photonics

MDPI

Article

Physics-Informed Deep Neural Network for Polarimetric Descattering Imaging in Dynamic Cement Dust Environments

Peikai Zhao, Chao Guan, Weiming Yuan, Liming Zhu, Khian-Hooi Chew and Rui-Pin Chen *

Zhejiang-Mexico Joint Laboratory of Intelligent Optoelectronic Sensing, Zhejiang Key Laboratory of Quantum State Control and Optical Field Manipulation, Department of Physics, Zhejiang Sci-Tech University, Hangzhou 310018, China

* Correspondence: chenrp@zstu.edu.cn

Abstract

Polarimetric descattering imaging has attracted growing interest due to its fundamental physical significance and potential applications. While deep learning has accelerated its development through powerful feature extraction and inference capabilities, existing methods still face limitations in practical scenarios, particularly under dynamic non-uniform scattering conditions such as cement dust environments. To address this, we propose a deep neural network based on the Mueller matrix model that effectively integrates polarization evolution information with deep learning. Specifically, local concentrations of the scattering medium in non-uniform cement dust are characterized by the evolution of the degree of linear polarization (DoLP), which is converted into pixel-wise weight biases to generate customized Mueller matrices adaptable to varying concentrations. The network predicts a pixel-wise dust concentration map and applies the corresponding concentration-specific Mueller matrix to each pixel for polarization-aware dehazing, ensuring physical consistency with Mueller matrix calculus throughout inference. This framework is further enhanced by a physics-constrained optimization loss and multi-scale feature fusion. Experimental results demonstrate the method's effectiveness and superiority in diverse dynamic non-uniform cement dust environments.

Keywords: deep learning; Mueller matrix; non-uniform scattering; cement dust

1. Introduction

Polarimetric descattering imaging and detections have garnered increasing interest due to fundamental physics and potential applications [1–3]. However, current descattering methods face many challenges in practical application scenarios [4], particularly under dynamic non-uniform scattering conditions such as cement dust environments. The presence of cement dust, a dense suspension of solid particles with a particle radius range of 1–50 μm [5], in confined engineering environments severely reduces imaging visibility. Unlike atmospheric haze, cement dust produces distinctive scattering effects, including non-uniform blurring and localized halo artifacts, which significantly impair operational safety and real-time monitoring [6]. Conventional dehazing methods have difficulty capturing such non-uniform, angle-dependent physics. In this work, the Mueller matrix model that describes the evolution of polarization state in various media is employed for dehazing imaging under cement dust environments without any assumptions, providing a robust, physics-based descattering imaging framework for handling diverse scattering scenarios.

Check for updates

Received: 2 March 2026
Revised: 5 April 2026
Accepted: 11 April 2026
Published: 15 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Many physical model-based de-scattering imaging methods have been developed. As a representative approach, the dark channel prior (DCP) [7] estimates transmission by leveraging the statistical property that haze-free images contain dark pixels in local patches. DCP is not adaptive for some scenes such as sky regions and low-light indoor environments where its fundamental assumption becomes invalid. Contrast-limited adaptive histogram equalization (CLAHE) [8] divides the image into small contextual regions and applies histogram equalization within each region independently. The automatic color enhancement (ACE) algorithm [9] employs a more sophisticated mechanism by estimating scene reflectance through the maximization of local color differences. This algorithm computes a series of pixel-wise comparisons to enhance local contrast. Nevertheless, these methods rely heavily on empirical assumptions or localized processing heuristics. As a result, they exhibit limited adaptability to non-uniform scattering conditions and frequently fail under challenging imaging scenarios.

The rapid development of deep learning has opened new avenues for descattering imaging. DehazeNet [10] pioneered the deep learning approach for image dehazing by designing a convolutional neural network to explicitly learn the mapping between hazy and clear images. A densely connected pyramid dehazing network [11] integrates dense connectivity with multi-scale pyramid processing to effectively handle extreme haze conditions. UWGAN [12] represents a specialized approach for underwater descattering, utilizing cycle-consistent adversarial networks to handle the unique color cast and scattering characteristics of water environments. However, such data-driven paradigms require large amounts of paired training data and lack physical interpretability because of their black-box nature. To address the limitations of purely data-driven methods, a new paradigm of physics-guided deep learning has emerged. AOD-Net [13] introduces a fundamental reformulation of the atmospheric scattering model by integrating transmission maps and atmospheric light into a unified learnable parameter K(x). This innovative approach allows the entire dehazing process to be encapsulated in a lightweight convolutional neural network (CNN) [14] that directly generates clear images from hazy inputs through matrix operations. ALPD-Net [2], a physics-guided deep learning framework that integrates multi-frame polarization features with a time-varying scattering model, addresses the limitations of conventional dehazing methods in scattering environments. In recent studies, the Polarization-Guided Stokes De-scattering (PGSD) framework proposed by Chen et al. enhances robustness in complex water environments by combining multi-parameter optimization and DoLP-gated airlight estimation [15]. Li et al. systematically reviewed the series, parallel, and hybrid cascade frameworks for underwater image restoration, and summarized the corresponding datasets and evaluation systems [16]. MTM-Net [17] embodies Mueller matrix formalism for polarization-based descattering. The network learns to estimate Mueller matrix parameters that characterize the scattering media and uses this physical representation to guide the descattering process. Prior physics-informed models, such as MTM-Net, typically assume a globally uniform scattering medium and obtain consistent Mueller matrix transformation parameters across the entire image despite the possibility of non-uniform concentration distributions. The concentration-driven Mueller matrix network (CDMMNet) overcomes this constraint by incorporating a concentration estimation (CE) module that explicitly derives a pixel-wise concentration map from polarimetric information. This enables the framework to apply spatially varying, concentration-aware Mueller matrices to each local area, ensuring physical consistency and robustness against the non-uniform concentration characteristic of industrial environments.

Motivated by the distinct challenges of cement dust and the limitations of prior art, we propose a novel concentration-driven deep neural network, the concentration-driven Mueller matrix network (CDMMNet), based on polarization physics (Mueller matrix for-

malism) for descattering imaging in dynamic, non-uniform cement dust environments. This network incorporates a novel concentration estimation (CE) module that computes spatial concentration maps using degree of linear polarization (DoLP) from polarimetric images, and a concentration-enhanced modulation convolution (CEMC) module that generates concentration-specific weight biases to dynamically encode Mueller matrix transformations into network parameters, enabling pixel-wise adaptation to spatially and temporally varying dust concentrations. This concentration-driven Mueller matrix learning framework is further enhanced by a physics-constrained optimization loss function to enforce physical plausibility in the learned transformations. The encoder–decoder architecture with skip connections enables multi-scale feature fusion, while inverted residual blocks and channel attention mechanisms maintain computational efficiency. Validated on our dataset, experimental results demonstrate the superior performance of this concentration-aware CDMMNet compared to conventional atmospheric scattering-based methods, especially under real-world industrial cement dust scenarios.

2. Theoretical Basis and Network Model

The Stokes vector is used to describe the evolution of polarization information in cement dust environments. At each pixel location (x,y), this four-dimensional vector is defined as:

$$\boldsymbol{s}(\mathrm{x},\mathrm{y}) = \begin{bmatrix} s_0(x,y) \\ s_1(x,y) \\ s_2(x,y) \\ s_3(x,y) \end{bmatrix} = \begin{bmatrix} I_0(x,y) + I_{90}(x,y) \\ I_0(x,y) - I_{90}(x,y) \\ I_{45}(x,y) - I_{135}(x,y) \\ I_{Q,45}(x,y) - I_{Q,135}(x,y) \end{bmatrix}, \quad (1)$$

where $I_0(x,y)$, $I_{45}(x,y)$, $I_{90}(x,y)$, $I_{135}(x,y)$ represent the light intensity information at the corresponding pixel positions with polarization angles of 0°, 45°, 90° and 135°. $I_{Q,45}(x,y)$ corresponds to the intensity of right-handed circularly polarized light, and $I_{Q,135}(x,y)$ corresponds to the intensity of left-handed circularly polarized light [18]. $s_0(x,y)$ quantifies the total energy of light, regardless of polarization state. $s_1(x,y)$ describes the difference in intensity between horizontally and vertically polarized light. $s_2(x,y)$ quantifies the difference in intensity between 45° and 135° linearly polarized light, characterizing diagonal linear polarization properties. $s_3(x,y)$ describes the left and right circular polarization components. The circular polarization component $s_3(x,y)$ is neglected in this work. To validate the physical justification, we measure the ratio of the s_3 component to total energy and conduct an ablation study comparing performance with and without the s_3 channel. Experiment results indicate that the circular polarization component s_3 contributes less than 0.1% of total energy, the differences in the restoration performance with and without the s_3 channel are negligible, as presented in Section 3. The evolution of Stokes vectors passing through the cement dust environment can be described using the 3 × 3 Mueller matrix: $\boldsymbol{s}_{\mathrm{clear}}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{M}(\mathrm{x},\mathrm{y}) \bullet \boldsymbol{s}_{\mathrm{haze}}(x,y)$ [17], where $\boldsymbol{s}(x,y) = [s_0(x,y), s_1(x,y), s_2(x,y)]^{\mathrm{T}}$ and $\boldsymbol{M}(x,y)$ is a position-dependent Mueller transmission matrix.

The DoLP serves as a critical indicator for dust concentration due to its sensitivity to scattering properties [19]. As the evolution of polarization state in a scattering medium sensitively depends on the property of the scattering medium, the DoLP value at each pixel in the captured image reflects localized scattering conditions, meaning the DoLP value at any point directly encodes the local dust density and scattering geometry. Thus, DoLP is adopted here as a concentration indicator under the inhomogeneous cement dust

environment. Therefore, the local dust concentration information can be evaluated based on the evolution of the DoLP:

$$c(x,y) = \mathrm{DoLP}(x,y) = \frac{\sqrt{s_1(x,y)^2 + s_2(x,y)^2}}{s_0(x,y)}. \tag{2}$$

Two pixel-specific modulation parameters are generated according to the evolution of DoLP through a lightweight convolutional neural network (ConvNet) [14]:

$$[\gamma(x,y), \beta(x,y)] = Split(ConvNet[\mathrm{c}(x,y)]), \tag{3}$$

where "*Split*" operation is defined as a channel-wise dimensional splitting that partitions an input feature tensor $C(x,y) \in \mathbb{R}^{H\times W\times C}$ (where H and W are spatial dimensions, C is the channel dimension):

$$Split : \mathbb{R}^{H\times W\times C} \to \prod_{m=1}^{M} \mathbb{R}^{H\times W\times C_m}, \qquad such\ that \sum_{m=1}^{M} C_m = C \tag{4}$$

"*ConvNet*" is a composite mapping of L convolutional layers and activation functions. For layer l, let K^l be the convolution kernel, σ^l the activation function, and $*$ the convolution operation. The composite mapping is:

$$ConvNet(X) = \sigma^L\left(K^L * \sigma^{L-1}\left(K^{L-1} * \cdots * \sigma^l\left(K^l * X\right)\cdots\right)\right). \tag{5}$$

The $\gamma(x,y)$ and $\beta(x,y)$ are two modulation parameters. $\gamma(x,y)$ serves as the baseline modulation parameter, providing a fundamental reference for concentration modulation and anchoring the basic characteristics of light–matter interactions that match the local dust concentration. In contrast, $\beta(x,y)$ acts as a fine-tuning parameter, enabling dynamic correction based on the baseline to adapt to subtle fluctuations or local specificities in the concentration distribution. Working in synergy, the two parameters not only avoid parameterization redundancy but also enhance adaptability to complex concentration fields through hierarchical modulation. Specifically, the ConvNet acts as a physical-to-parameter translator, bridging polarization physics and deep learning parameterization by learning the non-linear relationship between local dust concentration and the corresponding light–matter interaction characteristics. A Mueller matrix precisely matched to the specific non-uniform concentration field can be achieved on a pixel-wise basis:

$$\boldsymbol{M}(x,y) = \mathbf{W} \odot (1 + \sigma(\gamma(x,y))) + \alpha \cdot tanh(\beta(x,y)), \tag{6}$$

where $\boldsymbol{M}(x,y)$ denotes the concentration-aware Mueller matrix at pixel (x,y); $\mathbf{W}$ is the base weight matrix, which embeds learnable physical priors of the Mueller matrix under average scattering conditions; $\odot$ represents the element-wise multiplication operation; $\sigma(\cdot)$ is the sigmoid activation function; $\alpha = 0.1$ is the scaling factor; $\tanh(\cdot)$ is the hyperbolic tangent activation function. This formula is designed to achieve concentration-aware physical modulation while maintaining numerical stability throughout the training process. Specifically, the base weight matrix $\mathbf{W}$ provides a physically reasonable starting point for network optimization, encoding prior knowledge of the Mueller matrix under average scattering conditions. The concentration-driven modulation term $1 + \sigma(\gamma(x,y))$ modulates $\mathbf{W}$ via element-wise multiplication, enabling the matrix parameters to continuously adapt to concentration changes. The sigmoid activation function $\sigma(\cdot)$ constrains the modulation factor to a positive and bounded range, ensuring the physical realizability of the resulting matrix. The introduction of "1" in the modulation term $1 + \sigma(\gamma(x,y))$ acts as an identity bias, promot-

ing gradual optimization of the base matrix rather than destructive modification, thereby enhancing training stability. The additive bias term $\alpha \cdot \tanh(\beta(x,y))$ introduces fine-grained concentration-dependent adjustments to capture subtle scattering effects. The tanh function inherently constrains its output to the interval $(-1, 1)$, and the scaling factor $\alpha = 0.1$ further limits the adjustment range to $(-0.1, 0.1)$, preventing it from overwhelming the primary modulation mechanism. These designs restrict the impact of the additive term to a narrow operational range, acting as a targeted nonlinear correction to capture complex scattering behaviors and complement the multiplicative modulation's ability. This combined structure of multiplicative scaling and additive bias serves as a powerful yet constrained function approximator, enabling the Mueller matrix to dynamically specialize local concentration levels $\gamma(x,y)$ on a pixel-wise basis. The formula successfully integrates the physical priors of the base matrix **W** with the data-driven flexibility of the ConvNet, allowing the model to achieve generalization under diverse and dynamic scattering conditions in real-world industrial environments.

The polarization dehazing network architecture features a U-shaped structure that learns the Mueller transformation matrix directly from the input Stokes vectors in non-uniform cement dust concentration environments. The network employs a four-level encoder–decoder architecture, where the encoder progressively downsamples feature maps. At the same time, the decoder symmetrically upsamples them, as shown in Figure 1. The network extracts concentration distribution information through the CE module, performs concentration-aware feature modulation via the CEMC module, and accomplishes feature calibration and stable gradient transmission by relying on the NECK module (bottleneck layer). The network estimates the concentration distribution of the scattering medium from the input hazy image. Based on the concentration at each pixel location, it derives the weight biases $\gamma(x,y)$ and $\beta(x,y)$. These are then used to generate corresponding concentration-specific, pixel-wise Mueller matrices, which ultimately predict the dehazed image after several learning epochs with a specialized loss function. Meanwhile, this network framework integrates channel attention mechanisms and residual blocks to improve the dehazing ability in a non-uniform cement dust environment with low computational complexity.

The CE module extracts the polarization information from hazy images in dust environments with various concentration distributions and provides key parameter support for subsequent feature modulation. According to the characteristics of captured images, the 90°, 45°, 135°, and 0° linear polarization components of the $224 \times 224 \times 3$ input image are sequentially arranged into the 2×2 macro-blocks to obtain the Stokes vector $\boldsymbol{s}(x,y) = [s_0(x,y), s_1(x,y), s_2(x,y)]^{\mathrm{T}}$ where $s_0(x,y)$, $s_1(x,y)$, and $s_2(x,y)$ are all $224 \times 224 \times 3$ arrays, forming a complete Stokes vector with a dimension of $224 \times 224 \times 9$, as shown in Figure 1b. Since each element of this vector contains partial concentration information, the concentration-related features $c(x,y)$ are extracted from the DoLP image using the input Stokes vectors within the CE module. The resulting $224 \times 224 \times 1$ concentration distribution map as a key modulation parameter is transmitted to the CEMC module, providing a concentration basis for subsequent adaptive feature enhancement. The lightweight three-layer convolutional structure is adopted in CE: the first convolutional layer maps the DoLP information to basic features; the second convolutional layer further enhances feature expression after activation by the LeakyReLU functions; the third single-channel convolutional layer outputs the initial concentration distribution map, and a Sigmoid activation function is used for normalization to ensure that the concentration values strictly fall within the [0, 1] range.

The CEMC integrates with the U-shaped architecture by forming DCEMC modules within the encoder's downsampling path (where each stage reduces spatial resolution by half and doubles feature channels) and UCEMC modules within the decoder's upsampling

path (where each stage doubles spatial resolution and halves feature channels) to enable dynamic feature restoration. The concentration map $c(x,y)$ as an external conditional input is processed by a ConvNet processes $c(x,y)$, and a Split operation splits the concentration map $c(x,y)$ together with the output of concentration-related low-level features $\gamma(x,y)$ and $\beta(x,y)$, the baseline parameter $\gamma(x,y)$ and the dynamic fine-tuning parameter $\beta(x,y)$, enabling concentration-adaptive modulation in subsequent processing stages [20], thus sustaining concentration-aware dynamic feature modulation throughout the network's forward propagation. Meanwhile, the features from the corresponding levels of the encoder are fused with decoder features through skip connections to improve detail retention [21].

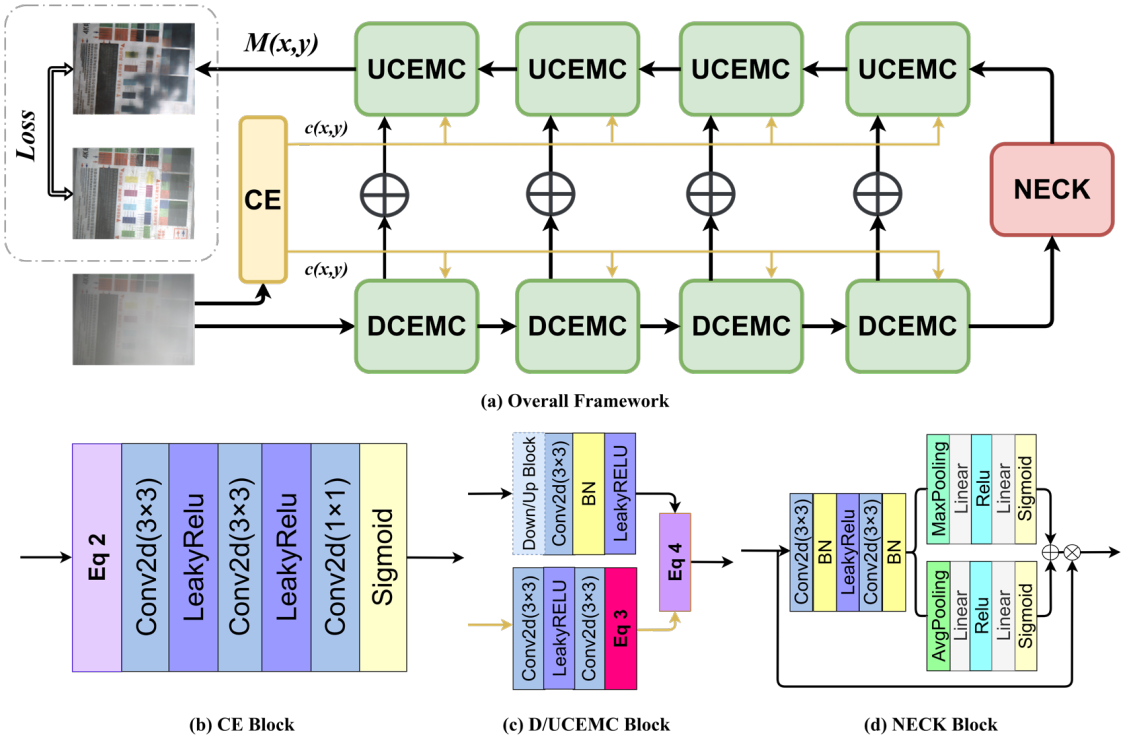


Figure 1. (**a**) Overall network architecture; (**b**) CE block is the unit for conditional signal generation in the network used to extract $c(x,y)$ based on the DoLP from polarized images; (**c**) D/UCEMC block is the core conditional enhancement unit in the network's sampling phase. It dynamically generates concentration modulation parameters γ and β in a pixel-wise manner based on $c(x,y)$; (**d**) NECK block connects the downsampling and upsampling phases and serves as the unit for feature purification and enhancement. Various modules are represented in different colors.

As a critical junction connecting the encoder's downsampling and the decoder's upsampling, the NECK module is located at the end of the encoder in the U-shaped network with a bottleneck layer, undertaking feature compression, calibration, and gradient stabilization. Feature compression is achieved through two successive 3×3 convolutional layers, which adopt a bias-free design paired with BatchNorm2d to maintain the feature receptive field while controlling the feature dimensions. Feature calibration relies on the channel attention module, which performs dual-path pooling (adaptive average + max pooling) on the dual 3×3 layers to extract dehazing-related features [22], generating channel-wise attention weights via fully connected layers. These weights are multiplied by the original features to strengthen effective channels and suppress noise, achieving precise feature distribution calibration. Gradient stabilization is ensured by the dual 3×3 residual block embedded with the attention module. A direct residual connection allows gradients to bypass complex operations and propagate directly to shallow layers during training, avoiding vanishing gradients caused by excessive depth.

To clarify the direction of model optimization, a triple-loss supervision mechanism is designed by integrating content loss, pixel loss, and perceptual loss. The content loss is the L1 distance between the VGG19 features of the output and target images:

$$L_{\text{content}} = \frac{1}{N}\sum_{i=1}^{N}\left\|\varphi_{18}\left(\mathrm{E}\left(x_{\text{hazy}}\right)\right)_i - \varphi_{18}(x_{\text{label}})_i\right\|_1, \tag{7}$$

where $\|\cdot\|_1$ represents the L1 norm, $\varphi_{18}(\cdot)$ is the feature map extracted from the 18th layer of a pre-trained VGG19 network, $\mathrm{E}(\cdot)$ denotes our proposed network, x_{label} is the ground truth clear image, x_{hazy} is the input hazy image, and N is the total number of pixels in the feature map. Here, the feature map from the 18th layer of the pre-trained VGG19 network is leveraged to maintain semantic consistency between the dehazed image and the clear reference image in the high-level feature spaces.

The pixel loss is formulated as the mean absolute error between the generated dehazed image and the ground truth image:

$$L_{\text{pixel}} = \frac{1}{N}\sum_{i=1}^{N}\left\|\mathrm{E}(x_{\text{hazy}})_i - x_{\text{label},i}\right\|_1 \tag{8}$$

This L1-norm-based loss function provides robust supervision at the pixel level for color distribution and structural contours between hazy and clear images [23], effectively preserving fine details.

Perceptual loss measures the discrepancy between the dehazed output and the clear ground-truth image in the high-level feature space:

$$L_{\text{perceptual}} = \frac{1}{C_j H_j W_j}\sum_{c=1}^{C_j}\sum_{h=1}^{H_j}\sum_{w=1}^{W_j}\left(\varphi_j\left(\mathrm{E}(x_{\text{hazy}})\right)_{c,h,w} - \varphi_j(x_{\text{label}})_{c,h,w}\right)^2, \tag{9}$$

where $\varphi_j(\cdot)$ denotes the feature extractor at the j-th layer of a pre-trained VGG-19 network, with the extracted feature map dimensions being $C_j \times H_j \times W_j$. The perceptual loss encourages the restored image to match the ground truth in terms of high-level feature representations [24]. The total loss is defined as:

$$L_{\text{total}} = \lambda_{\text{c}} L_{\text{content}} + \lambda_{\text{p}} L_{\text{pixel}} + \lambda_{\text{percept}} L_{\text{perceptual}}, \tag{10}$$

where λ_{c}, λ_{p}, λ_{percept} are weighting coefficients for the respective losses. Optimal performance is achieved with loss weights $\lambda_{\text{c}} = 0.5$, $\lambda_{\text{p}} = 0.3$, $\lambda_{\text{percept}} = 0.2$ through extensive experimentation. A sensitivity analysis with several representative loss function weight configurations is shown in Table 1; the proposed ratio (0.5:0.3:0.2) demonstrates the most stable convergence trend and superior performance across all metrics, effectively balancing high-level semantic recovery with low-level structural fidelity.

Table 1. Sensitivity analysis of loss function weight configurations regarding quantitative restoration performance.

λ_{c}:λ_{p}:λ_{percept}	PSNR	SSIM	LPIPS
0.5:0.3:0.2	**25.830**	**0.735**	**0.180**
0.7:0.2:0.1	22.140	0.587	0.200
0.33:0.33:0.33	24.790	0.698	0.196
0.2:0.6:0.2	24.440	0.692	0.224

Numbers in bold indicate the best performance for each metric.

We employ the Ranger optimizer (Lookahead + RAdam) with a weight decay coefficient 10^{-4}, initialized at a learning rate 3×10^{-4}. Cosine annealing adjusts the learning rate with an initial period T_0 = 10 epochs, period multiplier T_{mult} = 2, and range 3×10^{-4} to 10^{-6}. The base channel count is set to 32, with a batch size of 64 for training and 16 for validation. Training was conducted on NVIDIA RTX 3090 Ti GPUs.

3. Results and Discussion

The schematic diagram of the experimental platform for polarization imaging in a cement dust environment is shown in Figure 2. A transparent acrylic tank with dimensions of 60 cm × 40 cm × 40 cm is used as the experimental container, and the exterior of the tank is covered with black velvet cloth to eliminate interference from external stray light. The illumination source is a non-polarized white light LED (Ulanzi 40 W PRO, 3900 K, Ulanzi, Shenzhen, China) in this work. However, these results can be extended to polarized illumination sources because the underlying physics of dehazing restoration is based on the evolution of polarization states with Mueller matrix. The color polarization camera (LUCID PHX050S-QC, Lucid Vision Labs, Richmond, BC, Canada) with a resolution of 2448 × 2048 pixels can simultaneously capture images at the linear polarization angles of 0°, 45°, 90°, 135°. An air blower was arranged at the top corner of the acrylic tank near the target, and ordinary Portland cement (Grade 42.5) was placed below the blower. A continuous dust gradient from thin to dense was generated by the blower to simulate cement dust environments with different concentrations. The image dataset was collected experimentally by recording videos during the dynamic change process of dust concentration. The dataset constructed in this experiment contains 596 image groups, covering 29 types of target objects. All original images were cropped to a standard size of 224 × 224 pixels with a stride of 128 pixels. After data augmentation processing, they were randomly divided into a training set and a validation set at a ratio of 9:1, which were then used for subsequent model training and performance verification.

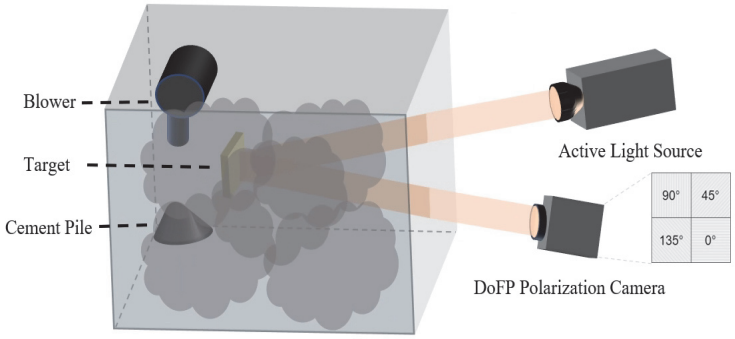


Figure 2. Experimental setup diagram.

Two parameters μ_c and ζ are introduced by considering the evolution of DoLP to quantitatively evaluate the overall haze intensity and the concentration inhomogeneity of dust.

To ensure consistent cross-image comparability, the final concentration μ_c and inhomogeneity ζ of the single-image are quantified as follows:

$$\mu_c = \frac{1}{HW} \sum_{x=1}^{H} \sum_{y=1}^{W} \frac{c_{gt}(x,y) - c_{haze}(x,y)}{c_{gt}(x,y)}, \tag{11}$$

$$\zeta = \frac{\sigma_c}{\mu_c}, \text{ where } \sigma_c = \sqrt{\frac{1}{HW} \sum_{x=1}^{H} \sum_{y=1}^{W} (c(x,y) - \mu_c)^2}. \tag{12}$$

where $c_{haze}(x,y)$ and $c_{gt}(x,y)$ denote the DoLP of evaluated image and its corresponding ground truth, respectively. H and W are the height (H) and width (W) of the evaluated image. Based on the global statistical analysis of the entire experimental dataset, the spatial dust distribution of inhomogeneous cement dust environments is classified with a high-concentration threshold of $\mu_c \geq 0.4$ and high-heterogeneity of $\zeta \geq 0.3$.

To validate the effectiveness of our proposed CDMMNet, we conducted comparisons against several representative methods: the physics-based DCP [7], CLAHE [8], the deep learning-based GCANet [25], and the polarization-based MTM-Net [17]. Three evaluation metrics were adopted to comprehensively assess restoration quality from different perspectives: peak signal-to-noise ratio (PSNR) is a pixel-level metric derived from MSE, quantifying pixel intensity discrepancies between restored and reference images; structural similarity index measure (SSIM) [26] is rooted in human visual perception of structure, evaluating similarity in terms of luminance, contrast and structural alignment, emphasizing key local structural details; and learned perceptual image patch similarity (LPIPS) [27] is a deep learning-based metric that computes the difference between two image patches in deep feature space by calculating the weighted distance between the corresponding features from multiple layers of a pre-trained convolutional neural network and the ground truth.

Comparative experiments were conducted between our CDMMNet and representative methods (DCP [7], CLAHE [8], GCANet [25], and MTM-Net [17]) under various cement dust concentrations and scattering inhomogeneity, as shown in Figures 3 and 4. Five different target objects under various scattering scenarios (low and uniform, high and uniform, low and non-uniform, high and non-uniform) were tested. The corresponding quantitative evaluation metrics are summarized in Tables 1 and 2, providing objective support for visual assessment. Scenarios with low concentration and uniform scattering environments represent the most ideal dehazing conditions, as shown in Figure 3a. All dehazing methods achieve favorable restoration performance. Although the DCP method suffers from a certain degree of darkening artifacts, it still provides acceptable results. The CLAHE method exhibits limited visual dehazing performance but can basically recover the information of the target object. In contrast, GCANet and MTM-Net demonstrate superior performance in terms of brightness and color fidelity, effectively removing haze while preserving the details of target objects. However, GCANet exhibits slight blurring artifacts in specific regions, particularly at edges with large concentration gradients. CDMMNet, while maintaining image sharpness, significantly reduces artifacts and delivers visual effects closely aligned with the ground truth.

Quantitative metrics also confirm its comprehensive advantages, as shown in Table 2. When the haze concentration is significantly increased while the distribution remains relatively uniform, the basic statistical assumptions of the DCP method are severely disrupted by cement dust, resulting in substantially darkened restored images, as shown in Figure 3b. The CLAHE algorithm amplifies image noise and exacerbates visual distortion. GCANet performs inadequately in this low-light environment, producing shadows and overall underexposure. Although MTM-Net outperforms the aforementioned two methods in brightness and color fidelity, it still exhibits a certain level of color distortion. On the other hand, CDMMNet retains excellent restoration performance even under high-concentration conditions, particularly excelling in detail preservation and artifact suppression with an advantage fully verified by quantitative evaluation, as shown in Table 2.

When the scattering environments become non-uniform spatial distributions, the DCP method demonstrates poor adaptability to complex lighting conditions, with evident darkening in the processed images, as shown in Figure 4a. CLAHE achieves moderate dehazing performance in low-concentration regions but introduces notable color distortion in the restored images. GCANet effectively mitigates thin haze but fails to address thick

haze scenarios, where it even exacerbates occlusion artifacts. MTM-Net outperforms other methods in overall brightness and color fidelity, yet its performance degrades under non-uniform conditions. However, CDMMNet achieves the optimal restoration performance in non-uniform scenarios due to its adaptive adjustment mechanism for handling spatially varying concentration distributions, as shown in Figure 4a and Table 3. Scenarios with both high concentration and high non-uniform distribution represent the most challenging conditions. In this extreme case, all methods face greater challenges. The DCP method barely produces usable restoration results, with images being severely darkened and lacking details. CLAHE method suffers severe detail loss and structural degradation under high concentration and non-uniformity conditions. GCANet also delivers unsatisfactory performance, showing not only severe shadows and underexposure but also a substantial number of blurring artifacts. Although MTM-Net improves the overall brightness of images to some extent, its effectiveness in non-uniform regions remains limited. In contrast, CDMMNet still significantly enhances image quality in this most challenging scenario by integrating advanced physical models with data-driven strategies. Based on comprehensive quantitative and visual evaluations, CDMMNet exhibits the most significant advantages under these extreme conditions, as demonstrated in Figure 4b. Experimental results indicate that GCANet and MTM methods achieve better texture restoration on target objects compared to the CLAHE method. The results of DCP are unsatisfactory and can be attributed to limitations of the dark channel prior under such dust conditions. Notably, CDMMNet consistently outperforms all comparative methods across the scattering scenarios by integrating concentration-adaptive Mueller matrix modeling, thereby achieving superior detail preservation and minimal artifacts validated through both quantitative metrics and visual assessments.

Table 2. Quantitative performance comparison of different networks (Uniform conditions).

Metrics		DCP	CLAHE	GCAN	MTM	Ours		DCP	CLAHE	GCAN	MTM	Ours
PSNR	$\mu_c = 0.19$ $\zeta = 0.08$	10.25	19.72	20.52	19.24	**24.23**	$\mu_c = 0.64$ $\zeta = 0.20$	9.74	15.48	16.02	19.88	**22.87**
SSIM		0.36	0.65	0.63	0.62	**0.65**		0.35	0.62	0.57	0.63	**0.65**
LPIPS		0.25	0.18	0.24	0.20	**0.12**		0.40	0.39	0.38	0.26	**0.16**
PSNR	$\mu_c = 0.14$ $\zeta = 0.13$	18.92	19.82	24.20	16.12	**29.94**	$\mu_c = 0.62$ $\zeta = 0.30$	26.14	16.46	16.33	11.23	**31.74**
SSIM		0.60	0.83	0.87	0.73	**0.93**		0.65	0.74	0.73	0.58	**0.93**
LPIPS		0.06	0.09	0.14	0.29	**0.02**		0.32	0.19	0.26	0.24	**0.04**
PSNR	$\mu_c = 0.10$ $\zeta = 0.09$	17.60	21.59	22.25	19.04	**28.33**	$\mu_c = 0.56$ $\zeta = 0.29$	17.18	18.86	17.28	15.24	**26.72**
SSIM		0.60	0.72	0.67	0.65	**0.80**		0.46	0.66	0.64	0.58	**0.76**
LPIPS		0.17	0.21	0.19	0.14	**0.06**		0.51	0.43	0.37	0.24	**0.11**
PSNR	$\mu_c = 0.19$ $\zeta = 0.10$	14.29	19.29	18.92	18.12	**25.02**	$\mu_c = 0.56$ $\zeta = 0.28$	13.50	16.48	15.14	13.00	**23.38**
SSIM		0.43	0.72	0.71	0.73	**0.87**		0.55	0.58	0.54	0.53	**0.78**
LPIPS		0.57	0.21	0.21	0.09	**0.04**		0.26	0.51	0.46	0.38	**0.13**
PSNR	$\mu_c = 0.13$ $\zeta = 0.11$	13.05	20.11	19.45	16.05	**23.18**	$\mu_c = 0.59$ $\zeta = 0.29$	13.48	18.13	18.65	16.81	**21.72**
SSIM		0.41	0.70	0.66	0.60	**0.81**		0.47	0.69	0.69	0.61	**0.69**
LPIPS		0.24	0.19	0.20	0.45	**0.08**		0.34	0.35	0.20	0.23	**0.11**

Numbers in bold indicate the best performance for each metric.

Besides restoration quality, the computational efficiency of CDMMNet is evaluated to verify its suitability for real-time industrial monitoring. The inference speed of CDMMNet with an NVfIDIA GeForce RTX 3090 Ti GPU achieves 139.3 FPS for an image with a resolution of 224 × 224 pixels, and the inference speed is 30.32 FPS for a larger monitoring

window of 512 × 512 pixels. Additionally, CDMMNet maintains a manageable model size with 8.39 M parameters. The computational cost, measured in FLOPs, is 54.24 G for 224 × 224 inputs and 283.37 G for 512 × 512 inputs. These metrics demonstrate that CDMMNet achieves a high level of efficiency suitable for practical industrial deployment. Furthermore, the training time per iteration for both the baseline U-Net and the proposed CDMMNet are examined by using a batch size of 32. Experiment results indicate that the baseline U-Net requires 0.1113 s per step whereas the full CDMMNet requires 0.1292 s. There is a 16.07% training time increase with CE/CEMC modules.

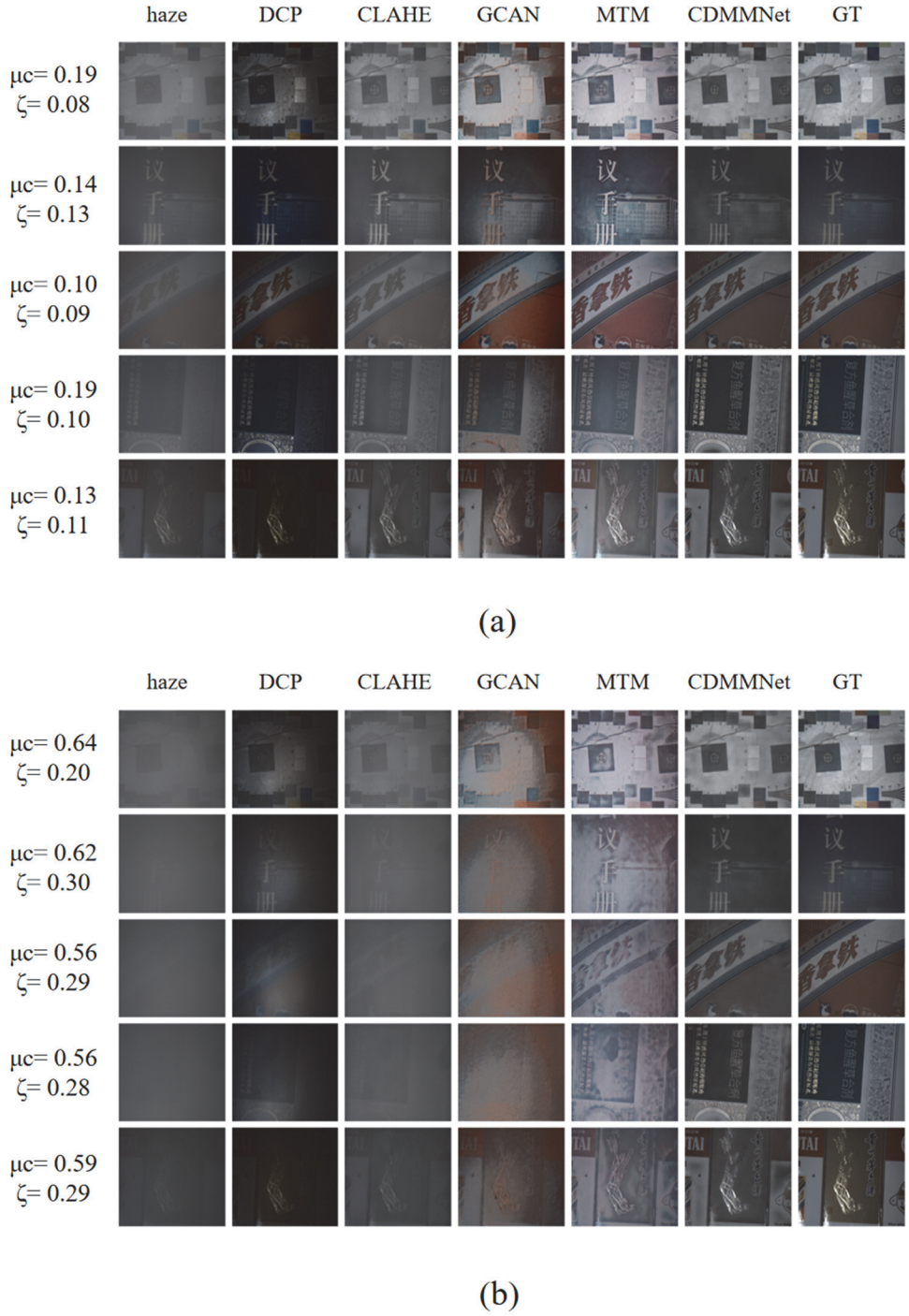


Figure 3. Visual comparison of results between our CDMMNet and other methods (DCP, CLAHE, GCANet, and MTM-Net) under: (**a**) the low and uniform scenario, (**b**) the high and uniform scenario. The quantitative evaluation results of haze concentration and heterogeneity for the above images are shown in the figure.

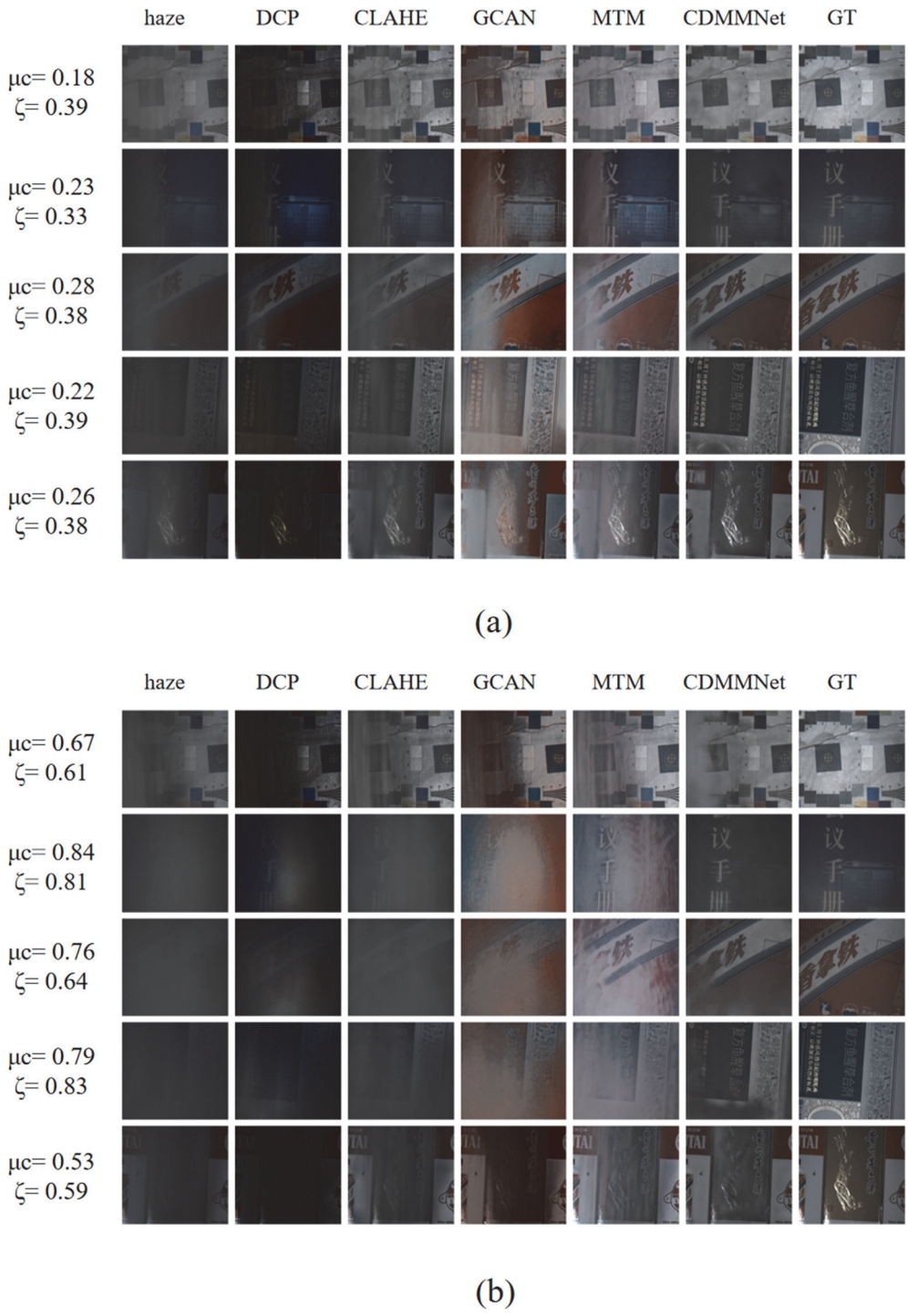


Figure 4. Visual comparison of results between our CDMMNet and other methods (DCP, CLAHE, GCANet, and MTM-Net) under: (**a**) the low and non-uniform scenario, (**b**) the high and non-uniform scenario.

A series of ablation studies were conducted to verify the effectiveness of the specific CE and CEMC modules of the proposed network. All experiments were performed with the same hyperparameter settings and trained on the same dataset. The visualization results and quantitative performance metrics are presented in Figure 5 and Table 4. The results for the network without the CE module show that it cannot effectively handle the degradation caused by abruptly appearing high-concentration dust as shown in Figure 5 (column 2). This confirms the key role of the DoLP-based concentration distribution estimation module for de-scattering imaging in a non-uniform cement dust environment.

In comparison, the network without the CEMC module produced a poor dehazing effect with severely degraded images, as shown in Figure 5 (column 3), because the network cannot accurately apply the physical model to the corresponding concentration regions. When the network lacks both CE and CEMC modules, it can only handle low-concentration dust on the left side of the image but is ineffective against high-concentration dust on the right, and it also introduces obvious color distortion due to non-uniform degradation, as shown in Figure 5 (column 4). The average PSNR, SSIM, and LPIPS metrics corresponding to each method in the ablation study are listed in Table 4, and these results collectively demonstrate the superior performance of the proposed method and its core modules in handling non-uniform cement dust.

Table 3. Quantitative performance comparison of different networks (non-uniform conditions).

Metrics		DCP	CLAHE	GCAN	MTM	Ours		DCP	CLAHE	GCAN	MTM	Ours
PSNR	$\mu_c = 0.18$ $\zeta = 0.39$	8.40	16.01	15.42	19.56	**19.99**	$\mu_c = 0.67$ $\zeta = 0.61$	7.96	12.85	11.70	15.55	**15.77**
SSIM		0.35	0.60	0.59	0.64	**0.65**		0.25	0.57	0.52	0.61	**0.63**
LPIPS		0.40	0.23	0.27	0.20	**0.19**		0.40	0.33	0.36	0.31	**0.256**
PSNR	$\mu_c = 0.23$ $\zeta = 0.33$	19.87	28.13	17.01	14.39	**32.34**	$\mu_c = 0.84$ $\zeta = 0.81$	20.02	24.22	14.01	14.20	**31.37**
SSIM		0.61	0.89	0.75	0.70	**0.93**		0.60	0.86	0.66	0.68	**0.93**
LPIPS		0.23	0.16	0.23	0.16	**0.04**		0.25	0.18	0.32	0.20	**0.06**
PSNR	$\mu_c = 0.28$ $\zeta = 0.38$	16.07	23.07	19.64	19.58	**26.68**	$\mu_c = 0.76$ $\zeta = 0.64$	15.63	21.00	19.40	12.58	**25.74**
SSIM		0.53	0.77	0.72	0.74	**0.83**		0.46	0.70	0.67	0.58	**0.78**
LPIPS		0.23	0.25	0.21	0.13	**0.07**		0.46	0.43	0.35	0.31	**0.18**
PSNR	$\mu_c = 0.22$ $\zeta = 0.39$	14.12	17.97	17.52	18.29	**23.07**	$\mu_c = 0.79$ $\zeta = 0.83$	13.01	15.62	16.18	14.88	**17.55**
SSIM		0.53	0.77	0.68	0.73	**0.85**		0.46	0.63	0.59	0.62	**0.77**
LPIPS		0.21	0.25	0.30	0.20	**0.06**		0.38	0.45	0.40	0.33	**0.23**
PSNR	$\mu_c = 0.26$ $\zeta = 0.38$	13.80	19.88	17.92	17.35	**22.85**	$\mu_c = 0.53$ $\zeta = 0.59$	11.78	18.58	15.73	15.19	**21.45**
SSIM		0.45	0.80	0.71	0.69	**0.81**		0.26	0.77	0.69	0.77	**0.84**
LPIPS		0.41	0.20	0.36	0.23	**0.09**		0.33	0.26	0.29	0.29	**0.09**

Numbers in bold indicate the best performance for each metric.

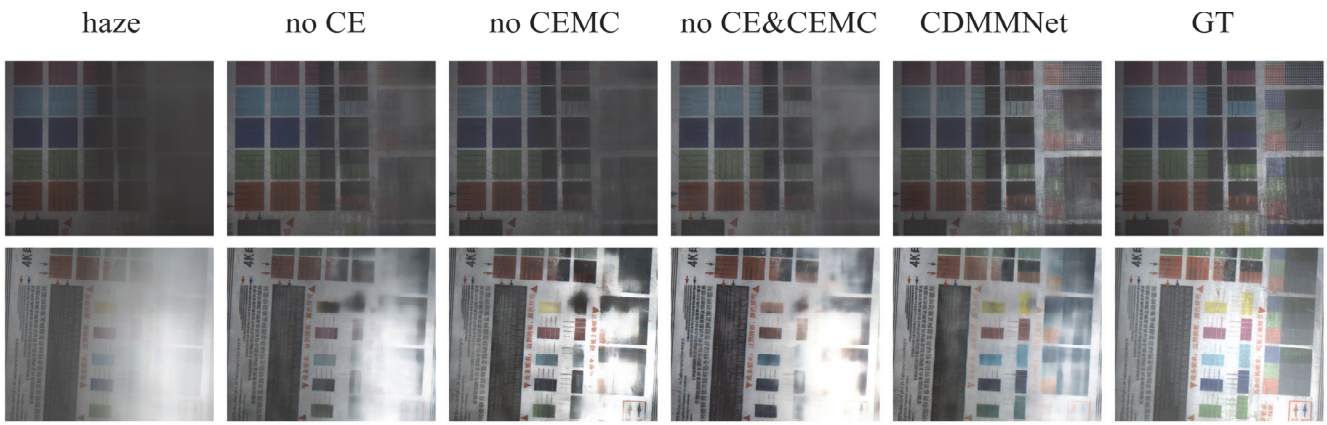


Figure 5. Ablation study on two challenging scenes with spatially varying cement dust (upper plots: $\mu_c = 0.54$ and $\zeta = 0.71$; lower plots: $\mu_c = 0.73$ and $\zeta = 0.91$). Each row is one scene. Columns from left to right show results for: hazy input, without CE module, without CEMC module, without both CE and CEMC, full CDMMNet (proposed), ground truth image. The full model effectively removes non-uniform haze while preserving texture and color.

Table 4. Performance metrics of ablation study.

CE	CEMC	PSNR	SSIM	LPIPS
		15.031	0.639	0.343
√		18.083	0.695	0.235
	√	18.598	0.670	0.243
√	√	**24.984**	**0.751**	**0.147**

Numbers in bold indicate the best performance for each metric. The symbol "√" denotes the case that the module was included in the ablation study.

Specifically, we conduct an ablation study comparing performance with and without the s_3 channel. The differences of restoring performances with and without the s_3 channel are negligible, as shown in Figure 6 and Table 5.

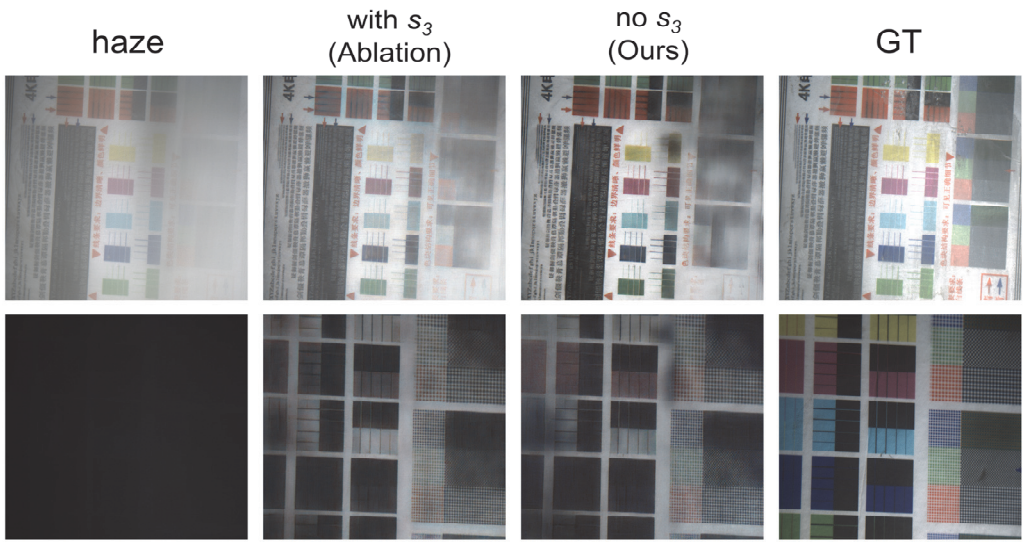


Figure 6. Visual validation of the s_3 component's impact.

Table 5. Quantitative comparison of the impact of the s_3 component.

MODULE	PSNR	SSIM	LPIPS
with s_3	20.339	**0.712**	0.191
no s_3	**21.816**	0.687	**0.179**

Numbers in bold indicate the best performance for each metric.

To further justify the specific internal designs of our CDMMNet, we conducted expanded component-level ablation studies focusing on the CE module's depth and the CEMC module's fusion strategy. The performance of the CE module with various convolutional layers is evaluated to demonstrate the impact of CE module layers. The quantitative results are summarized in Figure 7 and Table 6 for one, three, and five layers. According to the experiment results, the three-layer design (CE-3) is chosen in this work by considering the restoring performance and computational efficiency.

Ablation study on the CE module architecture with our proposed "Split + Conv" modulation and the standard feature concatenation are compared. Our proposed "Split + Conv" strategy outperforms simple concatenation by a significant improvement in evaluating indicators, as shown in Figure 8 and Table 7. Concatenation treats concentration as a generic data channel, relying on the network to implicitly learn the interaction. In contrast, the proposed physics-driven "Split + Conv" modulation strategy directly modulates the base

Mueller matrix **W** by splitting the concentration features into multiplicative scaling γ and additive bias β.

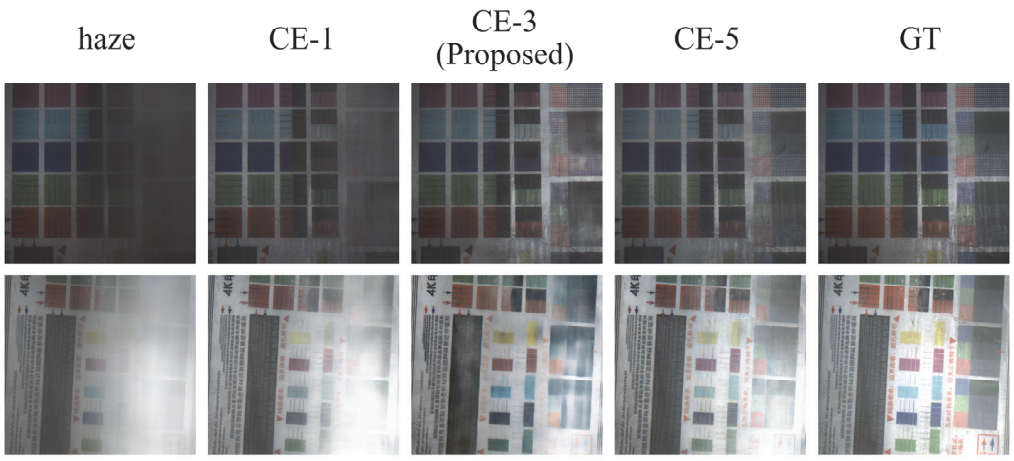


Figure 7. Ablation study on the CE module architecture: visual and quantitative comparison of different convolutional layer depths (one, three, and five layers).

Table 6. Quantitative performance comparison for different CE module layers.

MODULE	PSNR	SSIM	LPIPS
CE-1	21.229	0.740	0.223
CE-3 (Proposed)	24.984	0.751	0.147
CE-5	**25.905**	**0.760**	**0.127**

Numbers in bold indicate the best performance for each metric.

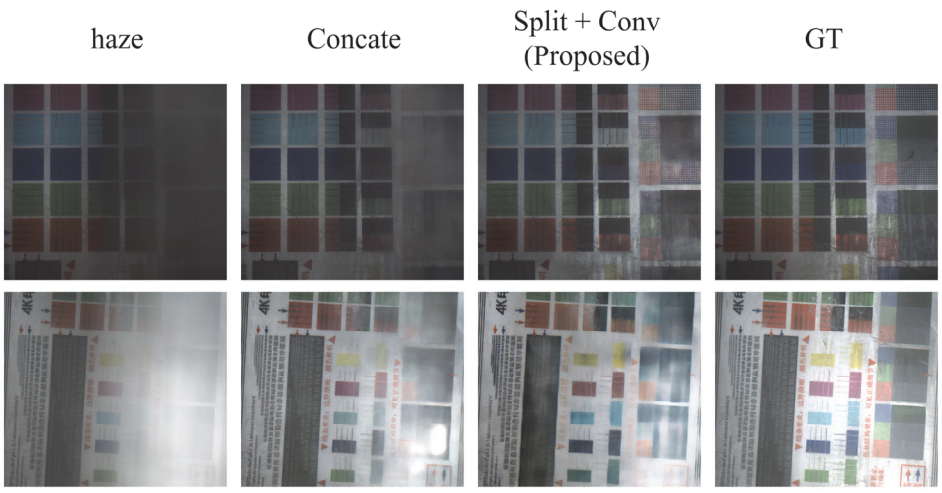


Figure 8. Ablation study on the CEMC fusion strategy: visual and quantitative comparison between conventional feature concatenation and the proposed physics-driven "Split + Conv" modulation.

Table 7. Quantitative performance comparison of CEMC fusion strategies.

MODULE	PSNR	SSIM	LPIPS
Concate	22.434	0.727	0.185
Split + Conv	**24.984**	**0.751**	**0.147**

Numbers in bold indicate the best performance for each metric.

4. Conclusions

Traditional de-scattering imaging methods are limited in non-uniform scattering scenarios. In this work, the CDMMNet de-scattering imaging scheme for non-uniform cement dust environments is introduced, based on Mueller matrix theory and deep learning techniques. This proposed network incorporates a novel CE module that determines local medium concentration distribution through the evolution of DoLP information, and a CEMC module that performs pixel-wise polarization information analysis of concentration data, enabling pixel-wise adaptation to dynamically varying dust concentrations. Experimental results demonstrate that our proposed CDMMNet significantly outperforms other methods in detail restoration, color fidelity, and perceptual similarity under diverse dynamic non-uniform cement dust environments. Ablation experiments confirm the effectiveness of the collaborative interaction between the CE and CEMC modules in scenarios with abrupt concentration gradients. This study establishes a unified "physical model + data-driven" framework for robust de-scattering imaging in complex, dynamic industrial environments, while the Mueller matrix model is theoretically feasible to any scattering environments and the performance of the physics-informed architecture of CDMMNet is primarily validated under laboratory environment. However, experiment results in real industrial environments indicate that its robustness is affected due to various types of cement dust and humidities. Therefore, more datasets are required to improve the generalization of this model according to various real industrial environments. In addition, its effectiveness is notably constrained in extreme scattering regimes where the scattering coefficient $\mu_c > 0.9$. In particular, the performance of CDMMNet may be degenerated into that of MTMNet when the DoLP–concentration relationship breaks down (e.g., specific lighting geometries or highly reflective targets).

Author Contributions: Conceptualization, R.-P.C.; methodology, P.Z. and C.G.; software, P.Z.; validation, P.Z. and C.G.; formal analysis, P.Z. and L.Z.; investigation, P.Z., C.G., W.Y. and L.Z.; resources, R.-P.C.; data curation, P.Z. and C.G.; writing—original draft preparation, P.Z.; writing—review and editing, R.-P.C., K.-H.C. and C.G.; visualization, P.Z. and W.Y.; supervision, R.-P.C.; project administration, R.-P.C.; funding acquisition, R.-P.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Natural Science Foundation of Zhejiang Province under Grant LD26E050004; Natural Science Foundation of China under Grant 12474300 and 62405281.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Tyo, J.S.; Goldstein, D.L.; Chenault, D.B.; Shaw, J.A. Review of passive imaging polarimetry for remote sensing applications. *Appl. Opt.* **2006**, *45*, 5453–5469. [CrossRef] [PubMed]
2. Wang, X.; Chen, Y.; Guan, C.; Zhu, L.; Chew, K.-H.; Chen, R.-P. Adaptive learning-driven polarimetric dehazing imaging in a dynamic turbid environment. *J. Opt. Soc. Am. A* **2025**, *42*, 672–678. [CrossRef] [PubMed]
3. Guan, C.; Li, J.; Tian, Y.; Wu, W.; Zhu, L.; Chew, K.-H.; Chen, R.-P. Polarization-enhanced super-resolution imaging reconstruction based on a denoising diffusion probabilistic model. *Opt. Express* **2025**, *33*, 39238–39248. [CrossRef] [PubMed]
4. Chen, Y.; Wang, X.; Guan, C.; Zhu, L.; Chew, K.-H.; Chen, R.-P. Polarization-resolved semantic segmentation for classification and detection of materials in a turbid underwater environment. *Opt. Express* **2025**, *33*, 25545–25557. [CrossRef] [PubMed]
5. Liu, J.; Zhang, Y.; Zhang, Q.; Wang, J. Scattering matrix for typical urban anthropogenic origin cement dust and discrimination of representative atmospheric particulates. *J. Geophys. Res.* **2018**, *123*, 3159–3174. [CrossRef]

6. Pizer, S.M.; Amburn, E.P.; Austin, J.D.; Cromartie, R.; Geselowitz, A.; Greer, T.; ter Haar Romeny, B.; Zimmerman, J.B.; Zuiderveld, K. Adaptive histogram equalization and its variations. *Comput. Vis. Graph. Image Process.* **1987**, *39*, 355–368. [CrossRef]
7. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353. [CrossRef] [PubMed]
8. Reza, A.M. Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *J. VLSI Signal Process. Syst. Signal Image Video Technol.* **2004**, *38*, 35–44. [CrossRef]
9. Getreuer, P. Automatic Color Enhancement (ACE) and its Fast Implementation. *Image Process. Line* **2012**, *2*, 266–277. [CrossRef]
10. Cai, B.; Xu, X.; Jia, K.; Qing, C.; Tao, D. DehazeNet: An end-to-end system for single image haze removal. *IEEE Trans. Image Process.* **2016**, *25*, 5187–5198. [CrossRef] [PubMed]
11. Zhang, H.; Patel, V.M. Densely connected pyramid dehazing network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Los Alamitos, CA, USA, 2018; pp. 3194–3203.
12. Wang, N.; Zhou, Y.; Han, F.; Zhu, H.; Yao, J. UWGAN: Underwater GAN for real-world underwater color restoration and dehazing. *IEEE Access* **2020**, *8*, 1–10.
13. Li, B.; Peng, X.; Wang, Z.; Xu, J.; Feng, D. AOD-Net: All-in-One dehazing network. In *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW)*; IEEE: Los Alamitos, CA, USA, 2017; pp. 4780–4788.
14. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [CrossRef]
15. Chen, Z.; Wu, J.; Hu, H.; Li, X. Underwater polarimetric descattering via scene adaptation and multi-parameter optimization. *Opt. Lasers Eng.* **2026**, *196*, 109410. [CrossRef]
16. Li, B.; Chen, Z.; Lu, L.; Qi, P.; Zhang, L.; Ma, Q.; Hu, H.; Zhai, J.; Li, X. Cascaded frameworks in underwater optical image restoration. *Inf. Fusion* **2025**, *117*, 102809. [CrossRef]
17. Gao, J.; Wang, G.; Chen, Y.; Wang, X.; Li, Y.; Chew, K.-H.; Chen, R.-P. Mueller transform matrix neural network for underwater polarimetric dehazing imaging. *Opt. Express* **2023**, *31*, 27213–27222. [CrossRef] [PubMed]
18. Li, X.; Han, Y.; Wang, H.; Liu, T.; Chen, S.-C.; Hu, H. Polarimetric imaging through scattering media: A review. *Front. Phys.* **2022**, *10*, 815296. [CrossRef]
19. Muinonen, K.; Tyynelä, J.; Zubko, E.; Lindqvist, H.; Penttilä, A.; Videen, G. Polarization of light backscattered by small particles. *J. Quant. Spectrosc. Radiat. Transf.* **2011**, *112*, 2109–2122. [CrossRef]
20. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *Proceedings of the International Conference on Learning Representations (ICLR)*; OpenReview.net: La Jolla, CA, USA, 2015.
21. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*; Springer: Cham, Switzerland, 2015; pp. 234–241.
22. Sun, D.; Gao, D. DPANet: Dual pooling attention network for semantic segmentation. *arXiv* **2022**, arXiv:2210.05437. [CrossRef]
23. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.P.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Los Alamitos, CA, USA, 2017; pp. 4681–4690.
24. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In *Computer Vision—ECCV 2016*; Springer: Cham, Switzerland, 2016; pp. 694–711.
25. Chen, D.; He, M.; Fan, Q.; Liao, J.; Zhang, L.; Hou, D.; Yuan, L.; Hua, G. Gated context aggregation network for image dehazing and deraining. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*; IEEE: Los Alamitos, CA, USA, 2019; pp. 1375–1383.
26. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef]
27. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; IEEE: Los Alamitos, CA, USA, 2018; pp. 586–595.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.