference-time fusion. The results reveal a trade-off between image fidelity and physical interpretability. Fusion improves the complete model's direct output. Six public real polarimetric examples show clearer contours but retain darkening and color bias. These results support polarization-guided restoration under the tested synthetic conditions; deployment in near-surface optical remote sensing still requires calibrated real-data validation.

***Index Terms*—**Atmospheric scattering, image dehazing, optical remote sensing, physics-guided deep learning, polarimetric imaging.

I. INTRODUCTION

In optical remote-sensing systems, land-surface targets are observed through a scattering atmosphere. Haze attenuates direct radiance and adds path radiance, reducing contrast, distorting color, and obscuring fine spatial structures that support scene interpretation. These effects are particularly pronounced in long-range terrestrial, airborne, and satellite imaging, where the optical path can vary across the scene. Reliable haze correction is therefore an important preprocessing step for extracting geoscientific information from optical imagery.

This work was supported in part by the Natural Science Foundation of Shandong Province under Grant ZR2023MF047, Grant ZR2024MA055, and Grant ZR2023QF139; in part by the Science and Technology Development Plan of Weifang City under Grant 2025ZJ1113; in part by the China University Industry-University-Research Innovation Foundation under Grant 2021ZYA11003; in part by the National Natural Science Foundation of China under Grant 61403283. *(Corresponding author: Wencheng Wang.)*

Wencheng Wang, Lun Li, Hongtao Liu are with College of Intelligent Science and Engineering, Weifang University, Weifang, China (e-mail: 20111158@wfu.edu.cn; 20200013@wfu.edu.cn; 20210124@wfu.edu.cn).

Zhenxue Chen is with College of Control Engineering, Shandong University, Jinan 250100, China (e-mail: chenzhenxue@sdu.edu.cn).

Xiaohui Yuan is with College of Engineering, University of North Texas, Denton, TX 76203 USA (e-mail: xiaohui.yuan@unt.edu).

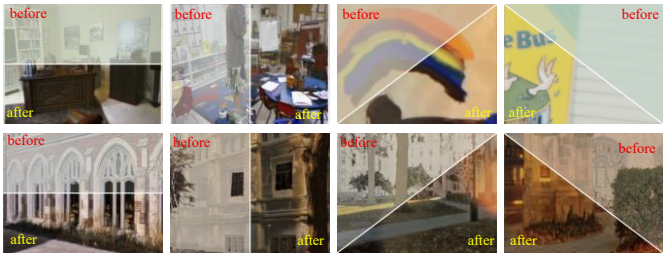


Fig. 1. Eight hazy inputs and the corresponding PPG-Net outputs

This study focuses on polarimetric image restoration for short-range terrestrial scenes. Potential applications include ground-based observation, infrastructure monitoring, and close-range mapping with polarization cameras. The benchmark does not reproduce long atmospheric paths, sensor spectral responses, or radiometric calibration in airborne and satellite systems. Evaluation therefore concerns image restoration and physical-variable estimation, not operational remote-sensing accuracy.

A single RGB image combines attenuated scene radiance with airlight from atmospheric scattering. Recovering the clear image therefore requires joint estimation of depth-dependent transmission and atmospheric light, an ill-posed inverse problem. Classical methods constrain the solution with assumptions about dark channels, color attenuation, or local contrast. These assumptions can fail in bright, low-texture, or spatially nonuniform regions. Deep networks learn more flexible mappings but usually operate on one unpolarized hazy image and do not directly measure the scattering component. Polarimetric imaging adds measurements of polarization magnitude and orientation, which help distinguish airlight from scene radiance. PPG-Net combines these measurements with channel-minimum priors and learned restoration. We evaluate the restored images, the consistency of the estimated physical variables, and changes in the output of a frozen detector. Figure 1 shows representative results.

The main contributions of this work are summarized as follows.

1) We propose PPG-Net as a polarimetric image-dehazing method with potential use in near-surface optical remote sensing. Unlike RGB-only restoration, it explicitly exposes polarization-derived airlight cues and a learned transmission estimate.

2) We construct a structured polarimetric prior by converting four-angle measurements into Stokes components, the degree and angle of linear polarization, a coarse transmission map and a coarse physical reconstruction. These variables provide interpretable cues about airlight and transmission.

3) We design a multiscale stem, a dilated residual backbone, and dual heads for haze-free radiance and transmission prediction. A physics-guided fusion stage corrects the final

output while retaining spatial resolution, restoration capacity, and physical interpretability.

4) We provide a reproducible synthetic forward model, source-RGB-D-image-disjoint splits, matched-protocol retrained controls, direct transmission validation, prior-parameter sensitivity, and out-of-domain real-sample analysis.

The remainder of this paper is organized as follows. Section II reviews related work. Section III derives the polarimetric image-formation model and prior tensor. Section IV describes the network and training objective. Section V presents the experimental protocol, restoration results, controlled analyses, and real-image transfer. Section VI concludes the paper.

II. RELATED WORK

Image dehazing research can be broadly grouped into traditional image-processing methods, physical-model and image-prior methods, deep learning-based methods, and polarization- or multimodal-based methods.

A. Traditional Image-Processing Methods

Nonlearning dehazing includes contrast optimization, filtering, intensity transformations, and image fusion. Wang et al. [1] use a linear transformation based on the minimum intensity channel, whereas Narasimhan and Nayar [2] analyze atmospheric effects on image color and visibility. Tan [3] recovers visibility by favoring high local contrast, and Tarel and Hautière [4] develop a fast filtering-based restoration procedure. Ancuti and Ancuti [5] combine complementary enhanced inputs through multiscale fusion without explicit depth estimation. These methods avoid paired-data training, but their assumptions and computational costs differ. Enhancement-driven approaches can improve visibility without uniquely recovering scene radiance; methods that estimate atmospheric variables additionally depend on the validity of their image priors. These distinctions motivate the physical-model methods reviewed next.

B. Physical Models and Image Priors

A large body of work has used the atmospheric scattering model to improve interpretability. The model expresses a hazy observation as attenuated scene radiance plus airlight, reducing dehazing to the estimation of transmission and atmospheric light [6]. Schechner et al. introduced polarization measurements to analyze atmospheric scattering and restore scene visibility [7]. He et al. proposed the dark channel prior (DCP) [8]. They observed that at least one color channel in most local patches outside the sky region has very low intensity, which enables transmission estimation. The DCP has since become a standard baseline. Meng et al. improved transmission estimation through boundary constraints and contextual regularization [9], Fattal explained hazy color distributions using color lines [10], and Zhu et al. proposed the color attenuation prior to infer depth from brightness and saturation [11]. Berman et al. introduced haze lines to recover transmission from color clusters [12]. More recent priors include the saturation-line prior [13], patch-line and fuzzy-clustering priors [14], and the gradient-line prior [15]. Physical models and hand-crafted priors are interpretable and require no large-scale training set. Their accuracy, however, depends strongly on whether the assumed image statistics hold. Biased estimates of transmission or atmospheric light then propagate directly to the restored image.

C. Deep Learning-Based Methods

Deep learning has substantially advanced single-image dehazing. DehazeNet learned a mapping from a hazy image to its transmission map with a convolutional neural network [16], and MSCNN estimated transmission in a coarse-to-fine multiscale architecture [17]. AOD-Net reparameterized variables in the atmospheric scattering model to enable end-to-end haze-free reconstruction [18]. DCPDN jointly estimated transmission, atmospheric light, and scene radiance by embedding the atmospheric scattering model in the network [19], whereas GFN fused complementary enhancement branches through learned gates [20]. The RESIDE benchmark facilitated supervised training and standardized evaluation [21]. EPDN, GridDehazeNet, FFA-Net, and MSBDN use adversarial image-to-image translation, grid connectivity, feature attention, and multiscale boosting, respectively [22]–[25]. Shao et al. used domain adaptation to reduce the synthetic-to-real gap [26]. AECR-Net introduced contrastive regularization for supervised dehazing [27], whereas PSD adapted a pretrained model to real hazy images using physical-prior losses [28]. TransWeather, DehazeFormer, MB-TaylorFormer, and GridFormer further exploited long-range dependencies and multiscale feature interaction [29]–[32]. Recent work has expanded toward diffusion restoration, real scenes, and remote sensing. RSHazeDiff applied a diffusion model to remote-sensing dehazing [33], Su et al. combined multiple priors with domain transfer for real scenes [34], Saxena et al. developed an adaptive self-attention conditional GAN [35], and Yang et al. investigated semantic-assisted unpaired dehazing [36]. Despite these advances, cross-scene stability and preservation of task-relevant structure remain sensitive to the training distribution. Cross-domain remote-sensing instance segmentation [37] demonstrates the relevance of appearance-domain differences, although it does not itself test haze removal. Recent advances also include unpaired optimal-transport dehazing [38], degradation-aware self-prompting for all-in-one weather restoration [39], and cross-level interaction with intralevel fusion for remote-sensing image dehazing [40]. Frequency-regularized optimal transport for rain removal [41] is not a dehazing method, but it provides a related example of structure-preserving unpaired restoration under weather degradation.

D. Polarization- and Multimodal-Based Dehazing

Polarimetric imaging adds measurements of the polarization state of the observed light. Under appropriate illumination and viewing geometry, partially polarized airlight can be distinguished from more weakly polarized scene radiance [7]. Shen et al. [42] jointly address polarimetric dehazing and noise suppression through iterative optimization. Wang et al. [43] fuse intensity and degree-of-polarization information, while Suo et al. [44] combine a dynamic dark-channel prior with polarization. Yan et al. [45] develop a polarized-haze simulation procedure for data generation and evaluation. Learning-based approaches extend these ideas: Wang et al. [46] estimate polarization-related parameters in dynamic turbid underwater environments, and Chen et al. [47] integrate

polarization inference with learned dehazing. The underwater results in [46] concern a different scattering medium and are not direct evidence of atmospheric dehazing performance. Zhou et al. [48] model both transmitted-light and airlight polarization in a learned dehazing pipeline, avoiding the assumption that the direct component is always unpolarized. PPG-Net instead retains that simplifying assumption in its synthetic generator and tests its limitations through real-image transfer. In remote sensing, multiangle polarimetry is more commonly used to retrieve aerosol microphysics and optical properties than to reconstruct a haze-free RGB image. Reviews of aerosol polarimetry [49], multiangle spaceborne spectropolarimetry [50], and POLDER land-aerosol retrieval [51] establish that polarization carries atmospheric scattering information, but their inversion targets and path lengths differ fundamentally from near-field image dehazing. PPG-Net should therefore be interpreted as an image-restoration model informed by polarimetric physics, not as an aerosol-retrieval algorithm.

These studies suggest complementary roles for physical priors, polarization measurements, and learned restoration. PPG-Net combines them through a structured input representation, joint radiance and transmission prediction, and inference-time fusion. The experiments examine how the additional measurements and physical constraints affect restoration and transmission estimation.

III. POLARIMETRIC IMAGE FORMATION AND PHYSICAL PRIOR DERIVATION

Polarization measurements provide additional scattering-related cues beyond ordinary RGB intensity, but their separation from surface polarization depends on the image-formation assumptions. We construct a pathway from four-angle observations to the network input. For each RGB channel, the measurements yield the Stokes components S_0, Q, U and the channelwise DoLP ρ_c. Averaging ρ_c gives scalar DoLP ρ; averaging the RGB Q and U components before evaluating the half-angle relation gives scalar AoP φ. A coarse transmission t_0 and reconstruction J_{prior} are then formed, producing the 40-channel tensor X in Fig. 2.

A. Ill-Posedness of Haze Degradation

Under the homogeneous-atmosphere approximation, one intensity observation mixes haze-free scene radiance, depth-dependent transmission, and scattered airlight. Without additional constraints, either a network or a prior-based solver must infer several coupled unknowns from a single measurement. The atmospheric scattering model is

$$I_c(x) = J_c(x)t(x) + A_c[1 - t(x)], c \in \{r, g, b\} \quad (1)$$

where x denotes pixel location, $c \in \{r, g, b\}$ the color channel, $I_c(x)$ the c-th channel of the nonpolarized hazy image, $J_c(x)$ the haze-free scene radiance, A_c the atmospheric-light component, and $t(x)$ the transmission. Its relationship to scene depth and the scattering coefficient is given by (2).

$$t(x) = \exp[-\beta d(x)] \quad (2)$$

Here, $d(x)$ is scene depth and β is a spatially constant scattering coefficient under the homogeneous-atmosphere model. Transmission varies spatially because depth varies; this is not a model of spatially varying aerosol concentration. Increasing depth or β reduces the direct component and increases the relative contribution of airlight.

When $t(x)$ and the RGB atmospheric-light vector A are known, the haze-free image can be recovered by the physical inversion in (3).

$$J_c(x) = \text{clip}_{[0,1]}\left(\frac{I_c(x) - A_c}{\max[t(x), t_{\min}]} + A_c\right) \quad (3)$$

Here, $t_{\min}$ is a numerical transmission floor and clip[0,1] denotes componentwise clipping to the normalized intensity range. Equation (3) matches the implemented inversion: the same bounded transmission is used in the attenuation and airlight terms before output clipping. It reduces to the inversion of (1) when the floor and clipping are inactive. The safeguard limits division by small values but does not ensure exact radiometric recovery.

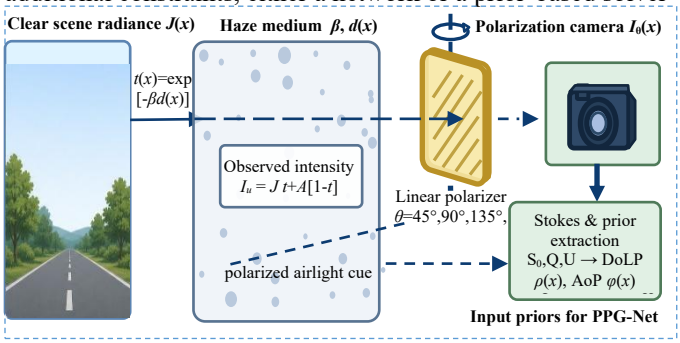


Fig. 2. Polarimetric image formation of the PPG-Net prior tensor.

B. Separable Information in Polarization Measurements

A single intensity observation does not uniquely separate direct radiance from airlight without additional constraints. Polarimetric imaging provides further measurements: atmospheric airlight is often partially polarized, whereas diffuse scene reflection may be more weakly polarized. Intensity changes across analyzer angles encode polarization magnitude and orientation. Their interpretation as scattering cues depends on the assumption about direct-light polarization stated below. The linear-polarizer measurement model is

$$I_\theta(x) = \frac{1}{2}[S_0(x) + Q(x)\cos(2\theta) + U(x)\sin(2\theta)] \quad (4)$$

Here, S_0 is total intensity, while Q and U are intensity differences along two orthogonal linear-polarization bases. We use this Stokes notation consistently. Ideal linear-polarizer intensity is 180° periodic, so I_{180} and I_0 are equivalent. Equation (5) is applied independently to each RGB channel; the implementation additionally clips reconstructed total intensity to [0,1].

$$S_0(x) = \frac{1}{2}\left[\left(I_{45}(x) + I_{135}(x)\right) + \left(I_{90}(x) + I_{180}(x)\right)\right]$$
$$Q(x) = I_{180}(x) - I_{90}(x), U(x) = I_{45}(x) - I_{135}(x) \quad (5)$$

Equation (6) defines polarization magnitude and DoLP for each color channel $c \in \{r,g,b\}$. The stabilizing constant ε prevents division by zero in weak-signal regions; the implementation uses $\varepsilon = 10^{-6}$ and clips channelwise DoLP to [0,1].

$$P_c(x) = \sqrt{[Q_c(x)]^2 + [U_c(x)]^2}$$
$$\rho_c(x) = \frac{P_c(x)}{\max[S_{0,c}(x), \varepsilon]} \quad (6)$$
$$\varphi(x) = \frac{1}{2}\text{atan2}\big(U(x), Q(x)\big) \quad (7)$$

The network retains the three ρ_c values as a three-channel DoLP feature and forms a scalar DoLP $\rho = (\rho_r + \rho_g + \rho_b)/3$ by arithmetic averaging. Likewise, Q_c and U_c are averaged over color channels to obtain $\bar{Q}$ and $\bar{U}$, from which the single-channel AoP is computed using (7). This notation keeps ρ_c distinct from ρ, Q_c from $\bar{Q}$, and U_c from $\bar{U}$.

Equation (8) provides an interpretable approximation relating polarization magnitude to scattered airlight. This relationship motivates the polarization features used by PPG-Net.

$$P_c(x) \approx \rho_A A_{s,c}(x), A_{s,c}(x) = A_c[1 - t(x)] \quad (8)$$

Here, ρ_A is the airlight DoLP and $A_{s,c}(x) = A_c[1 - t(x)]$ is the scattered-airlight component. The approximation in (8) assumes negligible polarization of the direct scene component.

C. From Physical Priors to Network Input

Real polarization measurements remain affected by noise, nonideal sensor response, and complex depth variation. PPG-Net therefore encodes physical quantities as learnable cues. Consistent with the implementation, the atmospheric-light candidate score $D(x)$ is the channel-wise minimum of total intensity S_0 at pixel x, as defined in (9).

$$D(x) = \min_{c\in\{r,g,b\}} S_{0,c}(x) \quad (9)$$

Pixels in the top 0.1% by $D(x)$ are retained as atmospheric-light candidates, and A is estimated from the candidate with the largest summed RGB intensity.

$$t_0(x) = \max\left(t_l, 1 - \min\left(a_l, \frac{\rho(x)}{p_A}\right)\right)$$
$$J_{\text{prior}}(x) = \mathcal{R}(S_0, A, t_0) \quad (10)$$

Equation (10) constructs t_0 heuristically from the observed scalar DoLP and then obtains J_{prior} using the clipped inversion operator R defined by (3), with floor t_l. The normalization p_A is an image-derived scale, not a measured ground-truth airlight DoLP; (10) is therefore not an exact inversion of (8). This input-stage floor is $t_l = 0.15$, distinct from $t_{\min} = 0.08$ in the final correction stage.

Equations (1)–(3) define the intensity-only inverse problem, whereas (4)–(8) introduce polarization cues related to scattered airlight. Equations (9) and (10) convert these cues into an atmospheric-light estimate, a coarse transmission cue, and a coarse reconstructed image for use by the network described in Section IV.

IV. PPG-NET ARCHITECTURE

Fig. 3 shows the overall PPG-Net architecture. Direct closed-form polarimetric inversion is sensitive to erroneous airlight polarization, whereas unconstrained RGB restoration can ignore measured scattering cues. The architecture therefore follows a staged design: explicit polarimetric feature construction reduces the ambiguity of the input; multiscale and dilated convolutions aggregate local edges and spatially extended haze without resolution loss; dual heads preserve an interpretable transmission variable; and a final physics-guided fusion corrects, but does not hard-constrain, the learned image.

The stem extracts shallow multiscale features from the structured tensor; the backbone enlarges the receptive field without downsampling; the dual heads jointly predict the network-restored image and transmission; and the final module combines direct prediction with two physically motivated alternatives. Because the fused image is not reprojected through the forward model during training, this design is physics-guided rather than a strict solution constrained to satisfy the atmospheric image-formation equation.

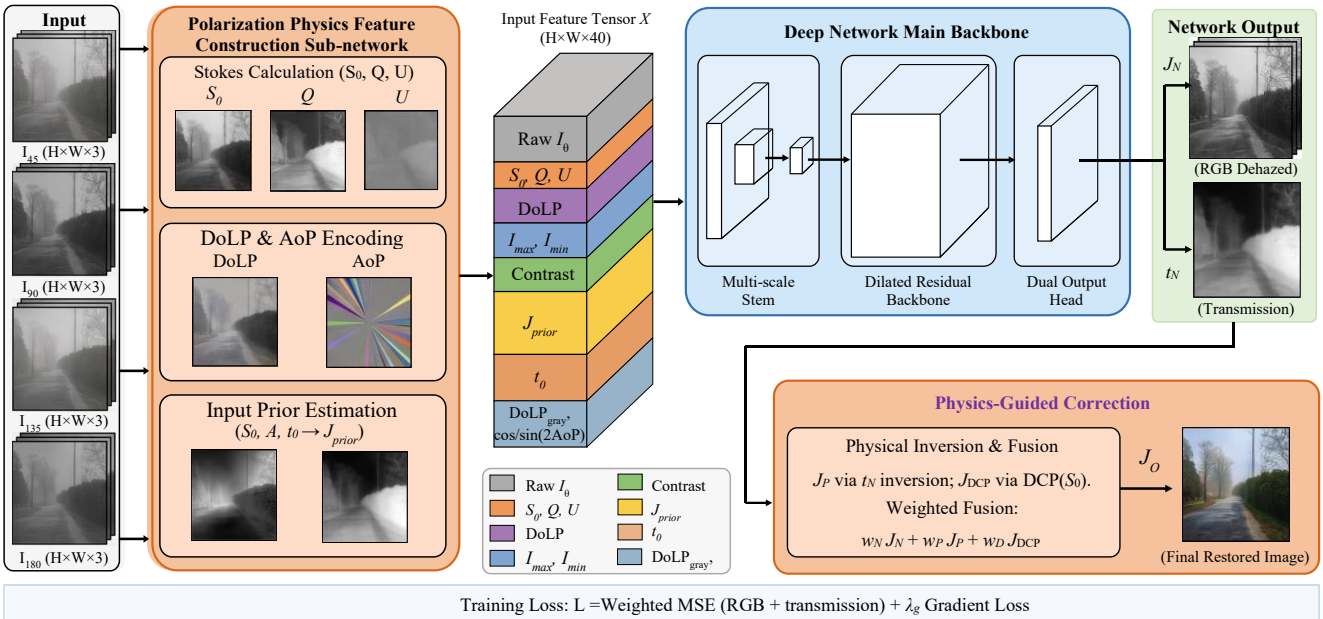


Fig. 3. Overall architecture of PPG-Net.

A. Construction of Polarization-Physics Features

The input on the left side of Fig. 3 is $\Phi(I_{45}, I_{90}, I_{135}, I_{180})$. Under ideal linear polarization, I_{180} and I_0 are equivalent, and J_{prior} denotes the coarse input-stage polarimetric inversion prior.

$$\boldsymbol{X}(x) = \text{Concat}\big(\boldsymbol{F}_{\text{raw}}, \boldsymbol{F}_{\text{Stokes}}, \boldsymbol{F}_{\text{pol}}, \boldsymbol{F}_{\text{prior}}\big)(x)$$
$$\boldsymbol{X} \in \mathbb{R}^{H\times W\times 40} \qquad (11)$$

F_{raw} comprises four RGB polarimetric images; F_{Stokes} contains RGB representations of S_0, Q, and U (9 channels); F_{pol} contains RGB DoLP, $I_{\max}$, $I_{\min}$, polarization contrast, scalar DoLP, and $\cos(2\text{AoP})/\sin(2\text{AoP})$ (15 channels); and F_{prior} contains the three-channel J_{prior} and single-channel t_0 (4 channels). Thus, X has 40 channels.

The representation exposes complementary information to the backbone. S_0, Q, U, DoLP, and AoP encode observable polarization relationships, while J_{prior} and t_0 provide approximate airlight and transmission cues. The latter remain uncertain input estimates rather than physical ground truth, and their effect is evaluated separately in the retrained controls.

B. Multiscale Stem

The multiscale stem is the first core subnetwork of PPG-Net (Fig. 4). It maps the 40-channel physical input to 64 shallow feature channels. Haze degradation combines local edge blur, low-frequency color shifts, and spatially varying veil diffusion, which are difficult to capture with a single kernel size. The stem therefore applies parallel 3 × 3 and 5 × 5 convolutions and a 3 × 3 dilated convolution with dilation 2. Their outputs are concatenated channel-wise and compressed by a 1 × 1 convolution followed by batch normalization (BN) and ReLU.

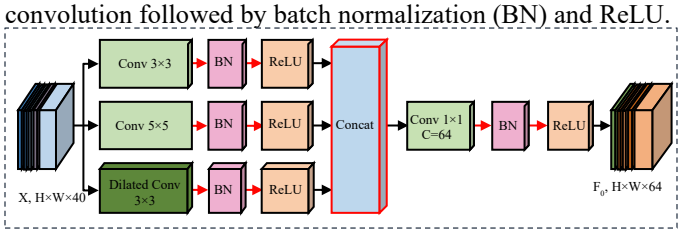


Fig. 4. Multiscale stem.

The fused multiscale-stem feature is expressed as

$$\boldsymbol{F}_0 = \text{Conv}_{1\times1}\big(\text{Concat}\big[B_{3\times3}(\boldsymbol{X}), B_{5\times5}(\boldsymbol{X}), B_{3\times3}{}^{d=2}(\boldsymbol{X})\big]\big) \qquad (12)$$

Equation (12) summarizes the stem. Each branch B comprises its indicated convolution, BN, and ReLU and produces 32 channels. Concatenation forms 96 channels. The Conv1×1 notation denotes the complete projection block, including its following BN and ReLU, giving the 64-channel feature F_0 shown in Fig. 4. Parallel receptive fields allow local edges and more extended intensity variation to enter the backbone together; their individual contributions are not isolated by the present ablations.

C. Dilated Residual Backbone

The residual backbone contains 14 dilated residual blocks (Fig. 5).

Each block has two 3 × 3 dilated convolutions. The first is followed by BN and ReLU, and the second by BN; its output is added to the block input and passed through ReLU. Unlike encoder–decoder downsampling, the backbone maintains the input resolution throughout, preventing excessive compression of fine texture and polarimetric priors during deep feature propagation.

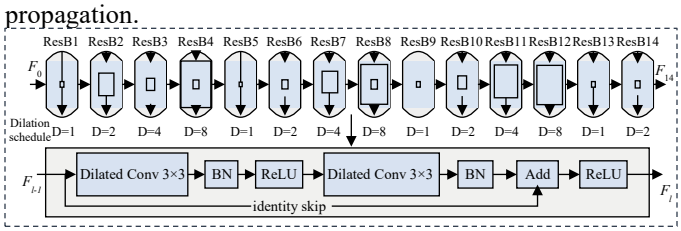


Fig. 5. Dilated residual backbone.

A dilated residual block is written as

$$\boldsymbol{F}_l = \text{ReLU}\big(\boldsymbol{F}_{l-1} + \mathcal{R}_l(\boldsymbol{F}_{l-1}; d_l)\big)$$
$$l = 1,2,\ldots,14, d_l \in \{1,2,4,8\} \qquad (13)$$

Equation (13) describes propagation through the dilated residual backbone. Residual connections retain reliable shallow information, while dilated convolutions enlarge the receptive field without reducing resolution and thereby model both local texture and large-scale nonhomogeneous haze. The dilation schedule d_l = [1, 2, 4, 8, 1, 2, 4, 8, 1, 2, 4, 8, 1, 2] samples local texture, mesoscale structure, and long-range haze at different receptive-field sizes.

This schedule reflects the spatial characteristics of dehazing. Small dilations preserve local edges and texture, whereas large dilations capture slowly varying trends in airlight and transmission. Residual connections allow the network to learn a strong haze-removal transform while retaining trustworthy low-level information from S_0 and the polarization-difference channels.

D. Dual Output Heads

The dual heads map backbone features to haze-free RGB scene radiance and transmission, linking the learned predictions to the physical inversion model. The prediction tail consists of Conv 3 × 3/64, ReLU, Conv 3 × 3/32, ReLU, Conv 1 × 1/4, and Sigmoid. The first three output channels form the network-dehazed image J_{N}, and the last channel is the transmission map t_{N}. PPG-Net thus retains an explicit physical intermediate variable while restoring texture and color.

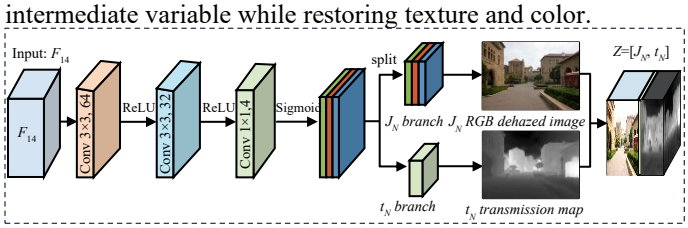


Fig. 6. Dual output heads.

The dual-head prediction is

$$\boldsymbol{Z} = \text{sigmoid}\big(H(\boldsymbol{F}_{14})\big), J_{\text{N}} = \boldsymbol{Z}_{1:3}, t_{\text{N}} = \boldsymbol{Z}_4 \qquad (14)$$

Equation (14) describes the dual output heads, where $\boldsymbol{F}_{14}$ is the output of the 14th dilated residual block, $H(\cdot)$ is the convolutional prediction head, and sigmoid(·) bounds the output. J_{N} is the direct network restoration, and t_{N} is used in the subsequent physical inversion.

E. Physics-Guided Correction and Image Fusion

Although the direct network output can itself serve as a dehazed result, the final PPG-Net output also uses a transmission-based inversion and an independent DCP

restoration of total intensity. These candidates emphasize, respectively, learned texture recovery, atmospheric-model inversion, and a conventional image prior. Their weighted combination is an inference-time correction; it does not guarantee exact satisfaction of the forward model.

$$J_{\mathrm{P}}(x) = \mathrm{clip}_{[0,1]}\left(\frac{S_0(x)-A}{\max[t_{\mathrm{N}}(x), t_{\min}]} + A\right) \qquad (15)$$

In (15), S_0 is the reconstructed RGB total intensity, A is the RGB atmospheric-light vector estimated from its channel-minimum candidate score, and $t_{\min} = 0.08$. The compact form uses bounded transmission consistently with the implementation and clips J_{P} to [0,1] before fusion. It is equivalent to the unfloored atmospheric inversion when t_{N} is above the floor and no output clipping occurs. J_{prior} is the input-stage polarimetric prior, J_{P} uses learned transmission, and J_{DCP} is the separate intensity-only DCP restoration; these are not interchangeable images.

$$J_{\mathrm{O}}(x) = w_{\mathrm{N}}J_{\mathrm{N}}(x) + w_{\mathrm{P}}J_{\mathrm{P}}(x) + w_{\mathrm{D}}J_{\mathrm{DCP}}(x) \qquad (16)$$

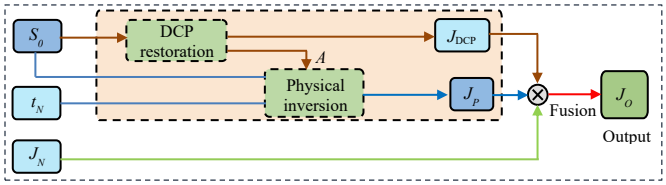


Fig. 7. Physics-guided correction and image fusion.

The nonnegative fusion weights sum to one. Each candidate is bounded to [0,1], and the final image is clipped before storage. The direct network restoration remains dominant, the transmission-based inversion adds radiative-transfer information, and DCP supplies an independent prior. The recorded fixed weights used for the final PPG-Net results are 0.60, 0.25, and 0.15, respectively.

F. Training Supervision and Implementation

Supervision consists of the haze-free RGB image J and ground-truth transmission t. During training, the predicted image and transmission are J_{N} and t_{N}, respectively. The three-branch fusion that produces J_{O} in (16) is applied only at inference and does not participate in backpropagation. To constrain scene-radiance restoration and transmission estimation jointly, we use the weighted mean-squared error in (17).

$$\mathcal{L}_{\mathrm{wMSE}} = \frac{1}{4}[3\mathrm{MSE}(J_{\mathrm{N}}, J) + \lambda_t \mathrm{MSE}(t_{\mathrm{N}}, t)] \qquad (17)$$

Here, MSE(J_{N},J) averages the squared RGB error over the batch, spatial positions, and three color channels; MSE(t_{N},t) averages the single-channel transmission error over the batch and spatial positions. Thus, the factor 3 restores the contribution of the three unit-weight RGB channels before division by four. With $\lambda_{\mathrm{t}} = 0.15$, (17) matches elementwise averaging with output-channel weights [1,1,1,0.15]. The complete objective is defined in (18).

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{wMSE}} + \lambda_g \mathcal{L}_{\mathrm{grad}} \qquad (18)$$

In (18), $\mathcal{L}_{\mathrm{grad}}$ computes the mean-squared errors of the first-order horizontal and vertical forward differences between J_{N} and the haze-free reference J. The two directions are weighted equally, with an overall coefficient $\lambda_g = 0.35$.

V. EXPERIMENTAL RESULTS AND VALIDATION SCOPE

A. Benchmark Construction and Evaluation Protocol

The benchmark was generated from DIODE [52], Middlebury [53], NYU Depth V2 [54], and TUM RGB-D [55]. All retained haze variants of each converted source RGB-D image were assigned to one split, yielding 5040 training, 720 validation, and 480 test samples from 1260, 180, and 280 source images, respectively. Each dataset contributes 120 test samples. This grouping ensures source-image separation, but related frames or views of the same physical scene can occur in different splits; it is not *a* physical-scene- or sequence-disjoint evaluation.

Algorithm 1 specifies the complete forward generator. The degree and angle of linear polarization of airlight are sampled once per haze variant and shared by all pixels and color channels; Q and U are not independently sampled. The direct scene component is treated as unpolarized, and only the airlight term is linearly polarized. For each channel, Q = ρ_{A} B cos(2 φ_{A}) and U = ρ_{A} B sin(2 φ_{A}), where B is the airlight image. Nonfinite depth values are set to zero before per-image min–max normalization; a constant depth map becomes zero. No additive sensor noise is introduced.

Algorithm 1. Synthetic polarimetric haze generation for one RGB-D image and one variant

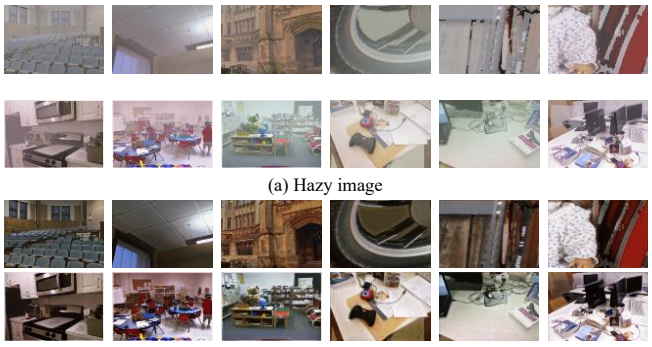
Step	Operation
1	Read clear RGB J and depth D; replace nonfinite depth with 0 and min–max normalize to $D_{01} \in [0,1]$.
2	Set $d = 0.5 + 5.5D_{01}$ m; sample $\beta \sim$ U(0.10, 0.38); set $t = \mathrm{clip}(\exp(-\beta d), 0.25, 1)$.
3	Sample $a \sim$ U(0.68, 0.88) and $u_c \sim$ U(0, 1) independently for $c \in$ {r, g, b}; set $A_c = \mathrm{clip}(a + 0.12(u_c - 0.5), 0, 1)$.
4	Sample airlight DoLP $\rho_{\mathrm{A}} \sim$ U(0.18, 0.55) and AoP $\varphi_{\mathrm{A}} \sim$ U(0°, 180°), once per haze variant.
5	Form direct radiance $L = Jt$ and airlight $B = A(1 - t)$; the unpolarized hazy image is $I_{\mathrm{h}} = L + B$.
6	For $\theta \in$ {45°, 90°, 135°, 180°}, form $I_\theta = \mathrm{clip}(0.5(L + B) + 0.5\rho_{\mathrm{A}} B \cos[2(\theta - \varphi_{\mathrm{A}})], 0, 1)$.
7	Store J, I_{h}, and I_θ as 8-bit PNGs; store D_{01} and t as 16-bit PNGs. Use random seed 7.

The generator represents normalized short-range terrestrial imaging, with optical thickness βd spanning approximately 0.05–2.28 before transmission clipping. It tests controlled scattering severity across four image sources. Its homogeneous atmosphere, normalized depth, and noise-free rendering do not reproduce kilometer-scale paths, vertical aerosol structure, spectral response, or radiometric-calibration error.

DCP [8] and CAP [11] are evaluated alongside seven deep baselines using public pretrained models without fine-tuning or test-time ensembling. All RGB-only methods receive the same unpolarized hazy image; PPG-Net uses four-angle measurements. The local UDPNet [56] evaluation supplies the benchmark's normalized depth in place of its estimated-depth prior, making it an adapted RGB+depth comparison with privileged depth. Outputs are evaluated in RGB [0,1] at 256 × 256. Table II compares complete restoration pipelines with different original training corpora and input modalities. Table III examines feature groups within the PPG-Net backbone using the same training split and update budget.

PPG-Net and every retrained control were implemented in MATLAB R2026a and trained from scratch on one NVIDIA GeForce GTX 1080 Ti. The fixed protocol uses seed 7, 128 × 128 random crops, horizontal and vertical flips with probability 0.5, batch size 24, Adam, an initial learning rate of 1.5×10^{-4}, eight epochs (1680 parameter updates), and a factor-of-two learning-rate reduction after epoch 4. Validation is performed once per epoch; the checkpoint with minimum loss is retained. For the full model this occurs at update 840 (validation loss 0.0069), so extending to eight epochs provides four further validation checks without selecting the test set. For the full model only, a fixed-order training-data pass at learning rate 1×10^{-12} restored missing batch-normalization running statistics. The new controls retained the statistics returned with their best-validation checkpoints. This implementation difference limits strict attribution to architecture alone; neither procedure used test images. Inference uses 256 × 256 inputs. The recorded fusion weights (0.60, 0.25, 0.15) were retained for the main test; Table IV reports their validation sensitivity rather than a globally optimal selection.

Figure 8 presents 12 hazy inputs in the first two rows and the corresponding PPG-Net outputs in the next two rows. The examples cover all four source datasets.



(a) Hazy image

(b) PPG-Net output

Fig. 8. Synthetic hazy images and the corresponding PPG-Net restorations.

Across the 12 examples in Fig. 8, PPG-Net reduces the gray-white veil and improves the separation of foreground and background regions. Edges at depth discontinuities become clearer, while large bright regions retain plausible color and exposure. Similar changes appear across the four source datasets. Because these examples show only final outputs, the contribution of each restoration branch is examined separately through the quantitative analysis in Section V-D.

Full-reference evaluation uses PSNR and SSIM [57]. For Tables I and II, saved predictions and references are read as RGB arrays in [0,1] and evaluated with the MATLAB psnr and ssim functions at 256 × 256 without a luminance conversion. Dataset summaries average per-image scores, rather than computing PSNR from a pooled error. Table III controls are evaluated directly on floating-point outputs, whereas the full-model image rows use saved 8-bit outputs.

For a haze-free reference J and restoration $\hat{J}$ with dynamic range L, PSNR is defined by (19), where MSE is the mean-squared error. A higher PSNR indicates lower pixelwise reconstruction error.

$$\mathrm{PSNR}(J,\hat{J}) = 10\log_{10}\frac{L^2}{\mathrm{MSE}(J,\hat{J})} \tag{19}$$

SSIM is defined by (20), where μ, σ^2, and $\sigma_{J\hat{J}}$ denote local means, variances, and covariance, and C_1 and C_2 are stabilizing constants. The metric jointly evaluates luminance, contrast, and structure; a higher value indicates better structural preservation.

$$\mathrm{SSIM}(J,\hat{J}) = \frac{(2\mu_J\mu_{\hat{J}}+C_1)(2\sigma_{J\hat{J}}+C_2)}{(\mu_J{}^2+\mu_{\hat{J}}{}^2+C_1)(\sigma_J{}^2+\sigma_{\hat{J}}{}^2+C_2)} \tag{20}$$

Whereas PSNR primarily measures pixelwise error, SSIM also accounts for luminance, contrast, and structural consistency and is therefore informative about edge hierarchy and local structure in dehazed images.

B. Qualitative Comparison

The qualitative comparison includes DCP, CAP, AOD-Net, FFA-Net, DehazeFormer, PromptIR [58], DEA-Net [59], MB-TaylorFormer [31], and UDPNet [56].

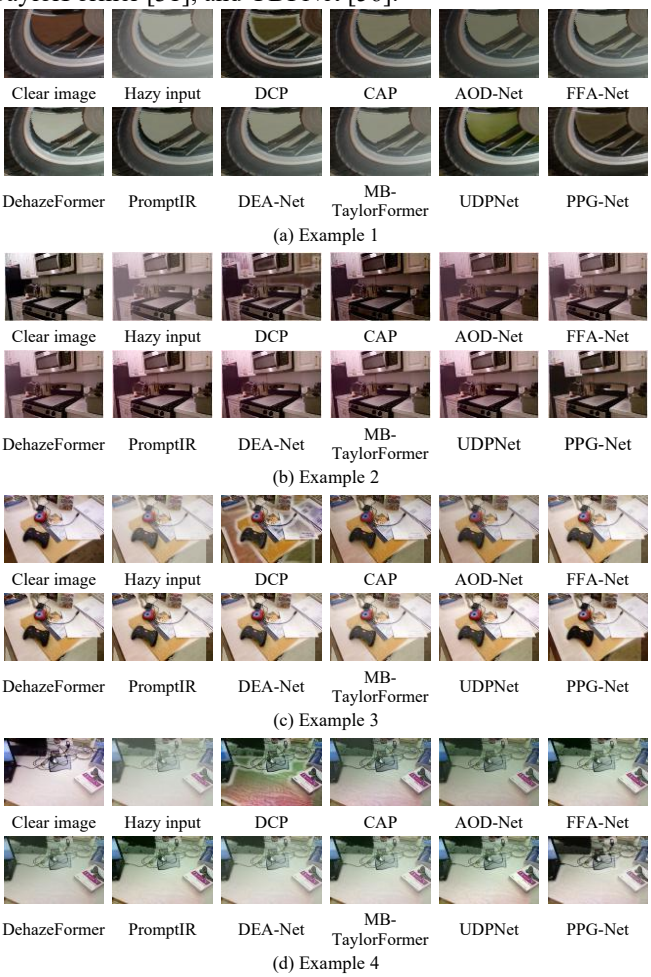


Fig. 9. Qualitative comparison of representative dehazing methods.

Figure 9 shows distinct failure modes among the comparison methods. DCP and CAP increase foreground contrast but leave blockwise haze or introduce color shifts in sky, wall, and other weak-texture regions. AOD-Net, FFA-Net, DehazeFormer, and PromptIR produce more uniform color but smooth some thin structures. DEA-Net and MB-TaylorFormer preserve sharper local edges, although a gray veil remains near large depth transitions. UDPNet improves separation between

near and distant regions by using a depth prior. In the displayed examples, PPG-Net removes more of the veil at these transitions while retaining edge continuity and avoiding pronounced color shifts.

Figures 10 and 11 provide enlarged comparisons on synthetic hazy samples from DIODE and NYU Depth V2, respectively. Red boxes identify the regions of interest, and the row below each full image shows the corresponding crop at a common size for comparing edge continuity, texture fidelity, and residual-haze suppression.

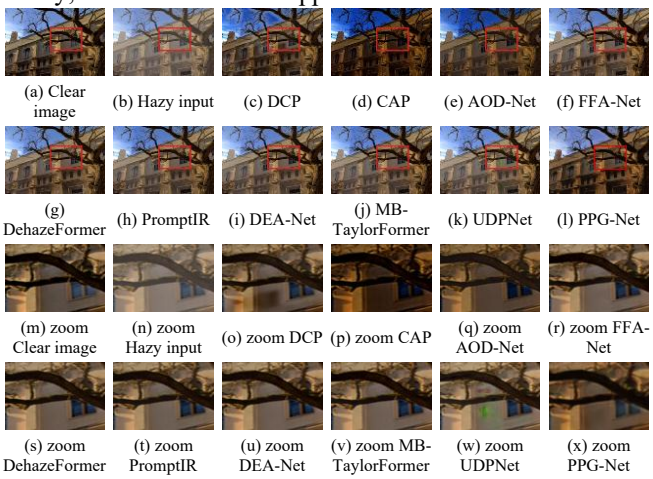


Fig. 10. Enlarged comparison on a DIODE example.

The crops in Figs. 10 and 11 make these differences easier to identify. DCP and CAP produce strong local contrast but can introduce halos and color discontinuities. AOD-Net and FFA-Net remove much of the veil while smoothing thin lines and weak textures. DehazeFormer, PromptIR, DEA-Net, and MB-TaylorFormer preserve structure more consistently, but mild haze remains under strong scattering or abrupt depth changes. UDPNet retains good separation of distant layers in regions with pronounced depth variation. In the displayed crops, PPG-Net maintains continuous contours and more visible background texture without the strong local contrast produced by the traditional priors. Section V-C reports full-reference metrics for these scenes.

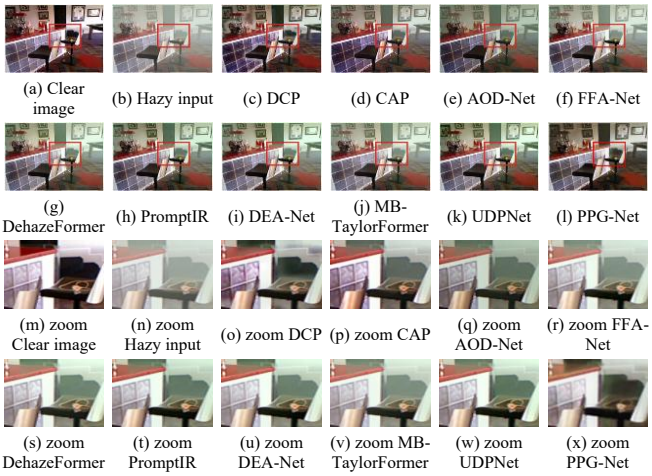


Fig. 11. Enlarged comparison on an NYU Depth V2 example.

C. Objective Evaluation

To support the qualitative observations in Figs. 9–11 with full-reference metrics for the same scenes, Table I reports average PSNR and SSIM over the six displayed scenes. In the performance tables, bold, underlined, and italic values indicate the first, second, and third distinct ranks within each metric and comparison group, respectively; equal displayed values share a rank. PSNR and SSIM are ranked separately.

PPG-Net has the highest average PSNR and SSIM over the six examples in Table I, consistent with the haze suppression and contour preservation in Figs. 9–11. The gains are not uniform across metrics: on the DIODE example in Fig. 10, PPG-Net has higher PSNR but lower SSIM than DehazeFormer. Reduced pixel error therefore does not always imply better local structural agreement.

Table II extends the comparison from the six displayed examples to all 480 fixed test images, with 120 images from each source dataset.

TABLE I
PER-IMAGE AND AVERAGE PSNR/SSIM OVER THE SIX EXAMPLES IN FIGS. 9–11

Method	Fig. 9(a) PSNR↑/SSIM↑	Fig. 9(b) PSNR↑/SSIM↑	Fig. 9(c) PSNR↑/SSIM↑	Fig. 9(d) PSNR↑/SSIM↑	Fig. 10 PSNR↑/SSIM↑	Fig. 11 PSNR↑/SSIM↑	Average PSNR↑/SSIM↑
DCP	17.322/0.740	18.404/0.803	13.948/0.756	11.601/0.533	20.852/0.897	20.357/**0.894**	17.081/0.770
CAP	*19.133*/0.805	16.442/0.746	17.152/0.862	16.386/0.704	15.378/0.757	17.879/0.844	17.061/0.786
AOD-Net	18.963/0.729	17.793/0.750	17.966/0.872	14.591/0.618	17.993/0.841	16.223/0.749	17.255/0.760
FFA-Net	17.632/0.750	19.586/0.797	17.228/0.878	20.044/0.678	*22.162*/*0.895*	14.858/0.745	18.585/0.791
DehazeFormer	14.489/0.731	20.747/0.817	18.486/0.881	18.149/*0.685*	23.731/**0.908**	16.326/0.780	18.655/*0.800*
PromptIR	18.107/0.746	23.937/*0.821*	*18.611*/0.878	*20.247*/0.663	21.227/0.881	17.991/0.801	*20.020*/0.798
DEA-Net	17.148/0.758	*23.851*/0.817	17.814/*0.885*	19.365/0.670	21.048/0.881	17.177/0.792	19.400/*0.800*
MB-TaylorFormer	15.980/0.731	23.812/*0.821*	16.429/0.873	21.075/0.680	20.849/0.881	15.320/0.756	18.911/0.791
UDPNet	21.808/*0.798*	23.013/0.823	18.694/0.887	19.877/0.655	19.623/0.856	*18.647*/0.817	20.277/0.806
PPG-Net	**22.738**/**0.819**	**24.609**/**0.852**	**25.057**/**0.909**	**22.493**/**0.798**	**25.074**/0.886	**21.437**/*0.831*	**23.568**/**0.849**

On the full 480-image test set, PPG-Net has the highest overall mean PSNR and SSIM among the compared methods in Table II. Its PSNR advantage over PromptIR, the strongest external comparison in the overall mean, is 2.33 dB. This comparison evaluates the complete pipeline under the input and training conditions in Section V-A. Table III examines the effects of input information, transmission supervision, and inference-time fusion within the controlled protocol.

TABLE II
PSNR AND SSIM ON THE FOUR SOURCE DATASETS

Method	Avg PSNR↑/SSIM↑	DIODE PSNR↑/SSIM↑	Middlebury PSNR↑/SSIM↑	NYU Depth V2 PSNR↑/SSIM↑	TUM RGB-D PSNR↑/SSIM↑
DCP	16.10/0.699	17.77/0.700	14.87/0.568	16.54/0.780	15.22/0.750
CAP	17.22/0.740	17.65/0.726	*18.82*/0.705	16.90/0.767	15.50/0.760
AOD-Net	17.52/0.767	19.12/0.803	16.49/0.702	17.10/0.749	17.35/0.816
FFA-Net	19.86/0.781	20.42/0.805	18.92/*0.713*	18.52/0.759	21.58/0.848

Method	Avg PSNR↑/SSIM↑	DIODE PSNR↑/SSIM↑	Middlebury PSNR↑/SSIM↑	NYU Depth V2 PSNR↑/SSIM↑	TUM RGB-D PSNR↑/SSIM↑
DehazeFormer	20.07/0.780	19.68/0.810	14.20/0.609	**23.80**/**0.832**	22.62/0.870
PromptIR	20.78/0.794	21.48/*0.821*	18.57/0.702	20.85/0.798	*22.24*/*0.857*
DEA-Net	*20.47*/*0.789*	21.11/*0.813*	18.65/0.704	20.38/0.788	21.74/0.852
MB-TaylorFormer	20.34/*0.789*	*21.19*/0.811	18.73/0.714	19.62/0.779	21.80/0.853
UDPNet	20.16/0.775	19.03/0.787	16.17/0.665	23.50/0.827	21.95/0.821
PPG-Net	**23.11**/**0.829**	**24.14**/**0.850**	**22.06**/**0.773**	*21.17*/*0.822*	**25.07**/**0.871**

D. Controlled Ablation and Fusion Sensitivity

Table III separates two questions. The first four retrained controls compare input information under the same architecture and 1680-update protocol: total intensity only, raw four-angle intensity, explicit Stokes/DoLP/AoP without coarse priors, and the complete tensor without the coarse transmission and reconstruction. A fifth control removes transmission supervision; its direct image output is evaluated without the associated inversion branch. The final two rows compare the direct and fused outputs of the complete model.

TABLE III
CONTROLLED RETRAINED ABLATIONS AND PPG-NET OUTPUTS

Variant	Input or change	Training	PSNR↑	SSIM↑	t MAE↓
Input only	Unpolarized hazy image	None	15.798	0.7093	—
Coarse physical prior	J_{prior} only	None	13.691	0.4517	0.2004
S_0-only	Channels 13–15; all others zero	Matched	20.350	0.7856	0.1206
Raw four-angle	$I_{45}/I_{90}/I_{135}/I_{180}$ only	Matched	22.232	0.7943	*0.0841*
Explicit polar features	S_0, Q, U, DoLP, AoP; no coarse priors	Matched	22.279	0.7988	0.0852
No coarse priors	Full tensor except J_{prior} and t_0	Matched	*22.555*	*0.8127*	0.0825
No t supervision	Direct; transmission loss weight 0	Matched	**24.196**	**0.8321**	0.2265
PPG-Net raw	Complete tensor; direct network output	Matched	22.454	0.7816	**0.0786**
PPG-Net	Complete tensor; physics-guided fusion	Matched	23.114	0.8290	**0.0786**

Table III shows that raw four-angle input outperforms total intensity alone, while explicit Stokes and polarization features provide a smaller additional gain. Fusion improves both image metrics over the complete model's direct output. The physical constraints nevertheless involve *a* tradeoff: removing coarse priors improves direct image scores but slightly worsens transmission MAE, and removing transmission supervision gives the highest image scores but the least accurate learned transmission. Joint supervision thus favors transmission accuracy rather than maximizing image metrics alone. These observations are subject to the single-seed and batch-normalization-state limitations in Section V-A.

TABLE IV
SENSITIVITY OF THE PHYSICS-GUIDED FUSION WEIGHTS

ID	w_N	w_P	w_D	N	PSNR↑	SSIM↑
1	1.00	0.00	0.00	720	22.739	0.8034
2	0.80	0.10	0.10	720	**23.724**	0.8399
3	0.70	0.20	0.10	720	23.697	0.8453
4	0.70	0.30	0.00	720	23.225	0.8296
5	0.70	0.00	0.30	720	23.136	*0.8487*
6	0.65	0.20	0.15	720	*23.626*	**0.8501**
7	0.60	0.25	0.15	720	23.432	0.8493
8	0.55	0.30	0.15	720	23.167	0.8467
9	0.50	0.35	0.15	720	22.840	0.8423
10	0.45	0.40	0.15	720	22.462	0.8362

Table IV varies only the three inference-time fusion weights, which sum to one, on the independent 720-image validation set. The mean is image-weighted across the 720 validation samples; the four sources are not equally represented. Network parameters remain fixed; no test image is used for weight selection.

TABLE V
TRANSMISSION ESTIMATION AGAINST SYNTHETIC GROUND TRUTH

Source	Estimator	MAE↓	RMSE↓	SSIM↑
DIODE	t_0	0.2523	0.2810	0.5665
DIODE	t_N	**0.0798**	**0.0955**	**0.6338**
Middlebury	t_0	0.1985	0.2187	0.7431
Middlebury	t_N	**0.0843**	**0.1016**	**0.7863**
NYU Depth V2	t_0	0.1839	0.2320	0.5698
NYU Depth V2	t_N	**0.0721**	**0.0903**	**0.6124**
TUM RGB-D	t_0	0.1668	0.2475	**0.6003**
TUM RGB-D	t_N	**0.0783**	**0.0969**	0.5773
All	t_0	0.2004	0.2448	0.6199
All	t_N	**0.0786**	**0.0961**	**0.6525**

Table IV shows that fusion improves on the direct-output setting over *a* range of weights. ID 2 has the highest validation PSNR, whereas ID 6 has the highest SSIM, indicating a tradeoff between pixel error and structural agreement. The recorded test setting, ID 7, is retained to keep the reported outputs unchanged; it is not validation-optimal, since ID 6 scores higher on both metrics. Numerical emphasis denotes metric ranks, not the choice of setting.

In Table V, the learned transmission has lower MAE and RMSE than the coarse prior on every source dataset, with both overall errors reduced by about 61%. The network therefore corrects substantial numerical error in the initial estimate. Transmission SSIM improves overall but decreases on TUM RGB-D, showing that lower absolute error does not guarantee better local structural agreement.

Validation-only analysis also identifies the sensitivity of the coarse prior. Increasing the scalar-DoLP normalization percentile from 97 to 99.5 reduces t_0 MAE from 0.2240 to 0.1994; raising the transmission floor from 0.10 to 0.25 reduces it from 0.2158 to 0.1777. The latter trend reflects the

simulator's target-transmission floor of 0.25. At the preset floor of 0.15, airlight-fraction caps of 0.85–0.95 produce identical clipped transmission. The preset percentile 99, cap 0.92, and floor 0.15 are retained without test-set retuning.

E. Frozen-Detector Diagnostic

Remote-sensing image enhancement should improve radiometric fidelity while retaining semantic structure for subsequent information extraction. Because the test set has no object-level annotations, we do not evaluate detection accuracy. Instead, we use a frozen detector to determine whether PPG-Net changes the responses of a fixed downstream model. Specifically, we use the torchvision SSDLite320-MobileNetV3 Large detector pretrained on COCO [60], with a MobileNetV3-Large backbone [61] and an SSD single-stage detection head [62]. The detector is not fine-tuned and remains in evaluation mode. The 480 hazy/PPG-Net pairs, or 960 images, use identical preprocessing, batch size, and a confidence threshold of 0.35. We record the number of above-threshold candidate boxes and their mean class confidence, as reported in Fig. 12 and Table VI. In Table VI, emphasis compares mean confidence within each pair, not detection accuracy; box counts are not ranked.

The six examples in Fig. 12 were selected because box counts did not decrease, mean class confidence increased, and the candidate boxes corresponded to visually recognizable objects. They therefore illustrate favorable cases rather than typical performance. Across all 480 pairs, the changes are more modest: box counts increase in 114 images and decrease in 75, while mean confidence rises only slightly. This suggests that dehazing can make some objects more apparent to the fixed detector, but the response is not uniformly beneficial. Without object-level ground truth, additional boxes or higher confidence cannot establish improved detection or localization accuracy.



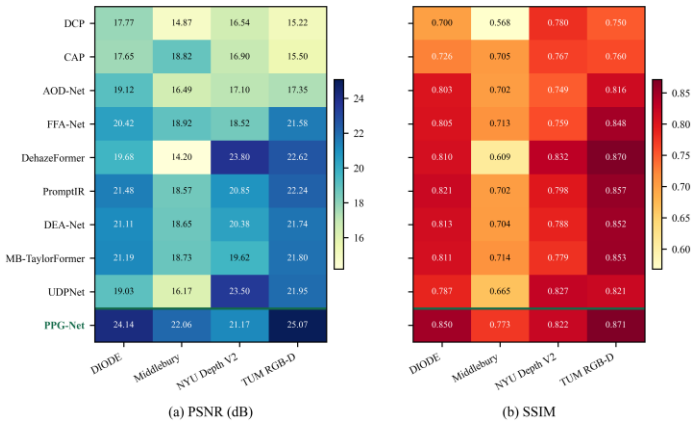
Fig. 12. Frozen-detector responses before and after PPG-Net dehazing. Top: hazy inputs. Bottom: PPG-Net outputs.

TABLE VI
FROZEN-DETECTOR CANDIDATE RESPONSES BEFORE AND AFTER PPG-NET DEHAZING

Case	Sample	Objects	Hazy boxes/ image	Hazy mean score	PPG-Net boxes/ image	PPG-Net mean score	Score difference
1	scene_ 000324_v02	refrigerator	0	0.000	1	**0.849**	+0.849
2	scene_ 000852_v04	banana	0	0.000	1	**0.892**	+0.892
3	scene_ 001063_v02	chair, tv	1	0.384	3	**0.695**	+0.311
4	scene_ 001359_v03	couch	0	0.000	1	**0.632**	+0.632
5	scene_ 001719_v01	tv	1	0.452	1	**0.724**	+0.272
6	scene_ 001759_v03	keyboard, tv	3	0.557	3	**0.677**	+0.119

F. Multidimensional Quantitative Analysis and Robustness to Haze Density

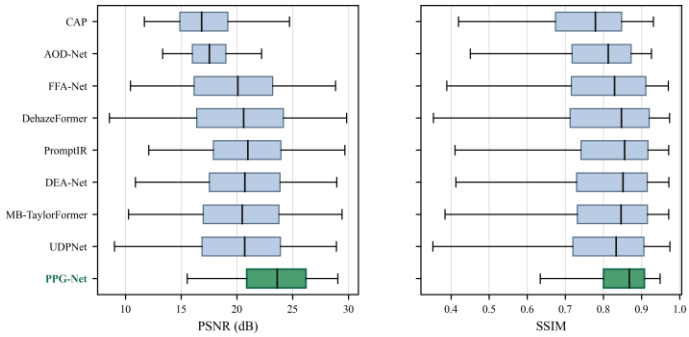
To prevent overall averages from obscuring dataset differences and difficult samples, we further compare the per-image distributions and haze-density-stratified results of DCP, CAP, AOD-Net, FFA-Net, DehazeFormer, PromptIR, DEA-Net, MB-TaylorFormer, UDPNet, and PPG-Net on the same 480 fixed test images. Samples are ranked by the physical attenuation implied by the generation parameters and divided equally into light-, moderate-, and dense-haze groups of 160 images each. Results are shown in Figs. 13–15.



(a) PSNR (dB) (b) SSIM

Fig. 13. PSNR and SSIM across methods and source datasets.

Figure 13 shows that the overall advantage is source dependent. PPG-Net has the highest mean PSNR and SSIM on DIODE, Middlebury, and TUM RGB-D. On NYU Depth V2, however, both metrics are lower than those of DehazeFormer and UDPNet. The aggregate gain therefore reflects strong performance on three sources, not a uniform advantage across all scene distributions.



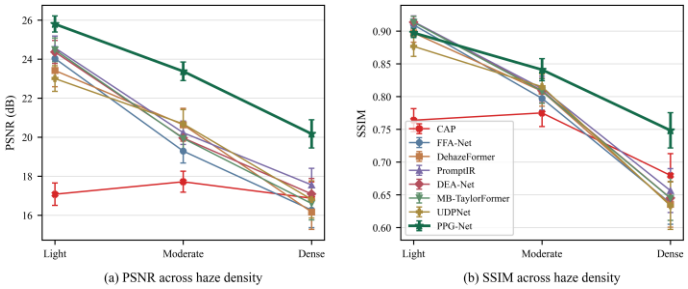
(a) Per-image PSNR distribution (b) Per-image SSIM distribution

Fig. 14. Per-image PSNR and SSIM distributions over the image test set.

In Fig. 14, PPG-Net's median and lower-quartile PSNR exceed those of PromptIR, so the overall advantage is not confined to *a* few high-scoring images. Median SSIM is also higher, although upper-quartile SSIM is lower. These are comparisons of marginal score distributions, not paired gains; they indicate an improvement in the center of the distribution without implying that every image benefits.

The clearest advantage in Fig. 15 occurs under moderate and dense haze, where PPG-Net has higher mean PSNR and SSIM than the comparison methods. This is consistent with polarization-derived scattering cues being useful when attenuation obscures scene structure. Under light haze, PPG-

Net retains a PSNR advantage but has lower SSIM than PromptIR, indicating less consistent structural benefit when degradation is mild. They are nominal image-level intervals; correlated haze variants and related frames preclude *a* physical-scene-level significance interpretation.



(a) PSNR across haze density (b) SSIM across haze density

Fig. 15. PSNR and SSIM across simulated haze densities.

G. Out-of-Domain Real Polarimetric Transfer

Six selected real polarimetric examples illustrate contour changes: four from the dense-haze subset and two from the heavy-haze subset of HFUT [63] (0°, 60°, 120°). Equation (4) converts each measured triplet to the four network input angles; these are not independent observations. A common gain is applied before clipping. The frozen synthetic-trained model uses 256 × 256 inputs and fusion weights (0.60, 0.25, 0.15). Outputs are resized to the input dimensions without display-only enhancement.



(a) Real hazy images

(b) PPG-Net outputs

Fig. 16. Real polarimetric examples from T. Liao's HFUT dataset [63] (CC BY 4.0).

In Fig. 16, roof edges, building outlines, and construction-site structures are more distinct after restoration. The crane lattice in column 4 and the facade and tree boundaries in column 5 also become easier to distinguish. Increased visibility is accompanied by residual haze, color shifts, darkening, and uneven brightness, however. These selected pairs show both contour enhancement and transfer artifacts. They do not establish dataset-wide superiority because aligned clear references are unavailable and the input angles are converted using an ideal Stokes model.

H. Low-Level Feature Correspondence Diagnostic

To examine whether dehazing exposes local structures usable by remote-sensing registration, mosaicking, or multitemporal analysis, we conduct an ORB [64] feature-correspondence probe on all 480 aligned hazy/PPG-Net/clear triplets. Up to 600 ORB keypoints are detected in each image. Binary descriptors are matched using a 0.80 ratio threshold and one-to-one assignment, and a match is considered geometrically consistent when its endpoint displacement from the aligned clear reference is no greater than 3 pixels. Because the synthetic haze process changes radiometry without changing scene geometry, this protocol measures repeatable local structure rather than image appearance alone. It is a task-interface probe, not a substitute for complete image-registration or remote-sensing task evaluation.

Table VII shows an 11.73% increase in the mean number of geometrically consistent correspondences after restoration, with 57.9% of test images improving. The gain is therefore common but not universal. The reported interval and two-sided Wilcoxon test treat image pairs as independent; correlated haze variants and related frames make them exploratory image-level summaries rather than scene-level significance tests.

TABLE VII
ORB CORRESPONDENCE TO THE ALIGNED CLEAR REFERENCE

Dataset (*n*)	Hazy	PPG-Net	Delta matches [nominal 95% CI]	*p* (image-level)
DIODE (120)	163.10	**192.13**	+29.03 [+14.15, +43.92]	0.0015
Middlebury (120)	30.11	**36.76**	+6.65 [+0.26, +13.04]	<0.001
NYU Depth V2 (120)	154.50	**212.35**	+57.85 [+47.02, +68.68]	<0.001
TUM RGB-D (120)	**352.73**	341.33	-11.39 [-22.68, -0.10]	0.0064
Overall (480)	175.11	**195.64**	+20.54 [+14.46, +26.61]	<0.001

Figure 17 illustrates additional repeatable correspondences along recovered edges and textured regions in DIODE, Middlebury, and NYU Depth V2. The source averages in Table VII follow this trend, whereas the mean count decreases on TUM RGB-D. The normalized correspondence rate changes little and provides no evidence of improvement in the exploratory test. The observed benefit is therefore a higher absolute number of usable correspondences overall.

I. Relevance to Atmospherically Degraded Optical Remote Sensing and Validation Scope

Restored optical images may support registration, segmentation, and detection. Cross-domain instance segmentation [37] and referring remote-sensing segmentation [65] are relevant examples of image-interpretation tasks. The SAR registration study [66] illustrates feature matching in a

different modality, not validation of optical dehazing. PPG-Net also provides a transmission estimate that could support quality assessment after radiometric calibration. Establishing these uses requires calibrated observations and task-level ground truth.

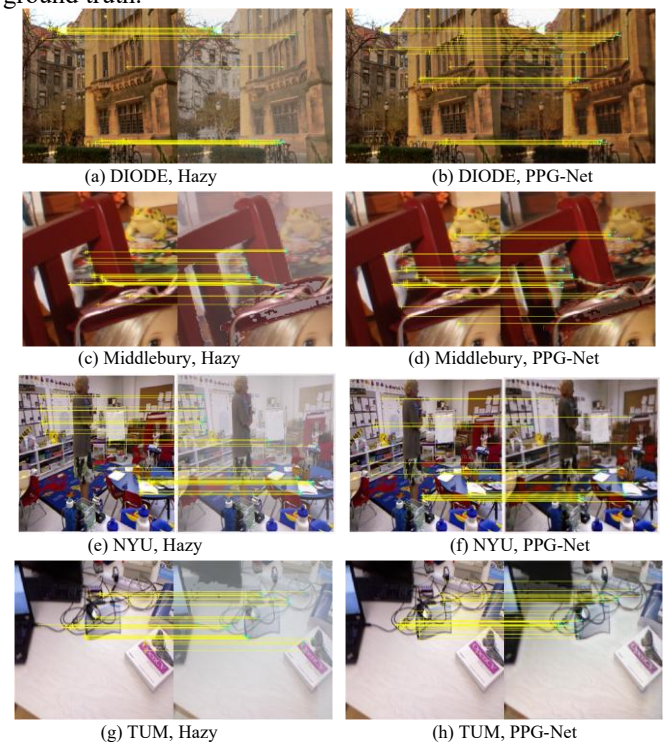


Fig. 17. Representative ORB correspondences to the clear reference, shown on the right half of each panel.

The present evidence is limited to synthetic terrestrial RGB-D data and selected real polarimetric examples. Depth normalization and a homogeneous-atmosphere model do not reproduce kilometer-scale propagation. Further validation should use calibrated polarimetric observations, account for sensor spectral response, and separate physical scenes or sequences across data splits. Task labels, runtime measurements, and cross-sensor tests are also needed to assess practical near-surface remote-sensing use.

VI. DISCUSSION AND CONCLUSION

PPG-Net combines four-angle polarization measurements with joint scene-radiance and transmission estimation and physics-guided output fusion. On the synthetic benchmark from four RGB-D sources, it achieves 23.11 dB PSNR and 0.8290 SSIM. Splits are disjoint at the source RGB-D image level, although related frames may remain across splits. Controlled comparisons support the value of angular measurements and show a tradeoff between image fidelity and transmission accuracy; fusion improves the complete model's direct restoration. ORB analysis finds more geometrically consistent correspondences overall, while the frozen-detector results show only changes in candidate responses. Real polarimetric examples reveal clearer contours alongside color and brightness artifacts. The findings support polarization-guided restoration under the tested conditions, with potential for near-surface optical remote sensing. The remaining limitations include synthetic training data, different training corpora for external baselines, single-seed controls with the batch-normalization caveat in Section V-A, and the lack of calibrated real remote-sensing observations and task labels. Future work should prioritize real polarimetric data and scene-disjoint evaluation, followed by radiometric and task-level validation across sensors.

REFERENCES

[1] W. Wang, X. Yuan, X. Wu, and Y. Liu, “Fast image dehazing method based on linear transformation,” *IEEE Trans. Multimedia*, vol. 19, no. 6, pp. 1142–1155, 2017.
[2] S. G. Narasimhan and S. K. Nayar, “Vision and the atmosphere,” *Int. J. Comput. Vis.*, vol. 48, no. 3, pp. 233–254, 2002.
[3] R. T. Tan, “Visibility in bad weather from a single image,” in *Proc. IEEE CVPR*, 2008, pp. 1–8.
[4] J.-P. Tarel and N. Hautière, “Fast visibility restoration from a single color or gray level image,” in *Proc. IEEE ICCV*, 2009, pp. 2201–2208.
[5] C. O. Ancuti and C. Ancuti, “Single image dehazing by multi-scale fusion,” *IEEE Trans. Image Process.*, vol. 22, no. 8, pp. 3271–3282, 2013.
[6] S. G. Narasimhan and S. K. Nayar, “Contrast restoration of weather degraded images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 6, pp. 713–724, 2003.
[7] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, “Polarization-based vision through haze,” *Appl. Opt.*, vol. 42, no. 3, pp. 511–525, 2003.
[8] 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, 2011.
[9] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, “Efficient image dehazing with boundary constraint and contextual regularization,” in *Proc. IEEE ICCV*, 2013, pp. 617–624.
[10] R. Fattal, “Dehazing using color-lines,” *ACM Trans. Graph.*, vol. 34, no. 1, pp. 1–14, 2014.
[11] 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, 2015.
[12] D. Berman, T. Treibitz, and S. Avidan, “Non-local image dehazing,” in *Proc. IEEE CVPR*, 2016, pp. 1674–1682.
[13] P. Ling, H. Chen, X. Tan, Y. Jin, and E. Chen, “Single image dehazing using saturation line prior,” *IEEE Trans. Image Process.*, vol. 32, pp. 3238–3253, 2023.
[14] M. Liao, Y. Lu, X. Li, S. Di, W. Liang, and V. Chang, “An unsupervised image dehazing method using patch-line and fuzzy clustering-line priors,” *IEEE Trans. Fuzzy Syst.*, vol. 32, no. 6, pp. 3381–3395, 2024.
[15] L. He, Z. Yi, P. Li, S. Wang, C. Chen, and M. Lu, “Efficient single image dehazing based on gradient line prior,” *IEEE Trans. Multimedia*, vol. 28, pp. 4596–4609, 2026.
[16] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, “DehazeNet: An end-to-end system for single image haze removal,” *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187–5198, 2016.
[17] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, “Single image dehazing via multi-scale convolutional neural networks,” in *Proc. ECCV*, 2016, pp. 154–169.
[18] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “AOD-Net: All-in-one dehazing network,” in *Proc. IEEE ICCV*, 2017, pp. 4780–4788.
[19] H. Zhang and V. M. Patel, “Densely connected pyramid dehazing network,” in *Proc. IEEE CVPR*, 2018, pp. 3194–3203.
[20] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M.-H. Yang, “Gated fusion network for single image dehazing,” in *Proc. IEEE CVPR*, 2018, pp. 3253–3261.
[21] B. Li et al., “Benchmarking single-image dehazing and beyond,” *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 492–505, 2019.
[22] Y. Qu, Y. Chen, J. Huang, and Y. Xie, “Enhanced Pix2pix dehazing network,” in *Proc. IEEE CVPR*, 2019, pp. 8152–8160.
[23] X. Liu, Y. Ma, Z. Shi, and J. Chen, “GridDehazeNet: Attention-based multi-scale network for image dehazing,” in *Proc. IEEE ICCV*, 2019, pp. 7313–7322.
[24] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, “FFA-Net: Feature fusion attention network for single image dehazing,” in *Proc. AAAI*, vol. 34, no. 7, 2020, pp. 11908–11915.
[25] H. Dong et al., “Multi-scale boosted dehazing network with dense feature fusion,” in *Proc. IEEE CVPR*, 2020, pp. 2154–2164.

[26] Y. Shao, L. Li, W. Ren, C. Gao, and N. Sang, "Domain adaptation for image dehazing," in *Proc. IEEE CVPR*, 2020, pp. 2805–2814.
[27] H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, and L. Ma, "Contrastive learning for compact single image dehazing," in *Proc. IEEE CVPR*, 2021, pp. 10546–10555.
[28] Z. Chen, Y. Wang, Y. Yang, and D. Liu, "PSD: Principled synthetic-to-real dehazing guided by physical priors," in *Proc. IEEE CVPR*, 2021, pp. 7176–7185.
[29] J. M. J. Valanarasu, R. Yasarla, and V. M. Patel, "TransWeather: Transformer-based restoration of images degraded by adverse weather conditions," in *Proc. IEEE CVPR*, 2022, pp. 2343–2353.
[30] 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.
[31] 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. IEEE ICCV*, 2023, pp. 12756–12767.
[32] T. Wang, K. Zhang, Z. Shao, W. Luo, B. Stenger, T. Lu, T.-K. Kim, W. Liu, and H. Li, "GridFormer: Residual dense transformer with grid structure for image restoration in adverse weather conditions," *Int. J. Comput. Vis.*, vol. 132, no. 10, pp. 4541–4563, 2024.
[33] J. Xiong, X. Yan, Y. Wang, W. Zhao, X.-P. Zhang, and M. Wei, "RSHazeDiff: A unified Fourier-aware diffusion model for remote sensing image dehazing," *IEEE Trans. Intell. Transp. Syst.*, vol. 26, no. 1, pp. 1055–1070, 2025.
[34] Y. Su, N. Wang, Z. Cui, Y. Cai, C. He, and A. Li, "Real scene single image dehazing network with multi-prior guidance and domain transfer," *IEEE Trans. Multimedia*, vol. 27, pp. 5492–5506, 2025.
[35] P. Saxena, A. Kumar Tiwari, and M. Narwaria, "Adaptive self-attention enhanced conditional GAN for image dehazing," *Pattern Recognit.*, vol. 177, Art. no. 113277, 2026.
[36] Y. Yang, L. Zhang, K. Pang, T. Chen, and X. Yue, "Semantic-assisted unpaired image dehazing," *Image Vis. Comput.*, vol. 165, Art. no. 105818, 2026.
[37] J. Wang, S. Ji, and T. Zhang, "From image transfer to object transfer: Cross-domain instance segmentation based on center point feature alignment," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, Art. no. 4407011, pp. 1–11, 2023.
[38] Y. Wen, T. Gao, S. Liang, D. Zhang, Z. Li, J. Qin, and T. Chen, "When optimal transport meets photo-realistic image dehazing with unpaired training," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 37, no. 9, pp. 4755–4769, 2026, doi: 10.1109/TNNLS.2026.3673760.
[39] Y. Wen, T. Gao, Z. Li, J. Zhang, K. Zhang, and T. Chen, "All-in-one weather-degraded image restoration via adaptive degradation-aware self-prompting model," *IEEE Trans. Multimedia*, vol. 27, pp. 3343–3355, 2025, doi: 10.1109/TMM.2025.3535316.
[40] Y. Wen, T. Gao, T. Chen, Z. Li, M. Liu, and L. Liu, "Cross-level interaction and intralevel fusion network for remote sensing image dehazing," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, Art. no. 5626214, pp. 1–14, 2025, doi: 10.1109/TGRS.2025.3576979.
[41] Y. Wen, T. Gao, Z. Li, Q. Zhang, J. Zhang, T. Chen, and L. Liu, "Structure-preserving frequency-regularized text-guided optimal transport for unpaired rain streaks and raindrops removal," *IEEE Trans. Multimedia*, vol. 28, pp. 3380–3395, 2026, doi: 10.1109/TMM.2026.3651132.
[42] L. Shen, Y. Zhao, Q. Peng, J. C. W. Chan, and S. G. Kong, "An iterative image dehazing method with polarization," *IEEE Trans. Multimedia*, vol. 21, no. 5, pp. 1093–1107, 2019.
[43] Y. Wang, Y. Yin, X. Yang, Q. Sun, X. Meng, G. Shentu, and B. Sun, "Polarimetric dehazing based on fusing intensity and degree of polarization," *Opt. Laser Technol.*, vol. 156, Art. no. 108584, 2022.
[44] H. Suo, J. Guan, M. Ma, Y. Huo, Y. Cheng, N. Wei, and L. Zhang, "Dynamic dark channel prior dehazing with polarization," *Appl. Sci.*, vol. 13, no. 18, Art. no. 10475, 2023.
[45] C. Yan, X. Zhang, X. Wang, G. Jiao, and H. He, "A novel simulation for polarization dehazing," *IEEE Signal Process. Lett.*, vol. 31, pp. 341–345, 2024.
[46] X. Wang, Y. Chen, C. Guan, L. Zhu, K.-H. Chew, and R.-P. Chen, "Adaptive learning-driven polarimetric dehazing imaging in a dynamic turbid environment," *J. Opt. Soc. Am. A*, vol. 42, no. 5, pp. 672–678, 2025.
[47] S. Chen, S. Qiao, L. Dong, and R. Chen, "Polarization-aware deep learning inference for dehazing images in the wild," *Opt. Laser Technol.*, vol. 203, Art. no. 115804, 2026.
[48] C. Zhou, M. Teng, Y. Han, C. Xu, and B. Shi, "Learning to dehaze with polarization," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 34, pp. 11487–11500, 2021.
[49] O. Dubovik et al., "Polarimetric remote sensing of atmospheric aerosols: Instruments, methodologies, results, and perspectives," *J. Quant. Spectrosc. Radiat. Transf.*, vol. 224, pp. 474–511, 2019, doi: 10.1016/j.jqsrt.2018.11.024.
[50] A. A. Kokhanovsky et al., "Space-based remote sensing of atmospheric aerosols: The multi-angle spectro-polarimetric frontier," *Earth-Sci. Rev.*, vol. 145, pp. 85–116, 2015.
[51] J.-L. Deuzé et al., "Remote sensing of aerosols over land surfaces from POLDER-ADEOS-1 polarized measurements," *J. Geophys. Res. Atmos.*, vol. 106, no. D5, pp. 4913–4926, 2001.
[52] I. Vasiljevic, N. Kolkin, S. Zhang, R. Luo, H. Wang, F. Z. Dai, A. F. Daniele, M. Mostajabi, S. Basart, M. R. Walter, and G. Shakhnarovich, "DIODE: A dense indoor and outdoor depth dataset," arXiv preprint arXiv:1908.00463, 2019.
[53] D. Scharstein, H. Hirschmüller, Y. Kitajima, G. Krathwohl, N. Nešić, X. Wang, and P. Westling, "High-resolution stereo datasets with subpixel-accurate ground truth," in *Proc. German Conf. Pattern Recognit. (GCPR)*, 2014, pp. 31–42.
[54] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2012, pp. 746–760.
[55] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2012, pp. 573–580.
[56] Z. Zuo, J. Jiang, G. Wu, and X. Liu, "UDPNet: Unleashing depth-based priors for robust image dehazing," arXiv preprint arXiv:2601.06909, 2026.
[57] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004, doi: 10.1109/TIP.2003.819861.
[58] V. Potlapalli, S. W. Zamir, S. Khan, and F. S. Khan, "PromptIR: Prompting for all-in-one image restoration," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 36, 2023, pp. 71275–71293.
[59] 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.
[60] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft COCO: Common objects in context," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2014, pp. 740–755.
[61] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, "Searching for MobileNetV3," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2019, pp. 1314–1324.
[62] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 21–37.
[63] T. Liao, "HFUT_Polarization_Datasets.rar," *figshare*, Nov. 2022, doi: 10.6084/m9.figshare.21644927.v1.
[64] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2011, pp. 2564–2571, doi: 10.1109/ICCV.2011.6126544.
[65] Z. Yuan, L. Mou, Y. Hua, and X. X. Zhu, "RRSIS: Referring remote sensing image segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, Art. no. 5613312, 2024.
[66] S. Paul and U. C. Pati, "SAR image registration using an improved SAR-SIFT algorithm and Delaunay-triangulation-based local matching," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 2958–2966, 2019.



Wencheng Wang (Member, IEEE) received a Ph.D. degree from Shandong University, and now he is a professor in the College of Machinery and Automation at Weifang University, acting as the principal investigator of the innovation team of robot vision perception and control. His main research interests include computer vision, pattern recognition, and intelligent computing. From 2015 to 2016, he was a visiting scholar at the University of North Texas, engaged in the research of intelligent control, image dehazing, and enhancement technology. He has participated in more than ten scientific research projects and published more than 50 papers in academic journals and conferences, two book chapters, and 16 patents. He is the young expert with outstanding contributions in Weifang and has been awarded the excellent teacher for seven years.



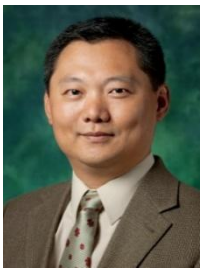
Lun Li, born in 1989, Doctor of Engineering. Graduated from Shenyang Polytechnic University in July 2020, majoring in Weapon Science and technology. Now he is an associate professor in the School of Information and Control Engineering of Weifang University, mainly engaged in the research of terahertz imaging detection technology and bionic robot control.



Hongtao Liu received a Ph.D. degree in electrical engineering from Shandong University, Jinan, China, in 2021. He is currently with the School of Machinery and Automation, Weifang University, Weifang, China. He has participated in six scientific research projects and published more than ten articles in academic journals and conferences. His research interests include modeling and analysis of physiological systems, physiological control of left ventricular assist devices, and research and application of intelligent algorithms.



Zhenxue Chen received the Ph.D. degree in pattern recognition and intelligent systems from Institute of Image Recognition and Artificial Intelligence at Huazhong University of Science and Technology, Wuhan, China, in 2007. From 2012 to 2013, he was a visiting scholar with the Michigan State University, East Lansing, Michigan, USA. He is currently a professor with the School of Control Science and Engineering, Shandong University. His main areas of interest include image processing, pattern recognition, and computer vision, with applications to face recognition. He has published over 100 papers in refereed international leading journals/conferences such as IEEE T-II, IEEE T-CSVT, IEEE T-IFS, IEEE T-VT, IEEE TITS, Information Sciences, Neurocomputing, Neural Computing and Applications, and SP-IC, etc.



Xiaohui Yuan (Senior Member, IEEE) received a Ph.D. degree in Computer Science from Tulane University in 2004. He worked as a postdoctoral researcher at the National Institutes of Health between 2004 and 2006 on medical image analysis. He joined the University of North Texas as an Assistant Professor in 2006 and was promoted to Associate Professor with tenure in 2012. His research interests include computer vision, data mining, machine learning, and artificial intelligence. He served as principal investigator in projects supported by the Air Force Lab, National Science Foundation, Texas Advanced Research Program, Oak Ridge Associated Universities, and UNT. His research findings are published in over 200 peer-reviewed papers. Dr. Yuan is a recipient of the Ralph E. Powe Junior Faculty Enhancement award in 2008 and the Air Force Summer Faculty Fellowship in 2011, 2012, and 2013. He also received two research awards and a teaching award from UNT in 2007, 2008, and 2012, respectively. He was featured in the 2020 Higher Education Review Magazine and EurekAlert! Science News 2017. He also has advised more than ten postdoctoral scholars.