LEARNING POLARIZATION IMAGE RESTORATION WITH GENERAL RESTORATION PRIORS

Anonymous authors
Paper under double-blind review

ABSTRACT

Polarization imaging captures distinctive surface and geometric cues that benefit a wide range of vision tasks. However, real-world polarization acquisition is often affected by multiple coupled degradations, making image restoration essential for practical polarization vision. Existing methods are largely tailored to specific degradations and remain constrained by the limited scale and quality of polarization data. To address these limitations, we develop an all-in-one polarization restoration framework for diverse and composite degradations. We first study the impact of different polarization representations on restoration performance and identify the normalized Stokes representation as an effective choice for separating intensity and polarization information. Accordingly, we devise a dual-branch architecture that separates intensity and polarization modeling. To overcome the limitations of polarization-specific training, the intensity branch leverages pretrained general restoration priors and a mixture-of-experts extension for composite degradations, while its restoration knowledge is adaptively distilled into the symmetric polarization branch via a cross-domain feature transform. In addition, we establish a composite-degradation polarization benchmark to support all-in-one restoration research. Extensive experiments on public datasets and our proposed benchmark demonstrate the effectiveness of the proposed method.

1 INTRODUTION

Polarization imaging reveals unique physical properties of a scene, such as reflectance, surface roughness, and geometry. These polarimetric cues benefit various downstream tasks, including object detection (Luo et al. (2025)), semantic segmentation (Liu et al. (2025)), reflection removal (Yao et al. (2025); Lyu et al. (2022); Lei et al. (2020); Wang et al. (2026)), and surface normal estimation (Deschaintre et al. (2021); Lyu et al. (2023); Lei et al. (2022)). However, the key to fully leveraging these benefits lies in the accurate acquisition of polarization properties. Polarization imaging mainly follows two paradigms, division-of-time (DoT) and division-of-focal-plane (DoFP), both of which acquire multi-angle intensity measurements to derive the degree of linear polarization (DoLP) and angle of polarization (AoP) (Rebhan et al. (2019)). This imaging mechanism makes polarization images more vulnerable to real-world degradations, including blur, low illumination, and noise, which limits their practical applications. Therefore, polarization image restoration is crucial for reliable deployment of polarization imaging.

The performance of image restoration models is largely determined by the quality of the training data. Existing polarization image restoration methods typically train models using paired polarization image datasets, either synthesized (Zhou et al. (2025a)) or captured from real scenes (Zhou et al. (2023); Xu et al. (2022); Rahman et al. (2025); Li et al. (2026b)). Although these datasets enable promising restoration performance, they still lag far behind general image restoration datasets. Taking image deblurring as an example, conventional image deblurring methods usually construct blurred-sharp image pairs through frame averaging from high-frame-rate videos, such as the GoPro (Nah et al. (2017)) and REDS (Nah et al. (2019)) datasets. However, due to the limitations of polarization cameras, existing polarization deblurring methods mainly rely on synthetic blur kernels to generate blurred-sharp pairs (Zhou et al. (2025a)). Moreover, the complex acquisition process of polarization data further restricts the scale and diversity of polarization datasets. Meanwhile, models carefully designed and trained on general restoration datasets have demonstrated strong performance

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

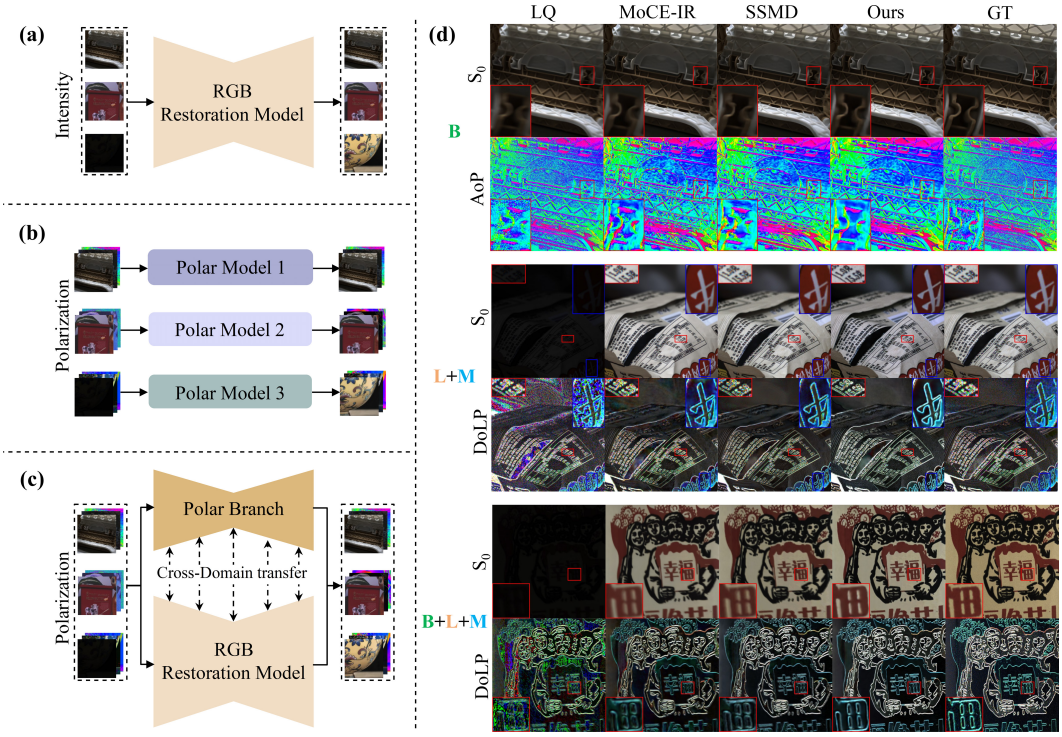


Figure 1: Polarization restoration approach. (a) Angle-wise general restoration overlooks polarization correlations. (b) Degradation-specific experts ("Polar Model") handle only individual degradations and remain constrained by polarization data distribution. (c) Our framework combines general restoration priors with polarization-specific modeling for restoration across diverse degradations. (d) Visual comparisons on blur (B), low light (L), mosaic (M), and their combinations: low quality image (LQ), MoCE-IR (Zamfir et al. (2025)), SSMD (Zhou et al. (2026)), Ours, and ground truth (GT) from left to right.

and generalization, naturally raising a question: *Can general restoration priors be effectively transferred to polarization image restoration to overcome exisiting domain-specific limitations?*

Beyond data, existing methods also face limitations in model design. Although current approaches optimize restoration networks towards the final polarization image quality, the representations learned by these models are largely chosen empirically, spanning the intensity domain (Li et al. (2025); Luo et al. (2024)), polarization domain (Zhou et al. (2023)), or combinations of both (Xu et al. (2022); Zhou et al. (2026)). The impact of different polarization representations on restoration performance has received limited attention. This leads to another fundamental question: *Which representation is better suited for polarization image restoration?* Furthermore, existing studies mainly focus on developing task-specific models with dedicated parameters for individual polarization degradations. In contrast, an all-in-one (AIO) polarization image restoration framework capable of handling diverse degradations remains largely unexplored.

Based on these observations, we propose to leverage general restoration priors for addressing polarization degradation. Specifically, we first conduct a systematic investigation into the impact of different polarization representations on restoration performance, and reveal that normalized Stokes-based representations are more suitable for diverse restoration tasks due to its intensity-polarization decoupling property. Based on this finding, we develop a dual-branch restoration framework that separately models intensity and polarization information. For the intensity branch, we adapt a pre-trained general restoration model (Conde et al. (2024)) and extend it into a mixture-of-experts (MoE) architecture (Puigcerver et al. (2024)) to better handle complex degradations, while introducing a symmetric polarization branch to recover polarization-specific information. To bridge general restoration priors and polarization-specific representations, we introduce a bidirectional interaction mechanism. A adaptive cross-domain feature transform selectively distills transferable intensity priors to guide polarization restoration, while complementary polarization cues are fed back via zero-

convolution modulation to preserve the pretrained intensity restoration capability. The proposed dual-branch architecture preserves domain-specific modeling to retain general restoration priors, while enabling effective information exchange between intensity and polarization representations. Finally, we introduce PolarCDD, a new polarization restoration dataset covering 11 degradation settings composed of blur, low light, noise, mosaic artifacts, and their combinations, to facilitate research on AIO and composite polarization restoration.

In summary, our contributions can be summarized as follows: (1) We investigate how different representation choices affect polarization restoration and find that the normalized Stokes representation provides the best trade-off for polarization fidelity. (2) We develop an AIO polarization restoration model that equips the intensity branch with pretrained restoration priors and an MoE extension, while leveraging an adaptive modulation mechanism to transfer its restoration capability to the polarization domain. (3) We establish PolarCDD, a dataset with diverse composite polarization degradations, to facilitate future research in polarization image restoration.

2 Related Work

Polarization Image Restoration. Most polarization image restoration methods are task-specific for individual degradations, and typically exploit either intensity information, polarization cues, or their complementary relationship. For example, intensity-guided restoration has been explored for polarization deblurring and enhancement, while dual-domain or parallel designs further model intensity and polarization jointly to improve fidelity (Zhou et al. (2025a); Xu et al. (2022); Zhou et al. (2023)). For denoising, prior work has explored residual dense feature aggregation for polarization parameter recovery and hybrid-attention Transformers with physics-based noise modeling for low-light polarization images (Li et al. (2020; 2023)). For demosaicking, prior work has investigated staged reconstruction, adversarial learning, diffusion models, and joint optimization with downstream tasks (Nguyen et al. (2022); Guo et al. (2024b); Li et al. (2025); Luo et al. (2026b)). More recently, Zhou et al. investigate general network architectures for polarization image restoration and propose a single-stage dual-domain design applicable to different restoration tasks (Zhou et al. (2026)).

Despite this progress, existing methods remain largely constrained by limited polarization data and task-specific architectural design. And the choice of polarization representation—an essential factor that determines how intensity and polarization information are encoded and restored—has received little systematic investigation. In addition, AIO restoration for diverse polarization degradations remains largely unexplored.

General Image Restoration. General image restoration has progressed from task-specific backbones toward unified models capable of handling diverse degradations. Representative architectures such as Restormer and NAFNet establish strong generic restoration backbones through efficient attention and simplified network design (Zamir et al. (2022); Chen et al. (2022)). Building upon these backbones, recent AIO methods introduce degradation-aware conditioning or adaptive routing to accommodate multiple restoration tasks. PromptIR and InstructIR employ learned prompts or textual instructions to condition restoration behavior (Potlapalli et al. (2023); Conde et al. (2024)), while OneRestore and AdaIR model degradation characteristics through multimodal or frequency-aware representations (Guo et al. (2024a); Cui et al. (2025)). MoCE-IR further introduces mixture-of-experts routing to dynamically allocate restoration capacity across degradations (Zamfir et al. (2025)). Beyond discriminative models, recent work has explored large-scale generative and foundation priors for image restoration, demonstrating the benefit of knowledge learned from broader restoration data (Fei et al. (2026); Chen et al. (2026)). However, these advances are primarily developed for conventional RGB images. This motivates us to investigate whether general restoration priors can be effectively transferred to polarization restoration, where data are substantially more limited and both intensity fidelity and polarization accuracy must be preserved.

3 METHODOLOGY

3.1 Preliminaries

Polarization imaging captures intensity images at multiple polarization angles and subsequently derives the polarization properties of the scene. Specifically, the polarization image x_θ at a given

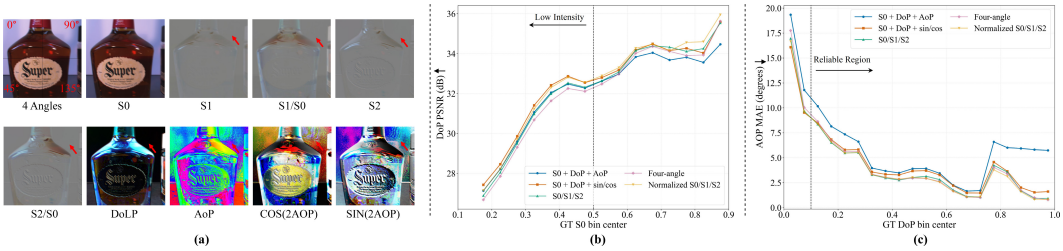


Figure 2: Representation study. (a) Representation visualization. (b) DoLP metric across intensity levels. Higher is better. (c) AoP metric across DoLP levels. Lower is better.

angle θ can be represented as:

$$x_\theta = I\{1 + p\cos 2(\theta - \phi)\}, \tag{1}$$

where I is the unpolarized light intensity, p and ϕ denote the degree and angle of polarization (DoLP and AoP), respectively. p and ϕ are governed by the material properties, surface geometry, and viewing direction , and thus encode rich polarization cues of the scene. We represent the captured intensity measurements at all polarization angles as $X = [x_{0^\circ}, x_{45^\circ}, x_{90^\circ}, x_{135^\circ}]$. X can be further transformed into a more suitable representation for polarization imaging, namely the Stokes vector (Collett (2005)):

$$\begin{aligned} S_0 &= \frac{1}{2}(x_{0^\circ} + x_{45^\circ} + x_{90^\circ} + x_{135^\circ}), \\ S_1 &= x_{0^\circ} - x_{90^\circ}, \quad S_2 = x_{45^\circ} - x_{135^\circ}, \end{aligned} \tag{2}$$

where $\{S_0, S_1, S_2\}$ denotes the Stokes vector. Accordingly, I, p, and ϕ in Eq. (1) can be derived from the Stokes vector:

$$I = \frac{S_0}{2}, p = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}, \phi = \frac{1}{2}\arctan(\frac{S_2}{S_1}). \tag{3}$$

When polarization imaging is affected by degradations, the degraded observation $Y = [y_{0^\circ}, y_{45^\circ}, y_{90^\circ}, y_{135^\circ}]$ can be generally formulated as:

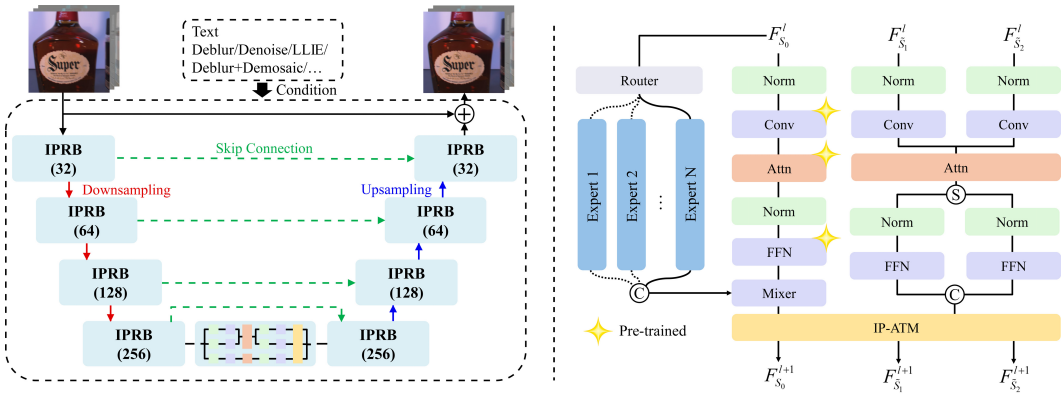
$$Y = \mathcal{D}(X) + n, \tag{4}$$

where n denotes the noise, and $\mathcal{D}$ can represent arbitrary degradation operators, such as exposure adjustment, blur kernels (Zhou et al. (2025a)), and imaging subsampling (Li et al. (2026a)). According to Eqs. (2)-(4), compared with intensity images, p and ϕ suffer from amplified degradation effects due to their nonlinear computation, resulting in lower image quality. Therefore, Polarization image restoration requires not only recovering intensity information but also ensuring accurate estimation of polarization parameters, which poses significant challenges for model optimization.

3.2 Representation Exploration

Before designing the network, we first address a fundamental question: Which polarization representation (or parameterization) is better suited for polarization image restoration? We compare five formulations: the intensity domain $[x_{0^\circ}, x_{45^\circ}, x_{90^\circ}, x_{135^\circ}]$, Stokes $[S_0, S_1, S_2]$, normalized Stokes $[S_0, \tilde{S}_1, \tilde{S}_2]$ with $\tilde{S}_i = S_i/S_0$, and the polarization-parameter domains $[S_0, p, \phi]$ and $[S_0, p, \cos(2\phi), \sin(2\phi)]$. Using NAFNet (Chen et al. (2022)) as a common backbone, we evaluate all representations on three tasks (deblurring (Zhou et al. (2025a)), LLIE (Zhou et al. (2023)), and demosaicking (Zhou et al. (2025b))) under identical settings.

As shown in Tab. 1, directly restoring angular intensities is redundant, whereas $[S_0, p, \phi]$ is difficult to optimize because AoP is periodic. Stokes-based representations perform better, with $[S_0, \tilde{S}_1, \tilde{S}_2]$ giving the best overall polarization accuracy by decoupling polarization from intensity, consistent with Eq. (1). Such decoupling makes the representation more robust to low intensity than the original Stokes representation, as illustrated in Fig. 2(b).



(a) Overall Pipeline (b) IPRB

Figure 3: Overview of the proposed method.(a) The normalized Stokes representation is processed by a U-Net composed of intensity-polarization restoration blocks (IPRBs), with restoration behavior conditioned by textual guidance through instruction condition block (ICB (Conde et al. (2024))) . (b) Each IPRB contains intensity and polarization branches built upon NAFNet-style attention and feed-forward blocks (Chen et al. (2022)). Residual connections are employed but omitted in the figure for clarity. The intensity branch incorporates pretrained restoration priors (Conde et al. (2024)) and an MoE extension, while the polarization branch separately models two polarization components and aggregates them during attention. Finally, the intensity-polarization adaptive transfer module (IP-ATM) enables domain interaction and adapts restoration priors to the polarization domain.

Moreover, $[\tilde{S}_1, \tilde{S}_2] = [p\cos(2\phi), p\sin(2\phi)]$ provides a continuous Cartesian encoding of AoP. Since AoP is unreliable in low-DoLP regions (Luo et al. (2026a)), even in the ground truth. Overfitting these regions during optimization can degrade restoration performance. p acts as a confidence weight and reduces the influence of noise. This explains its lower reliable-region error (MAE_r) than $[S_0, p, \cos(2\phi), \sin(2\phi)]$ in Tab. 1 and Fig. 2(c). We therefore adopt $[S_0, \tilde{S}_1, \tilde{S}_2]$ throughout.

3.3 Composite Degradation Polarization Image Dataset

Existing polarization datasets mainly target single degradations, while real-world polarization imaging often suffers from coupled degradations caused by low signal-to-noise ratio, long exposure, and DoFP sampling. To support composite restoration, we construct PolarCDD, a large-scale polarization dataset containing four basic degradations—low light, motion blur, noise, and mosaic artifacts—and their combinations, for a total of 11 degradation settings. Blur and mosaic degradations are synthesized following (Zhou et al. (2025a)) and (Luo et al. (2026b)), respectively; noise includes both synthetic and real observations, while low-light data are captured using a DoT system. Each degradation setting contains 50,000 training images and 100 test images[1].

Table 1: Average performance of different representations on three restoration tasks. From (a) to (e): $[x_{0^\circ}, x_{45^\circ}, x_{90^\circ}, x_{135^\circ}]$, $[S_0, S_1, S_2]$, $[S_0, \tilde{S}_1, \tilde{S}_2]$, $[S_0, p, \phi]$ and $[S_0, p, \cos(2\phi), \sin(2\phi)]$. The subscripts I and p denote intensity and DoLP, respectively. MAE_r is the AOP error in reliable regions with GT DoLP > 0.1.

Type	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$	$MAE_r\downarrow$
(a)	30.489	29.390	0.864	0.713	16.541	12.973
(b)	**30.603**	29.642	**0.864**	0.723	15.740	12.618
(c)	30.515	**30.110**	0.862	**0.727**	15.603	**12.101**
(d)	30.118	29.424	0.857	0.712	16.982	13.852
(e)	30.268	29.790	0.858	0.722	**14.987**	12.503

3.4 Network Design

3.4.1 Overview

Our network follows a U-Net architecture as shown in Fig. 3, with each stage composed of several stacked Intensity-Polarization Restoration Blocks (IPRBs). The normalized Stokes representation $[S_0, \tilde{S}_1, \tilde{S}_2]$ are concatenated and fed into the network. Specifically, each IPRB adopts a parallel

[1]Detailed pipelines are provided in the appendix.

tri-stream design, where three polarization representations are independently processed by their respective backbones and interact across domains through the Intensity-Polarization Adaptive Transfer Module (IP-ATM). To overcome the limited distribution of existing polarization data, we introduce general restoration priors. Since S_0 corresponds to a standard RGB image, pretrained RGB restoration priors can be directly applied to its restoration. Therefore, the IPRB backbone follows the RGB restoration model (in this work, InstructIR (Conde et al. (2024)) [2]) design, with the intensity branch initialized from its pretrained weights. For efficiency, we reduce the block number of the polarization branch by half, such that only half of the IPRBs operate solely on S_0 in each stage. Following InstructIR, textual embeddings are used to condition the restoration behavior of the network.

3.4.2 IPRB

Within each IPRB, different representations are processed by separate branches to reduce interference caused by distribution discrepancies across domains. Each branch follows the standard NAFNet (Chen et al. (2022)) design. For the intensity branch, we further extend the backbone into an MoE architecture to improve its capability in handling composite degradations. Each expert is initialized with pretrained InstructIR (Conde et al. (2024)) weights and mixed with the main branch through a convolutional layer. Intensity branch employs four experts, with only the top-1 expert selected for mixture. For the polarization branch, Stokes representations are also processed independently, but are concatenated before attention to enable joint aggregation within the domain.

3.4.3 IP-ATM

After backbone feature extraction, the intensity and polarization features are further complemented through a bidirectional intensity-polarization interaction mechanism. Our key motivation is to transfer the powerful restoration priors learned in the intensity domain to the polarization domain. Inspired by SFT (Wang et al. (2018)) and pixel AdaLN (Yu et al. (2026)), we use the intensity features to adaptively modulate the polarization representations through pixel-wise scaling and shifting at the end of each layer l. However, not all information encoded in the intensity features is beneficial to polarization restoration, and indiscriminate transfer may introduce intensity-specific biases into the polarization domain. We therefore first normalize the representation $[F^l_{S_0}, F^l_{\tilde{S}_1}, F^l_{\tilde{S}_2}]$ to suppress their domain-specific statistics. $[F^l_{S_0}, F^l_{\tilde{S}_1}, F^l_{\tilde{S}_2}]$ are then concatenated to predict adaptive gating weights that selectively identify transferable features:

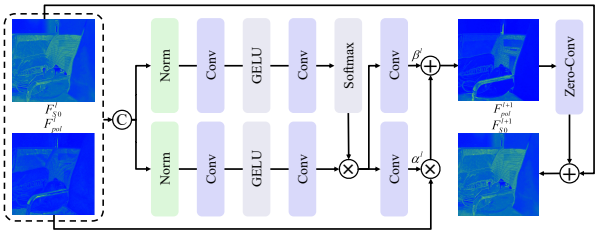


Figure 4: IP-ATM mechanism. Polarization representations selectively incorporate intensity restoration priors through adaptive scaling and shifting, followed by zero-convolution modulation that feeds polarization cues back into the intensity features.

$$G^l = Softmax(Convs(F^l)), M^l = G^l \odot Convs(F^l), \tag{5}$$

where F^l represents the concatenated features, $Convs$ denotes convolutional operations, and $\odot$ is the element-wise multiplication. $Softmax$ is applied along the channel dimension. The selected features M^l are further used to estimate the affine parameters $[\alpha^l, \beta^l]$ for modulating the polarization representations:

$$\alpha^l, \beta^l = Convs(M^l), F^{l+1}_{pol} = (1 + \alpha^l) \odot F^l_{pol} + \beta^l, \tag{6}$$

where F^l_{pol} and F^{l+1}_{pol} denote the polarization features $[F^l_{\tilde{S}_1}, F^l_{\tilde{S}_2}]$ before and after modulation.

In the reverse direction, the intensity branch already has strong restoration capability inherited from restoration priors. We therefore adopt a simple interaction strategy, where polarization information is injected into the intensity features through zero-convolution-based modulation (Zhang et al. (2023)):

$$F^{l+1}_{S_0} = F^l_{S_0} + Conv_{zero}(F^{l+1}_{Pol}). \tag{7}$$

[2] Our design is architecture-agnostic. We further investigate different restoration priors in the appendix.

Table 2: Quantitative comparison under the AIO3 setting. The superscript $*$ indicates task-specific weights. **Bold** indicates the best performance.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
PolDeblur*	29.193	27.232	0.864	0.694	15.803
SSMD*	32.601	29.550	0.909	0.744	14.416
PromptIR	25.098	27.554	0.802	0.690	15.858
InstructIR	29.128	28.388	0.873	0.709	15.596
AdaIR	25.106	27.516	0.801	0.690	15.874
MoCE-IR	31.451	29.012	0.902	0.729	15.184
CDIR	25.366	27.633	0.807	0.692	15.842
SSMD	27.880	27.588	0.842	0.699	15.495
Ours	**34.029**	**29.939**	**0.933**	**0.749**	**13.991**

(a) Deblurring dataset (Zhou et al. (2025a)).

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
PLIE*	38.967	26.876	0.974	0.689	16.855
CPNet*	39.868	28.568	0.980	0.754	13.954
PromptIR	38.211	27.856	0.974	0.727	14.352
InstructIR	39.171	28.202	0.976	0.737	14.405
AdaIR	37.893	27.541	0.972	0.720	14.560
MoCE-IR	39.373	28.758	0.979	0.745	14.270
CDIR	39.094	28.296	0.977	0.735	13.998
SSMD	39.549	27.975	0.977	0.744	14.059
Ours	**40.512**	**29.352**	**0.982**	**0.755**	**13.340**

(b) LLIE dataset (Zhou et al. (2023)).

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
PIDSR*	44.406	40.888	0.980	0.921	11.939
PUGDiff*	46.210	41.893	0.986	0.926	10.704
PromptIR	44.181	41.114	0.981	0.927	12.360
InstructIR	42.922	39.451	0.974	0.900	14.522
AdaIR	44.265	41.136	0.982	0.927	12.331
MoCE-IR	43.448	39.737	0.976	0.903	14.303
CDIR	44.188	40.964	0.981	0.924	12.027
SSMD	45.043	42.185	0.984	0.941	12.505
Ours	**47.088**	**42.505**	**0.990**	**0.942**	**9.966**

(c) Demosaicking dataset (Zhou et al. (2025b)).

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
NAFNet	30.567	30.697	0.859	0.747	15.128
Restormer	30.242	30.513	0.852	0.745	15.124
PromptIR	30.451	30.728	0.856	0.748	14.939
InstructIR	33.037	30.948	0.903	0.756	15.263
AdaIR	30.454	30.685	0.855	0.748	14.959
MoCE-IR	34.794	31.489	0.925	0.771	14.915
CDIR	30.701	30.780	0.860	0.749	14.826
SSMD	32.685	31.009	0.884	0.759	14.700
Ours	**37.518**	**32.820**	**0.949**	**0.794**	**13.009**

(d) Average performance.

Table 3: Quantitative comparison on the proposed PolarCDD and average performance. **Bold** indicates the best performance.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
PromptIR	34.248	30.627	0.919	0.757	16.914
InstructIR	33.237	31.309	0.923	0.779	17.146
AdaIR	34.729	31.185	0.925	0.774	15.687
MoCE-IR	34.542	32.250	0.938	0.800	16.615
CDIR	34.112	30.945	0.924	0.771	16.173
SSMD	34.455	32.031	0.930	0.804	15.525
Ours	**37.376**	**33.191**	**0.953**	**0.826**	**13.393**

(a) One degradation.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
PromptIR	27.559	28.029	0.849	0.680	20.843
InstructIR	27.480	27.888	0.855	0.689	21.438
AdaIR	27.843	28.575	0.858	0.703	19.094
MoCE-IR	28.616	28.777	0.875	0.715	20.737
CDIR	27.486	28.372	0.854	0.700	19.957
SSMD	28.531	29.049	0.876	0.730	18.939
Ours	**31.759**	**30.905**	**0.922**	**0.778**	**16.043**

(b) Dual degradations.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
PromptIR	24.989	26.688	0.795	0.647	20.838
InstructIR	25.279	26.749	0.807	0.655	21.247
AdaIR	24.865	26.940	0.798	0.659	19.899
MoCE-IR	26.411	27.260	0.828	0.673	20.761
CDIR	24.993	27.091	0.802	0.665	20.084
SSMD	26.053	27.057	0.831	0.675	19.656
Ours	**29.904**	**29.446**	**0.901**	**0.744**	**16.397**

(c) Triple degradations.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
PromptIR	29.524	28.730	0.865	0.702	19.413
InstructIR	29.173	28.925	0.871	0.716	19.843
AdaIR	29.805	29.227	0.871	0.721	18.002
MoCE-IR	30.370	29.764	0.889	0.738	19.243
CDIR	29.442	29.074	0.870	0.719	18.604
SSMD	30.235	29.771	0.888	0.747	17.828
Ours	**33.464**	**31.471**	**0.929**	**0.789**	**15.144**

(d) Average performance.

The zero initialization enables the interaction to start from an identity-like state, preserving the pretrained restoration behavior at the beginning of optimization while progressively incorporating complementary polarization cues to improve intensity restoration.

4 Experiment

4.1 Implementation Details

We train two separate model weights. The first is trained on a mixture of single-degradation datasets (AIO3 setting), including PolDeblur (Zhou et al. (2025a)) for deblurring, PLIE (Zhou et al. (2023)) and LLCP (Xu et al. (2022)) for low-light image enhancement, and PIDSR (Zhou et al. (2025b)) for polarization demosaicking. These datasets are jointly used to train a single AIO restoration model. The second model is trained on our proposed PolarCDD dataset. For both settings, The training and test sets are drawn from the same dataset. All models are trained for 200K iterations using the Adam (Kingma & Ba (2014)) optimizer with an initial learning rate of 2×10^{-4} and a cosine annealing schedule (Loshchilov & Hutter (2016)). The training patch size is set to 256×256 with a batch size

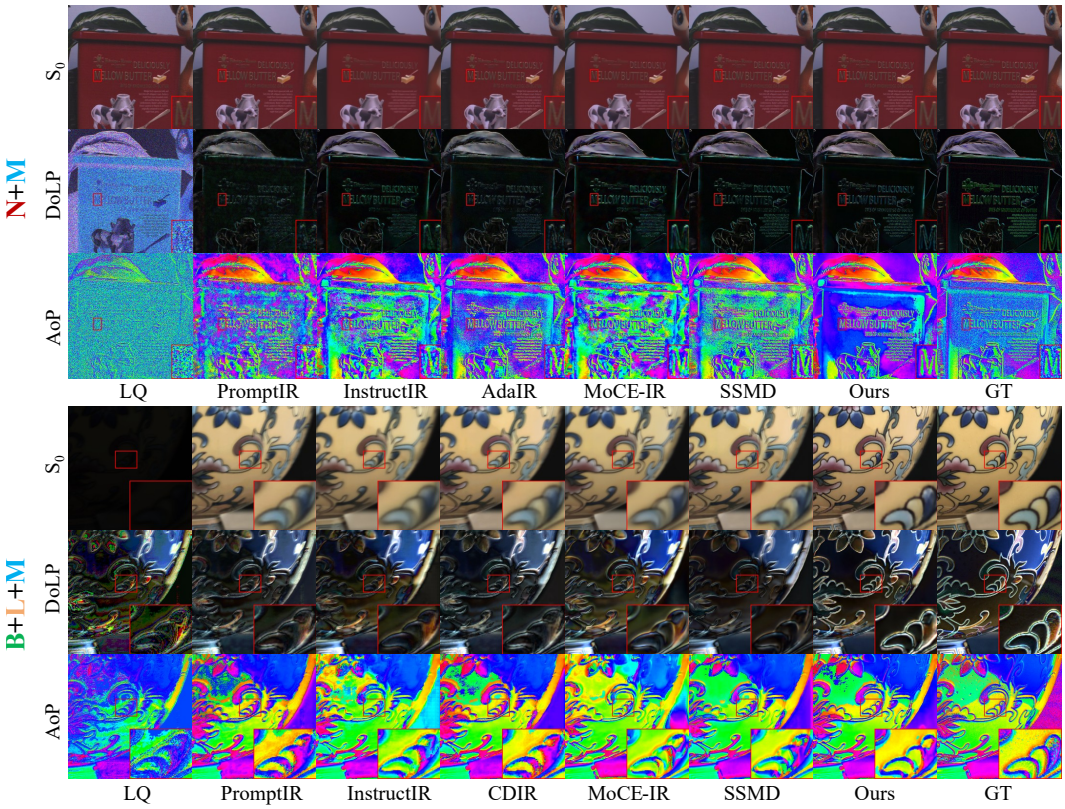


Figure 5: Visual comparisons of different methods on PolarCDD, including PromptIR (Potlapalli et al. (2023)), InstructIR (Conde et al. (2024)), AdaIR (Cui et al. (2025)), CDIR (Cui et al. (2026)), MoCE-IR (Zamfir et al. (2025)), SSMD (Zhou et al. (2026)) and Ours. N is noise, B is blur, L is low light, and M is mosaic.

of 8. All experiments are conducted on NVIDIA RTX 4090 GPUs. For the loss function, we employ an $\mathcal{L}_1$ loss to constrain the intensity images, the Stokes representations and DoP, while adopting a cosine loss for AoP to account for its periodicity.

4.2 Comparison with State-of-the-Art Methods

We compare our method against two restoration paradigms. The first category consists of RGB image restoration models, including NAFNet (Chen et al. (2022)), Restormer (Zamir et al. (2022)), PromptIR (Potlapalli et al. (2023)), InstructIR (Conde et al. (2024)), AdaIR (Cui et al. (2025)), MoCE-IR (Zamfir et al. (2025)), and CDIR (Cui et al. (2026)). To adapt these models to polarization restoration, we consider two strategies: (1) extending the input and output channels to match the polarization representation, i.e., the 9-channel normalized Stokes representation used in this work, (2) or treating each polarization angle as an individual RGB image and fine-tuning the pretrained models in batches. We train both variants and select the better one for each model.

Table 4: Ablation study of the proposed framework under the AIO3 setting.

Configuration	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
No promotion	36.649	31.125	0.945	0.757	14.316
+ Polarization-to-intensity	36.942	31.041	0.945	0.755	14.386
+ Intensity-to-polarization (ours)	**37.518**	**32.820**	**0.949**	**0.794**	**13.009**
IP-ATM → Zero-Conv	36.365	31.852	0.940	0.772	13.620
W/o gate	37.449	32.621	0.947	0.791	13.100
W/o MoE	35.610	32.319	0.938	0.786	13.362
W/o pretrained priors	33.797	31.564	0.901	0.764	13.909

The second category consists of polarization restoration methods, including task-specific models such as PolDeblur (Zhou et al. (2025a)) for deblurring, PLIE (Zhou et al. (2023)) and CPNet (Xu et al. (2022)) for LLIE, and PIDSR (Zhou et al. (2025b)) and PUGDiff (Li et al. (2026a)) for polarization demosaicking. In addition, SSMD (Zhou et al. (2026)) is a general polarization restoration framework. We train SSMD on the same mixed datasets to obtain an AIO version.

For evaluation, we use PSNR and SSIM to assess the quality of S_0 and DoLP, i.e. $PSNR_I$, $PSNR_p$, $SSIM_I$ and $SSIM_p$. For AoP, we use the mean angular error (MAE).

Qualitative Comparisons. Figs. 1 and 5 present qualitative comparisons under various degradations. Polarization properties are more vulnerable to degradation than intensity information, while our method produces cleaner results with sharper structural details and fewer noise artifacts.

Quantitative Comparisons. Tabs. 2 and 3 report quantitative results on public datasets and PolarCDD, respectively. Under the AIO setting, our method substantially outperforms general AIO models and even surpasses several task-specific methods in fidelity. We attribute this advantage to the effective exploitation of general restoration priors, which helps alleviate the limitations of existing polarization training data. In contrast, directly fine-tuning pretrained general restoration models such as InstructIR and MoCE-IR yields unsatisfactory performance, as these models operate in the intensity domain and cannot adequately capture inter-angle polarization discrepancies or higher-level polarization representations. These results suggest that simply transferring a pretrained RGB restoration model is insufficient; instead, a divide-and-conquer design that couples a general intensity restoration model with a dedicated polarization branch is necessary to effectively adapt general restoration priors to polarization restoration.

4.3 Ablation Study

Ablation on IP-ATM. We introduce a bidirectional promotion mechanism to facilitate information exchange between the intensity and polarization domains. As shown in Tab. 4, adding polarization-to-intensity modulation improves the intensity metrics, confirming that polarization cues provide complementary guidance for intensity restoration. Conversely, introducing intensity-to-polarization transfer brings substantial gains in polarization metrics, demonstrating the effectiveness of general restoration priors for polarization recovery. Replacing the intensity-to-polarization modulation with zero-convolution leads to consistent performance drops, suggesting that this direction requires stronger feature modulation. Removing the gating mechanism also degrades polarization-related metrics, indicating that selective modulation is preferable to unrestricted feature transfer. Fig. 6 visualizes how intensity features influence the polarization branch under different degradations. After interaction, the polarization features exhibit clearer structural details and reduced feature noise, while intensity-specific responses, such as the highlighted text in the first row, are not propagated into the polarization representation.

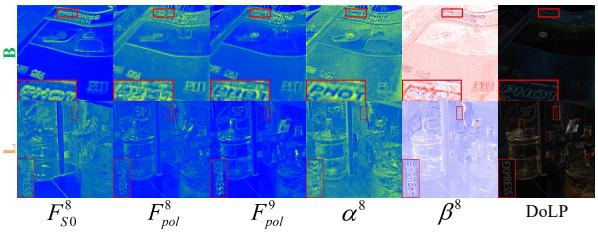


Figure 6: Feature visualization of IP-ATM at the 8th layer under different degradations.

Importance of General Restoration Priors. When the general restoration prior is removed and the intensity branch is trained from scratch, performance drops substantially, highlighting the importance of pretrained restoration knowledge. Nevertheless, the polarization metrics remain better than those obtained without bidirectional interaction, further validating the effectiveness of IP-ATM even in the absence of pretrained priors. In addition, the MoE extension consistently improves performance across all metrics and enhances robustness to degradations outside the pretraining distribution, such as demosaicking [3].

5 Conclusion

This work presents a unified polarization restoration network for handling diverse degradations. We first reveal the advantages of normalized Stokes representations for polarization restoration, particularly under low-intensity conditions and in reliable polarization regions. Based on this observation, we develop an intensity-polarization decoupled dual-branch architecture and introduce general restoration priors into the intensity domain to overcome the limitations of existing polarization data distributions. To effectively exploit these priors for polarization restoration, we further design an adaptive transfer mechanism that identifies transferable intensity features and selectively incorporates them into the polarization branch. In addition, we establish a composite-degradation polarization dataset to support research on polarization restoration. Extensive experiments on public datasets and our proposed benchmark demonstrate the effectiveness of the proposed approach.

[3] Additional ablation studies are provided in the appendix.

AI USE STATEMENT

In this work, we used generative AI tools solely for language polishing. No generative AI tools were used for the mathematical derivations or algorithm design. We carefully reviewed all AI-assisted content and take full responsibility for the final content of this work.

REPRODUCIBILITY STATEMENT

We emphasize reproducibility throughout this work. Detailed descriptions of the architecture, training setups, and evaluation protocols are provided in the main text (see corresponding sections) and in the Appendix. If the paper is accepted, we will release the source code and pretrained weights to facilitate transparent verification and future research.

REFERENCES

Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European Conference on Computer Vision (ECCV)*, pp. 17–33. Springer, 2022.

Xiang Chen, Jinshan Pan, Jiangxin Dong, Jian Yang, and Jinhui Tang. Foundir-v2: Optimizing pre-training data mixtures for image restoration foundation model. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8471–8480, 2026.

Edward Collett. *Field guide to polarization*, volume 15. SPIE press Bellingham, Washington, 2005.

Marcos V Conde, Gregor Geigle, and Radu Timofte. Instructir: High-quality image restoration following human instructions. In *European Conference on Computer Vision (ECCV)*, pp. 1–21. Springer, 2024.

Yuning Cui, Syed Waqas Zamir, Salman Khan, Alois Knoll, Mubarak Shah, and Fahad Khan. Adair: Adaptive all-in-one image restoration via frequency mining and modulation. In *International Conference on Learning Representations (ICLR)*, volume 2025, pp. 101306–101327, 2025.

Yuning Cui, Wenqi Ren, Boxin Shi, Jianhou Gan, and Alois Knoll. Cdir: Lora-inspired attention for efficient composite degradation image restoration. *IEEE Transactions on Image Processing*, 2026.

Valentin Deschaintre, Yiming Lin, and Abhijeet Ghosh. Deep polarization imaging for 3d shape and svbrdf acquisition. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15562–15571. IEEE, 2021.

Song Fei, Tian Ye, Lujia Wang, and Lei Zhu. Lucidflux: Caption-free universal image restoration via a large-scale diffusion transformer. In *International Conference on Learning Representations (ICLR)*, volume 2026, pp. 86233–86260, 2026.

Yu Guo, Yuan Gao, Yuxu Lu, Huilin Zhu, Ryan Wen Liu, and Shengfeng He. Onerestore: A universal restoration framework for composite degradation. In *European Conference on Computer Vision (ECCV)*, pp. 255–272. Springer, 2024a.

Yuxuan Guo, Xiaobing Dai, Shaoju Wang, Guang Jin, and Xuemin Zhang. Attention-based progressive discrimination generative adversarial networks for polarimetric image demosaicing. *IEEE Transactions on Computational Imaging*, 10:713–725, 2024b.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Chenyang Lei, Xuhua Huang, Mengdi Zhang, Qiong Yan, Wenxiu Sun, and Qifeng Chen. Polarized reflection removal with perfect alignment in the wild. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1747–1755. IEEE, 2020.

Chenyang Lei, Chenyang Qi, Jiaxin Xie, Na Fan, Vladlen Koltun, and Qifeng Chen. Shape from polarization for complex scenes in the wild. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12622–12631. IEEE, 2022.

Chenggong Li, Yidong Luo, Caiyun Wu, Junchao Zhang, Degui Yang, and Dangjun Zhao. Demosaicking customized diffusion model for snapshot polarization imaging. *Optics & Laser Technology*, 188:112868, 2025.

Chenggong Li, Yidong Luo, Junchao Zhang, and Yang Degui. Polarization uncertainty-guided diffusion model for color polarization image demosaicking. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(8):6028–6036, Mar. 2026a.

Chenggong Li, Yidong Luo, Junchao Zhang, Boxin Shi, and Degui Yang. Polarvsr: A unified framework and benchmark for continuous space-time polarization video reconstruction. *arXiv preprint arXiv:2605.10275*, 2026b.

Xiaobo Li, Haiyu Li, Yang Lin, Jianhua Guo, Jingyu Yang, Huanjing Yue, Kun Li, Chuan Li, Zhenzhou Cheng, Haofeng Hu, et al. Learning-based denoising for polarimetric images. *Optics Express*, 28(11):16309–16321, 2020.

Zhuoxiao Li, Haiyang Jiang, Mingdeng Cao, and Yinqiang Zheng. Polarized color image denoising. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9873–9882. IEEE, 2023.

Zhuoyan Liu, Bo Wang, Lizhi Wang, Chenyu Mao, and Ye Li. Sharecmp: Polarization-aware rgb-p semantic segmentation. *IEEE Transactions on Circuits and Systems for Video Technology*, 35 (10):10316–10329, 2025.

Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. *arXiv preprint arXiv:1608.03983*, 2016.

Yidong Luo, Junchao Zhang, Jianbo Shao, Jiandong Tian, and Jiayi Ma. Learning a non-locally regularized convolutional sparse representation for joint chromatic and polarimetric demosaicking. *IEEE Transactions on Image Processing*, 33:5029–5044, 2024.

Yidong Luo, Junchao Zhang, and Chenggong Li. Cpifuse: Toward realistic color and enhanced textures in color polarization image fusion. *Information Fusion*, 120:103111, 2025.

Yidong Luo, Chenggong Li, Yuchao Feng, Boxin Shi, Junchao Zhang, and Xin Yuan. Stokes-informed diffusion for robust linear polarization estimation. *arXiv preprint arXiv:2607.21239*, 2026a.

Yidong Luo, Chenggong Li, Yunfeng Song, Ping Wang, Boxin Shi, Junchao Zhang, and Xin Yuan. Polarapp: Beyond polarization demosaicking for polarimetric applications. *arXiv preprint arXiv:2603.23071*, 2026b.

Youwei Lyu, Zhaopeng Cui, Si Li, Marc Pollefeys, and Boxin Shi. Physics-guided reflection separation from a pair of unpolarized and polarized images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(2):2151–2165, 2022.

Youwei Lyu, Lingran Zhao, Si Li, and Boxin Shi. Shape from polarization with distant lighting estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11):13991–14004, 2023.

Miki Morimatsu, Yusuke Monno, Masayuki Tanaka, and Masatoshi Okutomi. Monochrome and color polarization demosaicking using edge-aware residual interpolation. In *2020 IEEE International Conference on Image Processing (ICIP)*, pp. 2571–2575. IEEE, 2020.

Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3883–3891, 2017.

Seungjun Nah, Sungyong Baik, Seokil Hong, Gyeongsik Moon, Sanghyun Son, Radu Timofte, and Kyoung Mu Lee. Ntire 2019 challenge on video deblurring and super-resolution: Dataset and study. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1996–2005. IEEE, 2019.

Vy Nguyen, Masayuki Tanaka, Yusuke Monno, and Masatoshi Okutomi. Two-step color-polarization demosaicking network. In *2022 IEEE International Conference on Image Processing (ICIP)*, pp. 1011–1015. IEEE, 2022.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman H Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one image restoration. *Advances in Neural Information Processing Systems (NeurIPS)*, 36:71275–71293, 2023.

Joan Puigcerver, Carlos Riquelme Ruiz, Basil Mustafa, and Neil Houlsby. From sparse to soft mixtures of experts. In *International Conference on Learning Representations (ICLR)*, volume 2024, pp. 28435–28445, 2024.

Simeng Qiu, Qiang Fu, Congli Wang, and Wolfgang Heidrich. Linear polarization demosaicking for monochrome and colour polarization focal plane arrays. In *Computer Graphics Forum*, volume 40, pp. 77–89. Wiley Online Library, 2021.

Muhamad Daniel Ariff Bin Abdul Rahman, Yusuke Monno, Masayuki Tanaka, and Masatoshi Okutomi. Polarization denoising and demosaicking: dataset and baseline method. In *2025 IEEE International Conference on Image Processing (ICIP)*, pp. 2724–2729. IEEE, 2025.

David Rebhan, Maik Rosenberger, and Gunther Notni. Principle investigations on polarization image sensors. In *Photonics and Education in Measurement Science 2019*, volume 11144, pp. 50–54. SPIE, 2019.

Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*, pp. 3982–3992, 2019.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10674–10685. IEEE, 2022.

Dongyue Wang, Yang Lu, and Jiandong Tian. Polarization state tracing for reflection removal and color-consistent reconstruction. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5680–5689, 2026.

Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 606–615. IEEE, 2018.

Xiuyu Xu, Minjie Wan, Junyu Ge, Hao Chen, Xingcheng Zhu, Xiaojie Zhang, Qian Chen, and Guohua Gu. Colorpolarnet: residual dense network-based chromatic intensity-polarization imaging in low-light environment. *IEEE Transactions on Instrumentation and Measurement*, 71:1–10, 2022.

Mingde Yao, Menglu Wang, King-Man Tam, Lingen Li, Tianfan Xue, and Jinwei Gu. Polarfree: Polarization-based reflection-free imaging. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10890–10899. IEEE, 2025.

Yongsheng Yu, Wei Xiong, Weili Nie, Yichen Sheng, Shiqiu Liu, and Jiebo Luo. Pixeldit: Pixel diffusion transformers for image generation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 14273–14282, 2026.

Eduard Zamfir, Zongwei Wu, Nancy Mehta, Yuedong Tan, Danda Pani Paudel, Yulun Zhang, and Radu Timofte. Complexity experts are task-discriminative learners for any image restoration. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12753–12763. IEEE, 2025.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5718–5729. IEEE, 2022.

Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3813–3824. IEEE, 2023.

Chu Zhou, Minggui Teng, Youwei Lyu, Si Li, Chao Xu, and Boxin Shi. Polarization-aware low-light image enhancement. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 3742–3750, 2023.

Chu Zhou, Minggui Teng, Xinyu Zhou, Chao Xu, Imari Sato, and Boxin Shi. Learning to deblur polarized images. *International Journal of Computer Vision*, 133(9):5976–5991, 2025a.

Chu Zhou, Yufei Han, Junda Liao, Linrui Dai, Wangze Xu, Art Subpa-Asa, Heng Guo, Boxin Shi, and Imari Sato. Architectural unification for polarimetric imaging across multiple degradations. *arXiv preprint arXiv:2603.05834*, 2026.

Shuangfan Zhou, Chu Zhou, Youwei Lyu, Heng Guo, Zhanyu Ma, Boxin Shi, and Imari Sato. Pidsr: Complementary polarized image demosaicing and super-resolution. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16081–16090. IEEE, 2025b.

A Appendix

In the appendix, we first provide detailed network architecture in A.1, followed by the dataset construction procedure in A.2 and the implementation details of the compared methods in A.3. A.4 analyzes model complexity, A.5 investigates the impact of different restoration priors, and A.6 examines the design of the polarization branch. A.7 introduces the loss formulation and presents corresponding ablation studies. A.8 analyzes the mixture-of-experts mechanism, A.9 discusses the limitations of this work and potential directions for future research, and A.10 presents additional qualitative results.

A.1 Network Architecture Details

Detailed network configurations are provided in Tab. 5. The backbone of intensity-polarization restoration block (IPRB) follows the NAFNet design (Chen et al. (2022)). Specifically, the convolutional layers consist of 1×1 convolutions and 3×3 depth-wise convolutions, the attention module adopts Simplified Channel Attention (SCA), and the activation function is replaced with SimpleGate. Residual connections are also employed but omitted in Fig. 3 for clarity.

For the mixture-of-experts (MoE) module, it is only employed in the decoder. Each expert shares the same architecture as the intensity branch. Following MoCE-IR (Zamfir et al. (2025)), the router uses a linear layer to map the input intensity features and text embeddings into expert selection probabilities. The Mixer is implemented as a convolutional layer to fuse the outputs of the selected experts with the intensity-branch features.

For the polarization branch, the two polarization representations are temporarily aggregated during the attention operation and then split back into separate features for independent processing. They are finally merged again to facilitate information transfer.

Following InstructIR (Conde et al. (2024)), we use textual guidance to control the restoration behavior. The input text is encoded into an embedding by a text encoder (Reimers & Gurevych (2019)) and injected into each IPRB through instruction condition block (ICB). Notably, ICB modulates only the intensity features, while the corresponding restoration mode is propagated to the polarization domain through intensity-polarization adaptive transfer module (IP-ATM). For each degradation setting, we use a fixed text prompt. For example, the prompt is "deblur" for motion deblurring and "denoise+demosaic" for the joint noise-and-mosaic degradation. This fixed-prompt design allows the text embeddings to be precomputed offline, enabling the text encoder to be removed during inference for improved efficiency.

A.2 PolarCDD Dataset Construction

The synthesis pipeline of PolarCDD is illustrated in Fig. 7. To isolate the effect of each degradation, particularly mosaic artifacts, our data are primarily collected using DoT imaging systems

Table 5: Configurations of the proposed network architecture. B_0 and B_{12} denote the numbers of NAFBlocks in the intensity (S_0) and polarization ($\tilde{S}_1/\tilde{S}_2$) branches, respectively. c_0 and c_{12} denote their feature-channel dimensions. ICB is the instruction condition block, IP-ATM is the intensity–polarization adaptive transfer module, and MoE uses E experts with top-k routing.

Component	Layer	Configuration	Output size
Text encoder	Language model	BGE-Micro-v2; mean pooling; frozen	384
	Projection head	Linear $384 \to 256$; ℓ_2 normalization	256
Input	Representation	$[S_0, \tilde{S}_1, \tilde{S}_2]$; $3+3+3$ channels	$h \times w \times 9$
	Branch split	S_0: 3 channels; $[\tilde{S}_1, \tilde{S}_2]$: 6 channels	$h \times w \times (3+6)$
	Intro	3×3 Conv for S_0; grouped 3×3 Conv ($g=2$) for $\tilde{S}_1/\tilde{S}_2$	$h \times w \times (32+32)$
Encoder 1	IPRB	$(B_0, B_{12}) = (2,1)$; $(c_0, c_{12}) = (32,32)$; ICB; IP-ATM	$h \times w \times (32+32)$
	Down 1	2×2 Conv, stride 2	$\frac{h}{2} \times \frac{w}{2} \times (64+64)$
Encoder 2	IPRB	$(B_0, B_{12}) = (2,1)$; $(c_0, c_{12}) = (64,64)$; ICB; IP-ATM	$\frac{h}{2} \times \frac{w}{2} \times (64+64)$
	Down 2	2×2 Conv, stride 2	$\frac{h}{4} \times \frac{w}{4} \times (128+128)$
Encoder 3	IPRB	$(B_0, B_{12}) = (4,2)$; $(c_0, c_{12}) = (128,128)$; ICB; IP-ATM	$\frac{h}{4} \times \frac{w}{4} \times (128+128)$
	Down 3	2×2 Conv, stride 2	$\frac{h}{8} \times \frac{w}{8} \times (256+256)$
Encoder 4	IPRB	$(B_0, B_{12}) = (8,4)$; $(c_0, c_{12}) = (256,256)$; ICB; IP-ATM	$\frac{h}{8} \times \frac{w}{8} \times (256+256)$
	Down 4	2×2 Conv, stride 2	$\frac{h}{16} \times \frac{w}{16} \times (512+512)$
Bottleneck	IPRB	$(B_0, B_{12}) = (4,2)$; $(c_0, c_{12}) = (512,512)$; IP-ATM	$\frac{h}{16} \times \frac{w}{16} \times (512+512)$
Decoder 1	Up 1	1×1 Conv; PixelShuffle(2); skip addition	$\frac{h}{8} \times \frac{w}{8} \times (256+256)$
	IPRB–MoE	$(B_0, B_{12}) = (2,1)$; $E=4$, top-$k=1$; ICB; IP-ATM	$\frac{h}{8} \times \frac{w}{8} \times (256+256)$
Decoder 2	Up 2	1×1 Conv; PixelShuffle(2); skip addition	$\frac{h}{4} \times \frac{w}{4} \times (128+128)$
	IPRB–MoE	$(B_0, B_{12}) = (2,1)$; $E=4$, top-$k=1$; ICB; IP-ATM	$\frac{h}{4} \times \frac{w}{4} \times (128+128)$
Decoder 3	Up 3	1×1 Conv; PixelShuffle(2); skip addition	$\frac{h}{2} \times \frac{w}{2} \times (64+64)$
	IPRB–MoE	$(B_0, B_{12}) = (2,1)$; $E=4$, top-$k=1$; ICB; IP-ATM	$\frac{h}{2} \times \frac{w}{2} \times (64+64)$
Decoder 4	Up 4	1×1 Conv; PixelShuffle(2); skip addition	$h \times w \times (32+32)$
	IPRB–MoE	$(B_0, B_{12}) = (2,1)$; $E=4$, top-$k=1$; ICB; IP-ATM	$h \times w \times (32+32)$
Output	Ending	Separate 3×3 Convs; grouped Conv ($g=2$) for $\tilde{S}_1/\tilde{S}_2$; residual addition	$h \times w \times 9$

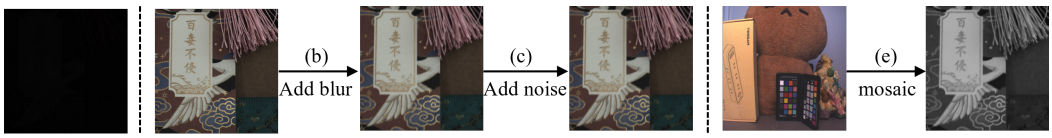


Figure 7: Degradation synthesis pipeline. low light: (a), blur: (b), noise: (c) or (d), mosaic: (e), low light+blur: (a)+(b), low light+mosaic: (a)+(e), blur+noise: (b)+(c), blur+mosaic: (b)+(e), noise+mosaic: (c)+(e) or (d)+(e), low light+blur+mosaic: (a)+(b)+(e), blur+noise+mosaic: (b)+(c)+(e).

(Morimatsu et al. (2020), Qiu et al. (2021), Li et al. (2025), Zhou et al. (2025b), and Rahman et al. (2025)). Most images in these datasets have resolutions above 1024×1024, allowing us to augment them through cropping. We use 256×256 crops for training and 512×512 crops for testing, while ensuring that the training and test sets contain no images from the same scene. Below, we detail the construction of each degradation type.

Motion blur. Due to the limited frame rate of polarization cameras, collecting paired sharp and blurred polarization images is challenging. We therefore follow the data synthesis strategy of PolDeblur Zhou et al. (2025a) to generate training pairs. The procedure is summarized as follows: (1) We first generate random camera motion trajectories. (2) The spatially carying blur model is applied to convert the motion trajectories into pixel-wise motion paths. (3) Each sharp polarization angles is warped along the corresponding pixel-wise trajectories to generate a sequence of latent sharp frames. (4) These latent frames are averaged to synthesized blurry images.

Low light. We capture multiple scenes at different exposure levels using a DoT system. Based on the exposure time, the images are categorized into four levels: extremely low light, low light, normal exposure, and overexposure.

Noise. Noise data are obtained in two ways. One is synthetically generated using Gaussian noise (Zhou et al. (2025a)), while the other is collected from a real-noise dataset (Rahman et al. (2025)). The real noise is further divided into three levels.

Mosaic artifacts. We synthesize mosaic observations from DoT images using the forward imaging operator (Luo et al. (2026b)) to construct training pairs.

Composite Degradations. We first select normally exposed or low-light scenes according to the exposure level, followed by blur synthesis, noise addition, and finally mosaic generation. When real-noise images are used, we do not apply additional blur or noise, as these operations may alter their inherent noise statistics. Similarly, no extra noise is added to low-light scenes, since low-light acquisition already introduces substantial noise. Finally, PolarCDD contains 11 degradation settings: low light, blur, noise, mosaic, low light+blur, low light+mosaic, blur+noise, blur+mosaic, noise+mosaic, low light+blur+mosaic, and blur+noise+mosaic.

A.3 Compared Methods Implementation Details

In this section, we detail the specific configurations of each comparison method. We reproduce each method following its official implementation or the configuration described in the corresponding paper. Given their different computational complexities, we aim to keep the training time and GPU memory consumption comparable across methods.

NAFNet. An activation-free NAFNet using 9-channel normalized Stokes input. It has width=32, four encoder levels with [1,1,1,28] blocks, one middle block, and [1,1,1,1] decoder blocks (Chen et al. (2022)).

Restormer. A four-level Restormer using MDTA and GDFN blocks with dim=32, block numbers [4,6,6,8], heads [1,2,4,8], and four refinement blocks. It takes 9-channel normalized Stokes input. (Zamir et al. (2022))

PromptIR. A Restormer-based model with learnable degradation prompts inserted at three decoder scales. It uses dim=48, blocks [4,6,6,8], heads [1,2,4,8], four refinement blocks, and 9-channel input/output (Potlapalli et al. (2023)).

InstructIR. A NAFNet-based instruction-conditioned model with ICB modules after each encoder and decoder level. It uses width=32, encoder blocks [2,2,4,8], four middle blocks, decoder blocks [2,2,2,2], and a 256-dimensional text embedding. Four RGB polarization angles are processed independently (Conde et al. (2024)).

AdaIR. A Restormer-based architecture with adaptive frequency modules in the bottleneck and decoder. It uses dim=48, blocks [4,6,6,8], heads [1,2,4,8], four refinement blocks, and 9-channel input/output (Cui et al. (2025)).

MoCE-IR. A Restormer-style encoder with shared decoder branches and frequency-domain mixture-of-experts adapters. It uses dim=48, encoder blocks [4,6,6,8], decoder blocks [2,4,4], four experts, and top-k=1 routing. Four RGB angles are processed independently (Zamfir et al. (2025)).

CDIR. A dynamic low-rank attention network using rank-8 DLA and multi-scale feed-forward blocks. It has width=32, encoder depths [2,3,6], six latent blocks, and decoder depths [3,6]. The input and output contain nine normalized Stokes channels (Cui et al. (2026)).

SSMD. A Transformer-based model that jointly processes four-angle images and Stokes features. It uses dim=32, six CDCI stages with [4,6,6,6,4,4] units and heads [1,2,4,2,1,1]. The input and output contain 12 four-angle channels following their original paper (Zhou et al. (2026)).

A.4 Complexity Analysis

Tab. 6 compares the computational complexity of different methods. Owing to its dual-branch architecture and MoE design, our model has the largest parameter count. Nevertheless, its convolution-dominated backbone maintains competitive memory consumption, computational cost, and inference speed. In contrast, many restoration models are Transformer-based. Despite having relatively few parameters, such as Restormer (Zamir et al. (2022)) and SSMD (Zhou et al. (2026)), their at-

Table 6: Complexity comparison. "9ch" denotes evaluation with a nine-channel normalized Stokes input adaptation. GPU memory, FLOPs, and latency are measured on 512×512 images using an NVIDIA RTX 4090 GPU.

Method	Params (M)	GPU memory (GB)	FLOPs (T)	Latency (ms)
NAFNet	17.112	1.068	0.294	82.91
NAFNet-9ch	17.115	0.321	0.074	29.14
Restormer	11.740	5.998	1.029	684.08
Restormer-9ch	11.745	1.533	0.259	175.47
PromptIR	35.377	9.179	2.530	915.41
PromptIR-9ch	35.385	2.474	0.635	228.93
InstructIR	33.334	1.353	0.259	129.70
InstructIR-9ch	33.337	0.430	0.066	29.59
AdaIR	28.785	9.232	2.355	977.52
AdaIR-9ch	28.865	2.495	0.591	255.68
MoCE-IR	25.354	4.101	1.591	917.63
MoCE-IR-9ch	25.359	1.226	0.355	197.97
CDIR	3.401	1.859	0.280	199.87
CDIR-9ch	3.409	0.484	0.072	49.02
SSMD	4.794	0.916	0.240	156.47
Ours	43.427	0.903	0.185	86.05

tention operations introduce substantial computational and memory overhead, making them less amenable to efficient scaling.

A.5 Analysis of Restoration Priors

Fig. 8 demonstrates the advantage of pretrained restoration priors over polarization-specific models. For motion deblurring, realistic blur is difficult to synthesize from polarization data to the same extent as in general image restoration. As a result, the pretrained restoration model (NAFNet (Chen et al. (2022))) produce better visual results on real-world polarization data than the expert model trained solely on polarization datasets. The polarization-specific expert fail to produce visually satisfactory results or restore blurred pixels to their correct spatial locations. Our method further transfers these strong restoration priors into the polarization domain, yielding substantial improvements under real-world degradations.

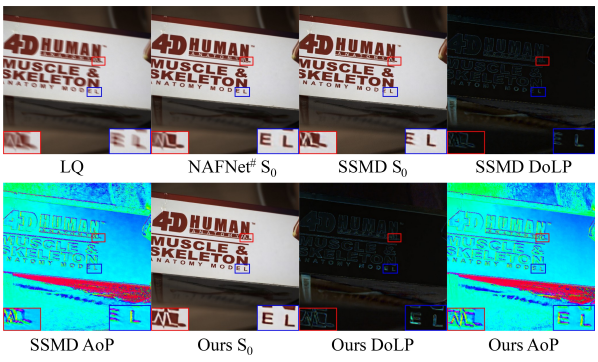


Figure 8: Visual comparison on real-world motion blur. The superscript # denotes the pretrained model.

Our method combines a general restoration model with a dedicated polarization branch to transfer restoration priors into the polarization domain. This design is architecture-agnostic, allowing us to further investigate the effect of different pretrained priors. As shown in Tab. 7, the 3-channel fine-tuned models generally outperform their 9-channel polarization-aware counterparts, highlighting the importance of pretrained restoration priors in overcoming the limitations of polarization-specific training. Moreover, our dual-branch design consistently achieves better performance when built upon NAFNet (Chen et al. (2022)), Restormer (Zamir et al. (2022)), MoCE-IR (Zamfir et al. (2025)), and InstructIR (Conde et al. (2024)). This not only validates the necessity of separately modeling intensity and polarization features, but also demonstrates that our framework can effectively adapt intensity-domain restoration priors to polarization restoration.

We ultimately select InstructIR as the restoration prior for two reasons. First, it provides strong restoration performance and generalization, outperforming MoCE-IR on polarization data. Second,

its lightweight backbone is easier to extend than Transformer-based architectures, allowing us to incorporate the polarization branch with relatively limited computational overhead.

Table 7: Comparison of different input and branch designs with general restoration priors. "9ch" directly processes the nine-channel input, "3ch" uses the three-channel intensity representation, and "Ours" uses our intensity–polarization dual-branch design. We report the performance of all experiments at 100k training iterations. **Bold** indicates the best performance in each subtable.

Configuration	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
NAFNet-9ch	25.788	27.869	0.818	0.696	15.948
NAFNet	33.180	28.554	0.914	0.728	15.566
NAFNet-Ours	35.468	30.626	0.944	0.771	13.577
Restormer-9ch	33.307	29.536	0.923	0.743	14.281
Restormer	32.299	29.190	0.903	0.736	15.069
Restormer-Ours	**36.574**	**31.014**	**0.951**	**0.780**	**13.246**

(a) Deblurring performance on PolDeblur dataset.

Configuration	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
MoCE-IR-9ch	28.608	26.396	0.812	0.605	22.433
MoCE-IR	29.483	28.627	0.862	0.708	20.645
MoCE-IR-Ours	31.306	29.332	0.890	0.732	18.160
InstructIR-9ch	28.799	26.483	0.816	0.594	22.624
InstructIR	29.058	28.315	0.865	0.703	20.536
InstructIR-Ours	**31.832**	**29.932**	**0.899**	**0.752**	**16.791**

(b) MoCE-IR and InstructIR restoration priors at PolarCDD settings.

A.6 Polarization Branch Design

As shown in Tab. 5, the polarization branch uses only half as many blocks as the intensity branch, as we find that this lightweight design already achieves strong performance. According to Tab. 8, halving the polarization branch leads to only a 0.1 dB drop in $PSNR_p$, while the MAE remains nearly unchanged. We further evaluate a tiny branch, where the polarization branch is deployed only at Encoder 1/2, Decoder 1/2, and the bottleneck. Its performance is consistently lower than the half configuration, indicating the importance of polarization processing across all feature levels.

To further examine the role of restoration priors, we remove IP-ATM from the full polarization branch. This results in a substantial performance drop, falling well below both the half and tiny variants. These results suggest that simply increasing the depth of the polarization branch is less effective than introducing restoration priors for overcoming the performance bottleneck.

Table 8: Ablation study on the polarization branch and feature promotion under the AIO3 setting. We report the performance of all experiments at 100k training iterations. **Bold** indicates the best performance.

Configuration	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Full polarization branch	**36.034**	**32.342**	0.934	**0.784**	**13.325**
Full branch w/o IP-ATM	35.141	31.119	0.929	0.756	14.323
Tiny polarization branch	35.744	32.017	0.934	0.777	13.573
Half polarization branch (Ours)	35.893	32.261	**0.934**	0.782	13.358

A.7 Loss Function Ablation

Our loss function is formulated as follows:

$$\begin{aligned}
&\mathcal{L}_{ang} = \left\|\hat{X} - X\right\|_1, \mathcal{L}_{stokes} = \left\|\hat{S} - S\right\|_1, \mathcal{L}_p = \|\hat{p} - p\|_1, \\
&\mathcal{L}_\phi = \left\|\cos(2\hat{\phi}) - \cos(2\phi)\right\|_1 + \left\|\sin(2\hat{\phi}) - \sin(2\phi)\right\|_1, \\
&\mathcal{L} = \mathcal{L}_{ang} + \lambda_1\mathcal{L}_{stokes} + \lambda_2\mathcal{L}_p + \lambda_3\mathcal{L}_\phi,
\end{aligned} \tag{8}$$

where $\|\cdot\|_1$ denotes the L1 norm. The GT $X = [x_{0^\circ}, x_{45^\circ}, x_{90^\circ}, x_{135^\circ}]$, $S = [S_0, S_1, S_2]$. $\lambda_1 = 0.5$, $\lambda_2 = 0.08$, and $\lambda_3 = 0.04$ are set to keep the different loss terms on comparable scales. The hat notation denotes network predictions. Although the network directly predicts the normalized Stokes representation, the corresponding intensity and polarization parameters can be analytically derived from it, allowing us to impose supervision in both the intensity and polarization parameter domains.

Tab. 9 reports the performance of different loss combinations. Progressively introducing $\mathcal{L}_{stokes}$ and $\mathcal{L}_{p,\phi}$ consistently improves the polarization-related metrics while maintaining strong intensity

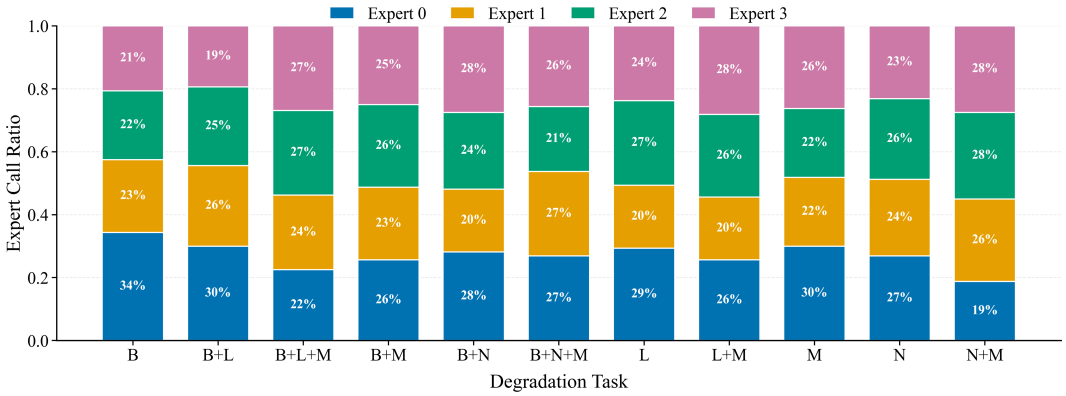


Figure 9: Activation patterns of each expert on PolarCDD. The probability indicates the proportion of images in a given task that are routed to the corresponding expert.

performance. In contrast, $\mathcal{L}_{ang}$ is critical for intensity restoration, and removing it leads to a substantial degradation in the intensity metrics. Based on these observations, we adopt the current loss configuration.

Table 9: Ablation study on different loss-function configurations under the AIO3 settings. We report the performance of all experiments at 100k training iterations. **Bold** indicates the best performance.

Configuration	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
$\mathcal{L}_{ang}$	**36.259**	31.224	**0.937**	0.757	14.986
$+\ \mathcal{L}_{stokes}$	36.129	31.543	0.936	0.766	14.518
$+\ \mathcal{L}_{p,\phi}$ (Ours)	35.893	**32.261**	0.934	**0.782**	13.358
w/o $\mathcal{L}_{ang}$	35.120	32.211	0.929	0.782	**13.357**

A.8 MoE Analysis

Fig. 9 shows the activation patterns of different MoE experts under various degradations. All experts are effectively utilized while exhibiting distinct specialization patterns.

Tabs. 10 and 11 report the effect of MoE under the AIO3 setting and on PolarCDD. With MoE, our method achieves consistent improvements across all metrics on AIO3. Since the pretrained restoration prior does not include demosaicking capability, the MoE extension helps adapt the model to this unseen degradation without disrupting the original prior. On PolarCDD, MoE brings only limited gains in intensity metrics but improves polarization-related metrics. This suggests that, under composite degradations, the additional MoE capacity makes the intensity features more transferable to the polarization domain, thereby indirectly enhancing polarization restoration.

A.9 Limitations and Future Work

We propose an AIO polarization restoration model that achieves strong performance across diverse degradation tasks. However, its gains over task-specific expert models remain marginal. In addition, the polarization branch introduces extra computational overhead, making our method particularly demanding when built upon Transformer-based backbones. This increased computational cost is a trade-off for achieving stronger restoration performance.

In addition, the restoration priors are still constrained by the distribution and scale of general restoration datasets, which remain relatively limited compared with large-scale vision corpora. Leveraging generative priors from large pretrained models therefore represents a promising direction. We have explored transferring a pretrained text-to-image diffusion model (i.e., Stable Diffusion (Rombach et al. (2022))) to polarization restoration. However, training tends to become unstable when the polarization branch remains lightweight, while scaling the polarization branch to a capacity comparable to Stable Diffusion introduces substantial deployment costs. How to effectively incorporate

Table 10: Comparison of Ours and Ours w/o MoE under the AIO3 setting. **Bold** indicates the best performance.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Ours-noMoE	32.032	29.470	0.919	0.739	14.241
Ours	**34.029**	**29.939**	**0.933**	**0.749**	**13.991**

(a) Deblurring dataset (Zhou et al. (2025a)).

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Ours-noMoE	**40.582**	29.044	**0.982**	0.753	13.380
Ours	40.512	**29.352**	0.982	**0.755**	**13.340**

(b) LLIE dataset (Zhou et al. (2023)).

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Ours-noMoE	44.852	41.846	0.983	0.936	10.717
Ours	**47.088**	**42.505**	**0.990**	**0.942**	**9.966**

(c) Demosaicking dataset (Zhou et al. (2025b)).

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Ours-noMoE	35.610	32.319	0.938	0.786	13.362
Ours	**37.518**	**32.820**	**0.949**	**0.794**	**13.009**

(d) Average performance.

Table 11: Comparison of Ours and Ours w/o MoE on the PolarCDD dataset. **Bold** indicates the best performance.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Ours-noMoE	37.143	33.075	0.951	**0.839**	14.024
Ours	**37.376**	**33.191**	**0.953**	0.826	**13.393**

(a) One degradation.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Ours-noMoE	**31.795**	30.411	0.920	0.772	16.915
Ours	31.759	**30.905**	**0.922**	**0.778**	**16.043**

(b) Dual degradations.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Ours-noMoE	29.655	28.871	0.898	0.728	17.178
Ours	**29.904**	**29.446**	**0.901**	**0.744**	**16.397**

(c) Triple degradations.

Method	$PSNR_I\uparrow$	$PSNR_p\uparrow$	$SSIM_I\uparrow$	$SSIM_p\uparrow$	MAE $\downarrow$
Ours-noMoE	33.351	31.100	0.927	0.788	15.912
Ours	**33.464**	**31.471**	**0.929**	**0.789**	**15.144**

(d) Average performance.

large-scale generative priors into polarization restoration therefore remains an important direction for future research.

A.10 Additional Visual Results

Figs. 10-13 present additional qualitative results. For deblurring (Fig. 11), we evaluate our method on the real-world PV dataset (Li et al. (2026b)), a vehicle-mounted polarization dataset in which motion blur frequently occurs. Our method removes motion blur more effectively than the compared approaches. We also evaluate on a real-world polarization mosaic dataset ((Luo et al. (2025)), Fig. 12), where our method produces the cleanest results and achieves visual quality comparable to task-specific expert models. Fig. 13 further presents results under composite degradations, showing that our method suppresses polarization-domain degradations more effectively than other AIO restoration models.

In addition, Fig. 14 demonstrates the necessity of jointly addressing composite degradations. Simultaneous low-light enhancement and demosaicking better recover image textures while reducing noise in the polarization parameters.

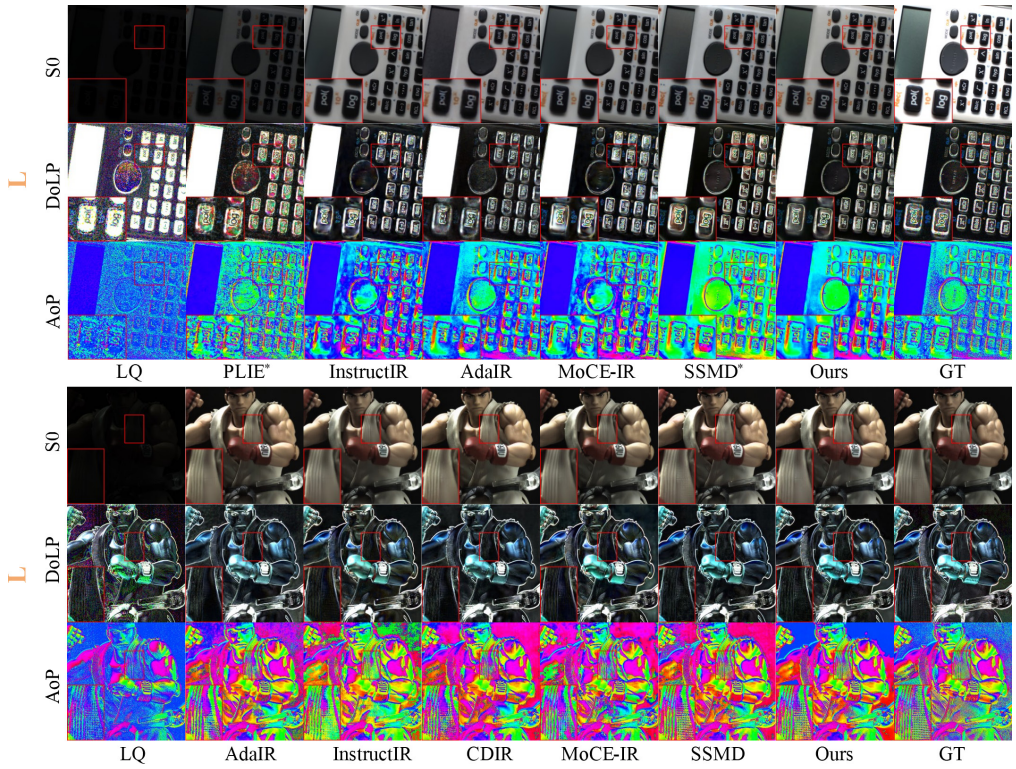


Figure 10: Visual comparisons of different methods on the dataset LLCP (Xu et al. (2022)) and PolarCDD. L is low-light. The superscript * indicates task-specific weights.

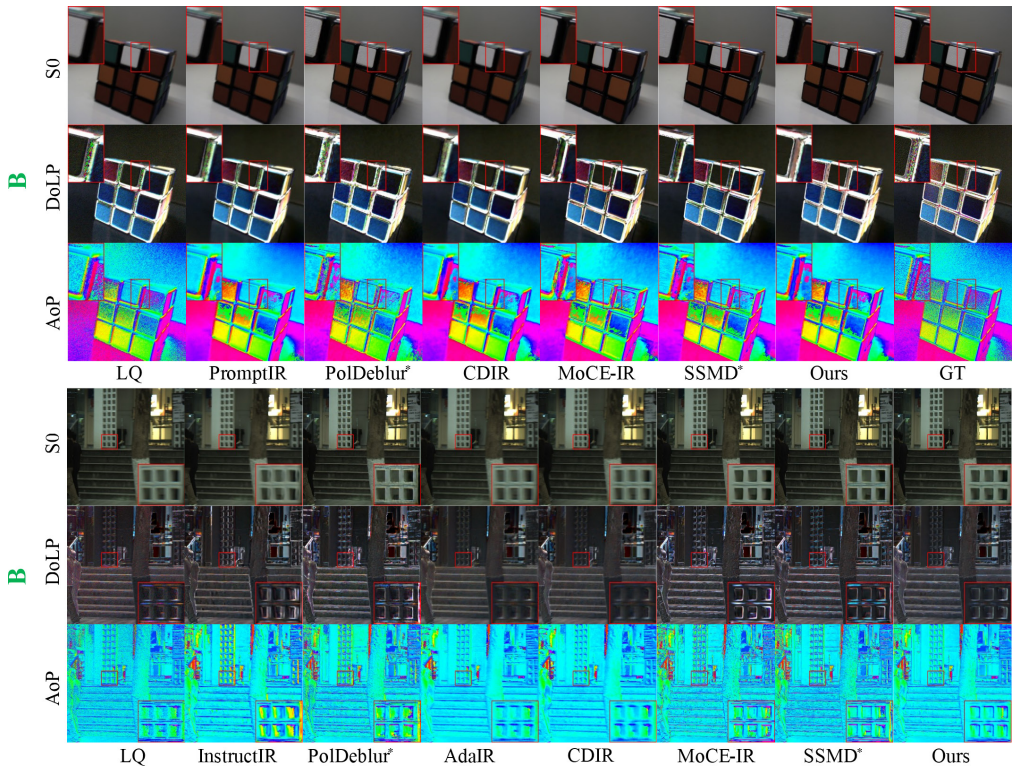


Figure 11: Visual comparisons of different methods on the dataset PolDeblur (Zhou et al. (2025a)) and real-world dataset PV (Li et al. (2026b)). B is blur. The superscript * indicates task-specific weights.

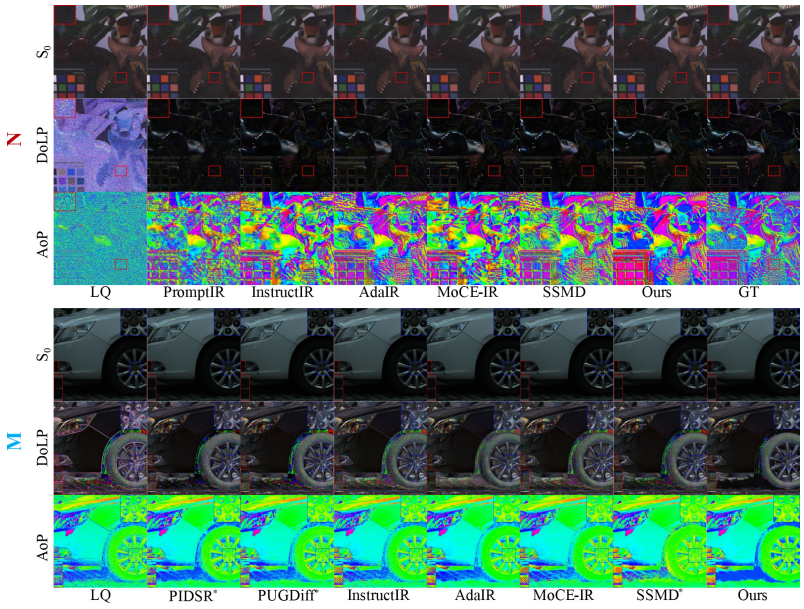


Figure 12: Visual comparisons of different methods on the dataset PolarCDD and real-world dataset (Luo et al. (2025)). N is noise and M is mosaic. The superscript * indicates task-specific weights.

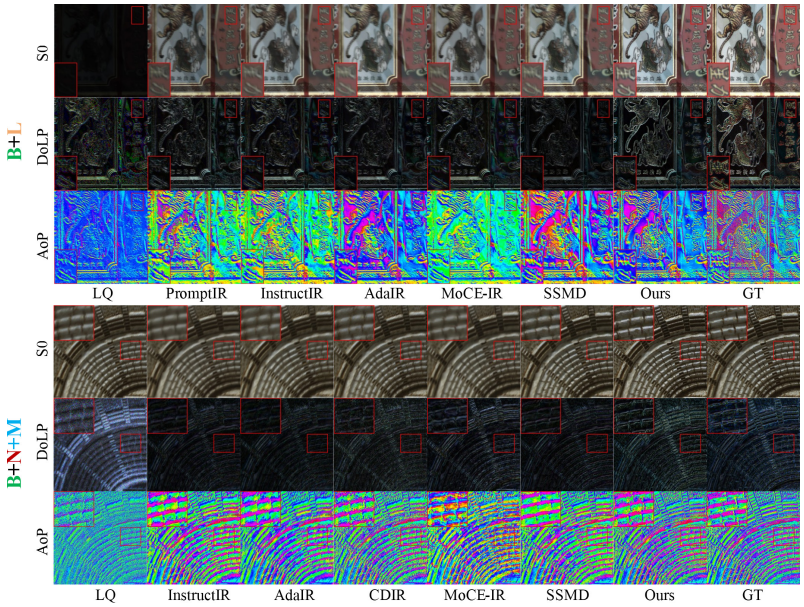


Figure 13: Visual comparisons of different methods on the dataset PolarCDD. B is blur, L is low light, N is noise and M is mosaic.

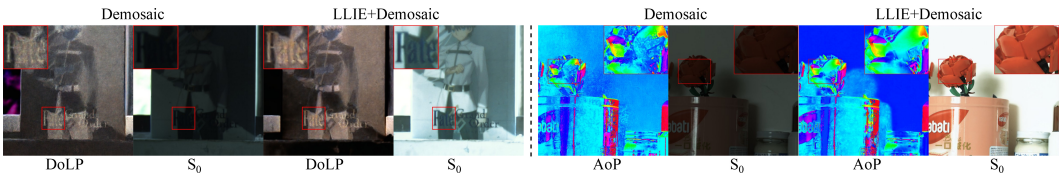


Figure 14: Text-conditioned restoration behavior.