Polarization Uncertainty-Guided Diffusion Model for Color Polarization Image Demosaicking

Chenggong Li[1,2], Yidong Luo[3,4], Junchao Zhang[1,2*], Degui Yang[1,2*]
[1]School of Automation, Central South University, Changsha, China
[2]Hunan Provincial Key Laboratory of Optic-Electronic Intelligent Measurement and Control, Changsha, China
[3]Zhejiang University, Hangzhou, China
[4]School of Engineering, Westlake University, Hangzhou, China
{244603040, junchaozhang, degui.yang}@csu.edu.cn, luoyidong@westlake.edu.cn

Abstract

Color polarization demosaicking (CPDM) aims to reconstruct full-resolution polarization images of four directions from the color-polarization filter array (CPFA) raw image. Due to the challenge of predicting numerous missing pixels and the scarcity of high-quality training data, existing network-based methods, despite effectively recovering scene intensity information, still exhibit significant errors in reconstructing polarization characteristics (degree of polarization, DOP, and angle of polarization, AOP). To address this problem, we introduce the image diffusion prior from text-to-image (T2I) models to overcome the performance bottleneck of network-based methods, with the additional diffusion prior compensating for limited representational capacity caused by restricted data distribution. To effectively leverage the diffusion prior, we explicitly model the polarization uncertainty during reconstruction and use uncertainty to guide the diffusion model in recovering high error regions. Extensive experiments demonstrate that the proposed method accurately recovers scene polarization characteristics with both high fidelity and strong visual perception.

Code — https://github.com/JJGNB/PUGDiff

1 Introduction

Polarization imaging can reveal unique object properties, such as material and reflectivity. By fully leveraging the degree of polarization (DOP) and angle of polarization (AOP), polarization imaging has been increasingly applied in fields like object detection, reflectivity removal, and 3D reconstruction (Luo, Zhang, and Li 2025; Yao et al. 2025; Wu et al. 2025). In practical applications, the photographer captures a mosaic array with polarization information using a division-of-focal-plane (DOFP) camera (Rebhan, Rosenberger, and Notni 2019). Fig. 1 shows the operation of the DOFP camera, which samples pixels from four polarization directions ($0°$, $45°$, $90°$, $135°$) into a mosaic array following an RGGB order, and the sampling process can be expressed as:

$$y = Ax + n, \tag{1}$$

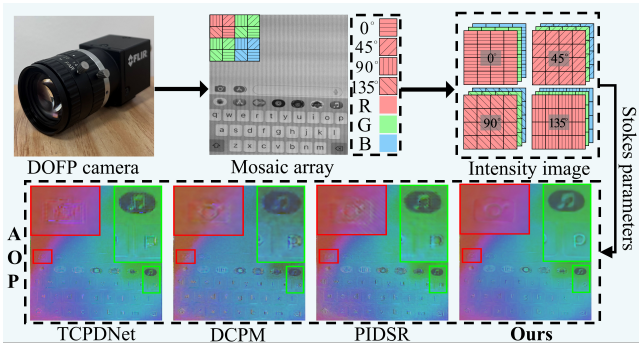

Figure 1: The workflow of color polarization imaging. The second row presents the AOP results of different CPDM methods, where AOP reflects the accuracy of polarization property reconstruction. Our method exhibits the least noise interference and achieves the most visually faithful results.

where A denotes the sampling operator, $x \in \mathbb{R}^{12\times H\times W}$, $y \in \mathbb{R}^{1\times H\times W}$ and $n \in \mathbb{R}^{1\times H\times W}$ represent the ground truth full-resolution color image, the mosaic array and the sensor noise, respectively. To obtain a full-resolution color image for better analysis of polarization characteristics, we need to recover the pixels lost during the sampling from the mosaic array, i.e., color polarization demosaicking (CPDM).

CPDM is a challenging inverse problem in polarization imaging. Some prior-based interpolation algorithms (Morimatsu et al. 2020; Wu et al. 2021) and model-based optimization algorithms (Wen, Zheng, and Lu 2021; Luo et al. 2024) have attempted to address this issue, but the reconstruction accuracy remains limited. Using neural networks f_θ to address CPDM is a promising approach (Zhang et al. 2018; Sun, Zhang, and Liang 2021) and the demosaicking process can be formulated as:

$$\begin{aligned} \tilde{x} &= I(y) \\ x &= f_\theta(\tilde{x}) + \varepsilon\eta \end{aligned}, \tag{2}$$

where I denotes the initial interpolation that restores the mosaic array into four direction intensity images $\tilde{x} \in \mathbb{R}^{12\times H\times W}$, ε follows the normal distribution with zero-mean and unit-variance, and η denotes the demosaicking error, which we term the uncertainty in the demosaicking process. By employing a predefined sampling operator to gener-

*Corresponding Author.

ate simulated datasets, network-based methods can learn to recover missing pixels from the data. However, even though these methods can reconstruct the intensity image well, they exhibit significant errors in recovering polarization characteristics like AOP, as shown in Fig. 1. This is because current network-based methods learn demosaicking capabilities from limited datasets (Morimatsu et al. 2020; Qiu et al. 2021; Li et al. 2025), which lack scene diversity and adequate scale, making the learned data priors insufficient to support these methods in achieving superior performance.

To achieve better demosaicking performance, we propose a polarization uncertainty-guided diffusion model, namely PUGDiff. PUGDiff is a dual-branch network shown in Fig. 2, a base branch trained from scratch provides the model with fundamental demosaicking capabilities, while the other branch consists of a text-to-image model (i.e., Stable Diffusion, SD), called SD branch. The SD branch is trained using Low-Rank Adaptation (LoRA) (Hu et al. 2022) to maximally preserve its diffusion prior, which helps overcome the limitations of existing data priors. We explicitly model a polarization uncertainty estimator to effectively leverage the two branches. Specifically, in regions with low uncertainty, the base branch can produce high-fidelity results, while in regions with high uncertainty, the SD branch helps correct polarization errors. The two branches are fused by a fusion module to generate the final output, which is trained using an uncertainty-guided loss, the loss transfers the uncertainty map into weights for different loss terms and thus directs the model to flexibly allocate the contributions of the two branches to the final result based on polarization uncertainty.

We summarize the primary contributions as follows:

- We introduce the text-to-image model into the CPDM, leveraging its diffusion prior learned from large-scale natural images to break through the limitations of the original data distribution.
- We model the uncertainty in the demosaicking process based on polarization characteristics, and transform it into an uncertainty-guided loss to supervise network fusion, enabling the network to adaptively select the dominant branch based on the polarization uncertainty.
- Qualitative and quantitative experiments on both simulated and real-world images demonstrate that our method achieves state-of-the-art (SOTA) performance.

2 Related Work

Color Polarization Demosaicking. Early research efforts on CPDM primarily relied on interpolation algorithms (Li et al. 2019; Morimatsu et al. 2020) or dictionary learning (Wen, Zheng, and Lu 2021; Luo, Zhang, and Tian 2023; Luo et al. 2024). With the blooming of deep learning, researchers have gradually shifted their focus toward network-based approaches. Nguyen et al. propose a two-stage demosaicking convolutional neural network (CNN) (Nguyen et al. 2022), achieving joint demosaicking for color and polarization. Guo et al. design an attention-based progressive discrimination generative adversarial network (GAN) to alleviate artifacts during the demosaicking process (Guo et al. 2024). Zhou et al. incorporate polarization characteristics into the network design, achieving mutual enhancement between demosaicking and polarization super-resolution through a carefully designed pipeline (Zhou et al. 2025). Li et al. propose a demosaicking customized diffusion model, which achieves more robust results through distribution modelling (Li et al. 2025). However, existing network-based methods only learn demosaicking capabilities from small-scale simulated datasets, leading to a performance bottleneck that makes them struggle to handle more complex and diverse scenarios.

Diffusion Prior in Inverse Problems. Using diffusion priors to assist in solving inverse problems is currently a hot topic in low-level vision. Many diffusion paradigms have been successfully applied to tasks like denoising, deblurring, inpainting, compressed sensing, super-resolution, etc (Kawar et al. 2022; Sun et al. 2023; Wu et al. 2024b; Chihaoui, Lemkhenter, and Favaro 2024; Sun et al. 2025; Chen et al. 2025b,a). Chung et al. integrate inverse problem solving into the diffusion solver through approximation of the posterior sampling, enabling effective handling of noisy inverse problems (Chung et al. 2022). Wang et al. utilize a pre-trained diffusion model to progressively restore the null-space of images while ensuring data consistency by range-space constraints (Wang, Yu, and Zhang 2022). Wu et al. introduce LoRA into Stable Diffusion and utilize variational score distillation loss to align the diffusion model with natural image priors (Wu et al. 2024a). Yue et al. design a deep noise predictor based on diffusion inversion, enabling noise customization for super-resolution and achieving high perceptual quality results (Yue, Liao, and Loy 2025). These achievements motivate us to explore the application of diffusion priors to polarization demosaicking, aiming to overcome the performance bottleneck caused by insufficient datasets.

3 Methodology

In this paper, we propose a dual-branch diffusion model guided by polarization uncertainty. As shown in Fig. 2, we first train a demosaicking network f_b from scratch, using a CNN-Transformer (Zamir et al. 2022) hybrid U-Net architecture. f_b serves as the algorithm's base branch, aiming to meet the high fidelity requirements of the CPDM task. The process of the base branch can be expressed as:

$$x_b = f_b(\tilde{x}), \tag{3}$$

$x_b \in \mathbb{R}^{12\times H\times W}$ is then fed into a modified SD (SD branch) to leverage its inherent diffusion prior for further refinement, the above process can be expressed as:

$$x_{sd} = f_{sd}(x_b). \tag{4}$$

To effectively utilize the two branches for better demosaicking performance, we develop a polarization uncertainty model and use the uncertainty to guide the fusion of the two branches' outputs. The fusion process can be formulated as:

$$x_{final} = f_{fuse}(x_b, x_{sd}). \tag{5}$$

In the following, we detail the modifications made to the SD and the derivation of the polarization uncertainty model. The loss functions for each module are then explained.

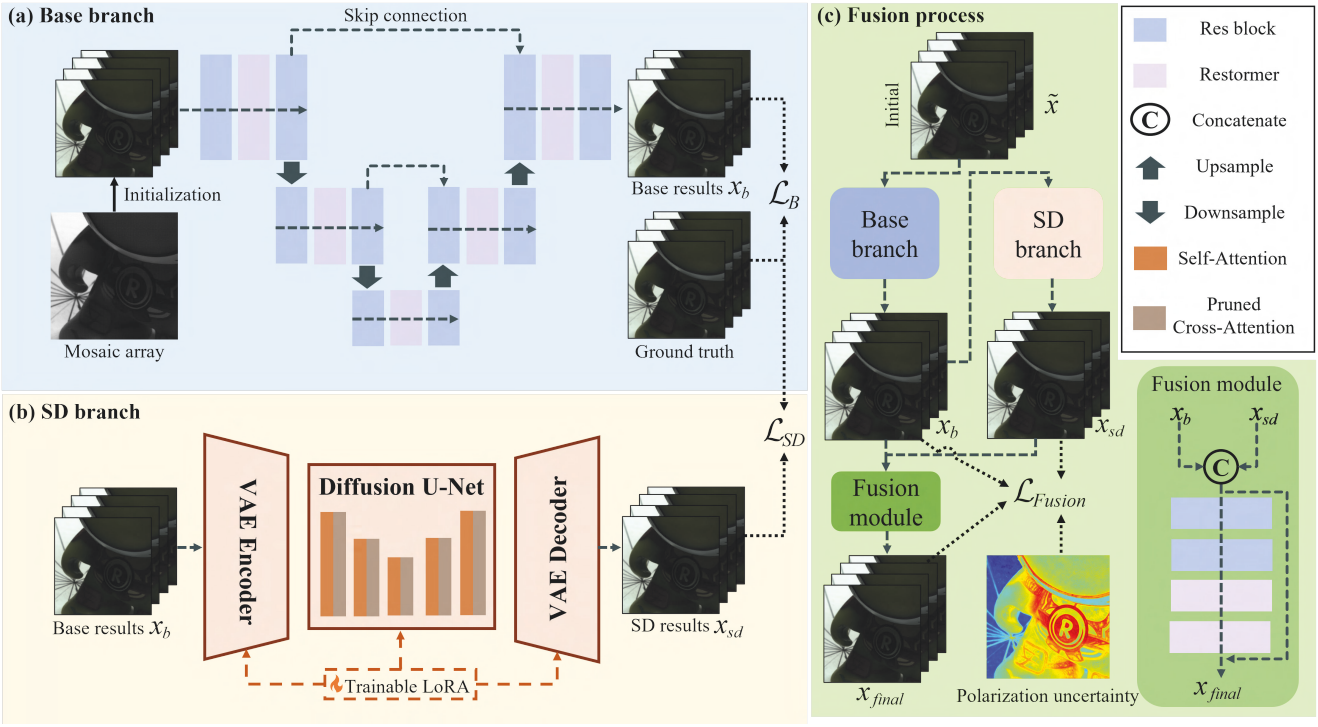

Figure 2: The framework of the proposed PUGDiff. (a) The architecture of the base branch, the mosaic array is initialized as a low-quality color polarization image and fed into the base branch for processing, and the network is updated via $\mathcal{L}_B$; (b) The architecture of the SD branch, results of the base branch is further refined by the SD branch. LoRA layers are injected into SD to adapt it to the CPDM task, while the cross-attention modules are pruned to improve efficiency, the LoRA parameters are updated via $\mathcal{L}_{SD}$; (c) The fusion process of the two branches, which is also used during inference. The contribution of each branch is adaptively weighted by polarization uncertainty. In low-uncertainty regions, the base branch is favored for high fidelity; in high-uncertainty regions, the SD branch is used to enhance the polarization property.

3.1 Modified SD

The SD consists of a variational autoencoder (VAE) and a latent diffusion model (LDM) shown in Fig. 2 (b). Since SD processes natural images, we treat the four direction intensity images as four separate batches and feed them into the SD model sequentially. The data flow of SD can be expressed as:

$$\begin{aligned} z_b &= E(x_b), \\ z_{sd} &= LDM(z_b), \\ x_{sd} &= D(z_{sd}), \end{aligned} \tag{6}$$

where E and D denote the VAE encoder and decoder, respectively, with z_b and z_{sd} representing the image features in the latent space. The LDM obtains the final result using only one step, which can be expressed as:

$$z_{sd} = LDM(z_b) \triangleq \frac{z_b - \beta_T \psi(z_b, T)}{\alpha_T}, \tag{7}$$

where T is the time, α_T and β_T denotes the noise scheduler of the SD, and ψ represents the diffusion U-Net. In the context of CPDM, text prompts are unnecessary for the SD, hence, we remove the text encoder and cross-attention module in the denoising U-Net, improving efficiency without compromising demosaicking performance.

Instead of training from scratch, we inject trainable LoRA modules into the VAE and diffusion U-Net, and fine-tune them on simulated datasets to accommodate the CPDM task. Training with LoRA effectively preserves the diffusion prior of the SD, enabling the network to leverage additional knowledge to recover lost pixels. This helps overcome the performance bottleneck caused by limited CPDM datasets and further enhances the models' generalization capability.

3.2 Polarization Uncertainty Model

Although SD can produce high-quality perceptual results, it compresses the image, leading to information loss and failing to meet the high-fidelity requirements of CPDM. In contrast, while the base branch recovers four polarization directions well, the polarization properties (DOP and AOP) are derived through nonlinear computation, which amplifies reconstruction errors. This finding inspires us to design a rule for fusing the outputs of the two branches. For regions with small reconstruction errors from the base branch, we prefer its results to ensure fidelity; for regions with large errors, we place more confidence in the SD output to improve the visual faithfulness of polarization properties.

Therefore, from the perspective of polarization reconstruction error (uncertainty), we establish a polarization un-

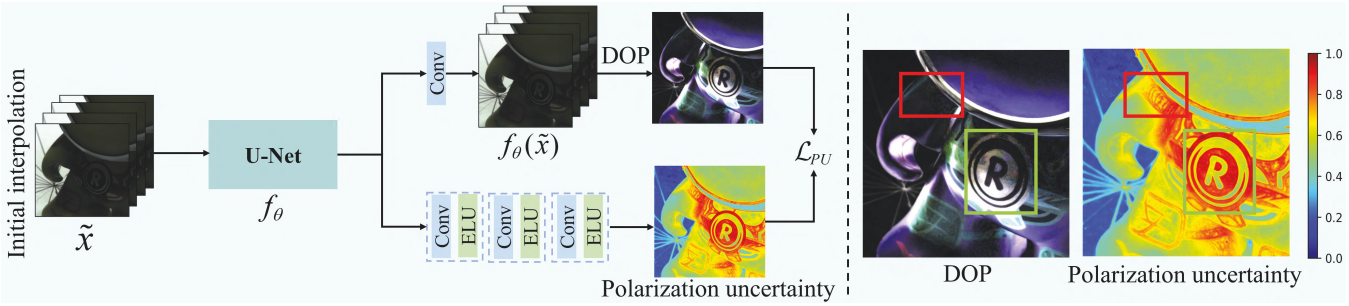

Figure 3: The architecture of the uncertainty estimation network. The network's backbone shares the same architecture as the base branch, but augmented with an additional estimation head to output polarization uncertainty. The entire network is updated via $\mathcal{L}_{PU}$, and during the training phase in Fig. 2 (c), only the polarization uncertainty is used. The right part of the figure shows the reconstructed DOP and the polarization uncertainty. High values in the heatmap indicate regions where DOP reconstruction is inaccurate.

certainty model to explicitly characterize the uncertainty. We reformulate Eq. (2) as:

$$x_i = f_\theta(\tilde{x}_i) + \varepsilon\eta, i = 0°, 45°, 90°, 135°, \tag{8}$$

where η is the intensity uncertainty, and ε follows the normal distribution with zero-mean and unit-variance, the subscript i indicates the polarization images at different directions. Eq. (8) can be expressed as $x_i \sim N(f_\theta(\tilde{x}_i), \eta^2)$. According to Collett's literature (Collett 2005), the Stokes parameters are defined as:

$$\begin{aligned} S_0 &= \tfrac{1}{2}(x_{0°} + x_{45°} + x_{90°} + x_{135°}) \\ S_1 &= x_{0°} - x_{90°} \\ S_2 &= x_{45°} - x_{135°} \\ DOP &= \tfrac{\sqrt{{S_1}^2 + {S_2}^2}}{S_0} \\ AOP &= \tfrac{1}{2}\arctan(\tfrac{S_2}{S_1}) \end{aligned} . \tag{9}$$

Substituting Eq. (8) into Eq. (9) and using the properties of the normal distribution gives:

$$S_0 \sim N(\tilde{S}_0, \eta^2), S_1 \sim N(\tilde{S}_1, 2\eta^2), S_2 \sim N(\tilde{S}_2, 2\eta^2), \tag{10}$$

where

$$\begin{aligned} &\tilde{S}_0 \triangleq \tfrac{1}{2}(f_\theta(\tilde{x}_{0°}) + f_\theta(\tilde{x}_{45°}) + f_\theta(\tilde{x}_{90°}) + f_\theta(\tilde{x}_{135°})) \\ &\tilde{S}_1 \triangleq f_\theta(\tilde{x}_{0°}) - f_\theta(\tilde{x}_{90°}), \tilde{S}_2 \triangleq f_\theta(\tilde{x}_{45°}) - f_\theta(\tilde{x}_{135°}) \end{aligned} . \tag{11}$$

Since f_θ has already reconstructed the intensity image (S_0) effectively, we can assume that the estimate of S_0 is accurate [1], and thus use $\tilde{S}_0$ as an approximation of the ground truth S_0, i.e. $\tilde{S}_0 \triangleq S_0$. Subsequently, we can derive:

$$\tfrac{S_1}{S_0} \sim N(\tfrac{\tilde{S}_1}{\tilde{S}_0}, \tfrac{2\eta^2}{\tilde{S}_0^2}), \tfrac{S_2}{S_0} \sim N(\tfrac{\tilde{S}_2}{\tilde{S}_0}, \tfrac{2\eta^2}{\tilde{S}_0^2}). \tag{12}$$

Based on Eqs. (9) and (12), the DOP (denoted as ϕ) follows a Rice distribution of the form (Hwang et al. 2025):

$$\phi \sim Rice(\tilde{\phi}, \tfrac{\sqrt{2}\eta}{S_0}), \tag{13}$$

[1]This observation is also supported by the visual comparisons in the experiment section, where most CPDM methods achieve S_0 results close to the ground truth.

and we denote $\frac{\sqrt{2}\eta}{S_0}$ as η_p, which represents the polarization uncertainty we aim to model in this paper. Compared to the intensity uncertainty η, η_p provides more direct polarization information for the subsequent fusion process.

The probability density function (PDF) of Eq. (13) can be expressed as:

$$P(\phi|\tilde{\phi}, \eta_p) = \tfrac{\phi}{{\eta_p}^2}\exp(-\tfrac{(\phi^2+\tilde{\phi}^2)}{2{\eta_p}^2})I_0(\tfrac{\phi\tilde{\phi}}{{\eta_p}^2}), \tag{14}$$

where I_0 denotes the modified Bessel function of the first kind, order zero. In applications, the signal amplitude is much larger than the uncertainty; therefore, according to the definition of the modified Bessel function (Bowman 2012), I_0 has an asymptotic form for large arguments, and Eq. (14) can be approximated as [2] :

$$P(\phi|\tilde{\phi}, \eta_p) \approx \sqrt{\tfrac{\phi}{2\pi\tilde{\phi}{\eta_p}^2}}\exp(-\tfrac{(\phi-\tilde{\phi})^2}{2{\eta_p}^2}). \tag{15}$$

To obtain η_p, we use an uncertainty estimation network to directly predict it through supervised learning shown in Fig. 3. The estimation network's weights are initialized from the base branch for consistency, and the optimization objective is to minimize the negative log likelihood of Eq. (15)[3], i.e.,

$$\mathcal{L}_{PU} = \tfrac{1}{2}(\ln\tilde{\phi} - \ln\phi) + 2s + \tfrac{1}{2}\exp(-2s)\left\|\tilde{\phi} - \phi\right\|_2^2, \tag{16}$$

where s denotes the log polarization uncertainty $s = \ln\eta_p$.

3.3 Uncertainty Guided Loss Function

We use the mean square error (MSE) loss to train both the base branch f_b and SD branch f_{sd}, i.e.,

$$\mathcal{L}_{B\&SD} = \|x - x_r\|_2^2, x_r \in \{x_b, x_{sd}\}, \tag{17}$$

where $x \in \mathbb{R}^{12\times H\times W}$ denotes the ground truth of four directions.

[2]The detailed derivation can be found in the supplementary material.

[3]The detailed derivation can be found in the supplementary material.

	Method	$PSNR_{mean}$	$PSNR_{S_0}$	$PSNR_{DOP}$	MAE	$SSIM_{mean}$	$SSIM_{S_0}$	$SSIM_{DOP}$
MQ's	Polanalyser	38.2149	35.2560	32.5532	14.8220	0.9653	0.9541	0.8387
	NLCSR	36.8877	33.5610	34.0964	11.1297	0.9523	0.9310	0.8557
	CPDNet	39.4374	36.6062	32.6586	12.5138	0.9700	0.9631	0.8465
	TCPDNet	42.5602	39.7694	36.0613	10.1765	0.9815	0.9765	0.8921
	DCPM	42.0221	40.0263	36.1496	11.7376	0.9807	0.9759	0.8958
	PIDSR	42.4990	**40.6269**	36.5413	9.6463	0.9822	**0.9794**	0.8950
	Ours	**42.6794**	39.4785	**37.4700**	**9.2029**	**0.9827**	0.9768	**0.9059**
PIDSR's	Polanalyser	41.0285	39.4167	35.8349	17.9892	0.9655	0.9621	0.8347
	NLCSR	40.0285	38.5526	36.0323	16.0549	0.9601	0.9569	0.8775
	CPDNet	41.4389	39.6154	37.0999	**14.9234**	0.9694	0.9639	0.8529
	TCPDNet	43.1434	41.2190	39.1320	15.5998	0.9785	0.9734	0.8883
	DCPM	43.4987	41.7152	38.4340	15.5325	0.9804	0.9750	0.8791
	PIDSR	43.0540	41.2950	39.8428	16.4804	0.9792	0.9741	0.9015
	Ours	**44.2325**	**41.9823**	**40.6696**	17.0324	**0.9815**	**0.9753**	**0.9107**
DCPM's	Polanalyser	36.6754	35.4464	29.3446	17.2526	0.9653	0.9652	0.7344
	NLCSR	37.8591	36.8419	29.5389	15.6778	0.9688	0.9693	0.7294
	CPDNet	39.3842	39.4640	29.9120	16.2399	0.9767	0.9806	0.7463
	TCPDNet	40.2257	40.7204	30.5104	15.9773	0.9789	0.9829	0.7532
	DCPM	43.2512	**43.7558**	32.5309	14.7762	0.9857	0.9870	0.7765
	PIDSR	40.5454	42.5008	31.7020	14.5157	0.9815	0.9856	0.7624
	Ours	**43.3532**	42.8488	**33.2396**	**12.6792**	**0.9871**	**0.9877**	**0.7934**

Table 1: Quantitative comparison with state-of-the-art methods. The higher value means better effects. The bold black indicates the best result. The proposed method shows the overall optimal performance.

To better leverage the strengths of both branches, we define the following uncertainty-guided fusion loss shown in Fig. 2 (c):

$$\mathcal{L}_{Fusion} = \bar{s} \left\| x_{final} - x_{sd} \right\|_2^2 + (1 - \bar{s}) \left\| x_{final} - x_b \right\|_2^2, \tag{18}$$

where $\bar{s}$ is the normalized s. The above function fuses the two branches based on polarization uncertainty. In regions with small η_p, the base branch ensures high fidelity, while in regions with large polarization estimation uncertainty, the SD branch is emphasized to improve the quality of polarization reconstruction. $\mathcal{L}_{Fusion}$ implicitly incorporates polarization uncertainty as a gating mechanism within the fusion module. During inference, there is no need to explicitly output the uncertainty, thereby reducing the network's forward computation.

4 Experiments

4.1 Implementation details

For the experiment part, the proposed PUGDiff is trained with Monno's (Morimatsu et al. 2020), Qiu's (Qiu et al. 2021) and PIDSR (Zhou et al. 2025) synthetic datasets. Cropping, flipping and rotation are applied to augment the training data. The final training set consists of 28020 image patches of size 256×256. The mosaic arrays are first interpolated using Polanalyser (Maeda 2019) for initial reconstruction before being fed into the network. To ensure training stability, the base branch, SD branch, and uncertainty estimation network are trained separately. For the rank of LoRA in the SD branch, we set it to 4. After training the above modules, we proceed to train the fusion module. All models are trained for 200k iterations with a fixed learning rate of 5×10^{-5} and a batch size of 2. All experiments are conducted on an NVIDIA RTX 4090 GPU.

4.2 Experimental configurations

We conduct tests on the Monno's, Qiu's (MQ), PIDSR, and DCPM (Li et al. 2025) simulated datasets, and validate the generalization on the real-captured mosaic array. The proposed method is compared with Polanalyser (Maeda 2019), NLCSR (Luo et al. 2024), CPDNet (Wen et al. 2019), TCPDNet (Nguyen et al. 2022), DCPM (Li et al. 2025) and PIDSR (Zhou et al. 2025). We evaluate the quality of CPDM with PSNR and SSIM. The subscript "mean" indicates that the metric is calculated on the intensity images of four directions and then averaged, while "S_0", "DOP" and "AOP" denote the metrics calculated for S_0, DOP and AOP, respectively. The higher value of both metrics means better CPDM effects. We also use Mean Angular Error (MAE) to evaluate AOP quality, where lower values indicate better performance.

4.3 Comparison with State-of-the-art Methods

The CPDM results of all methods are shown in Fig. 4. Although these methods successfully reconstruct the intensity images (S_0), the AOP and DOP, which reflect the scene's polarization properties, still exhibit significant errors. On the simulated data, our method produces the sharpest text edges and exhibits the smallest deviation from the ground truth. On the real-captured image, most algorithms are overwhelmed by noise. The closest performer is PIDSR; upon magnification, our method shows clearer DOP details on clothing and facial regions, and more distinct AOP compared to PIDSR. Table 1 presents the quantitative results on three datasets, where our method achieves the best overall performance, particularly in DOP and AOP metrics. This improvement is attributed to the polarization uncertainty-guided diffusion prior, which enables our method to overcome the performance ceiling of conventional network-based methods.

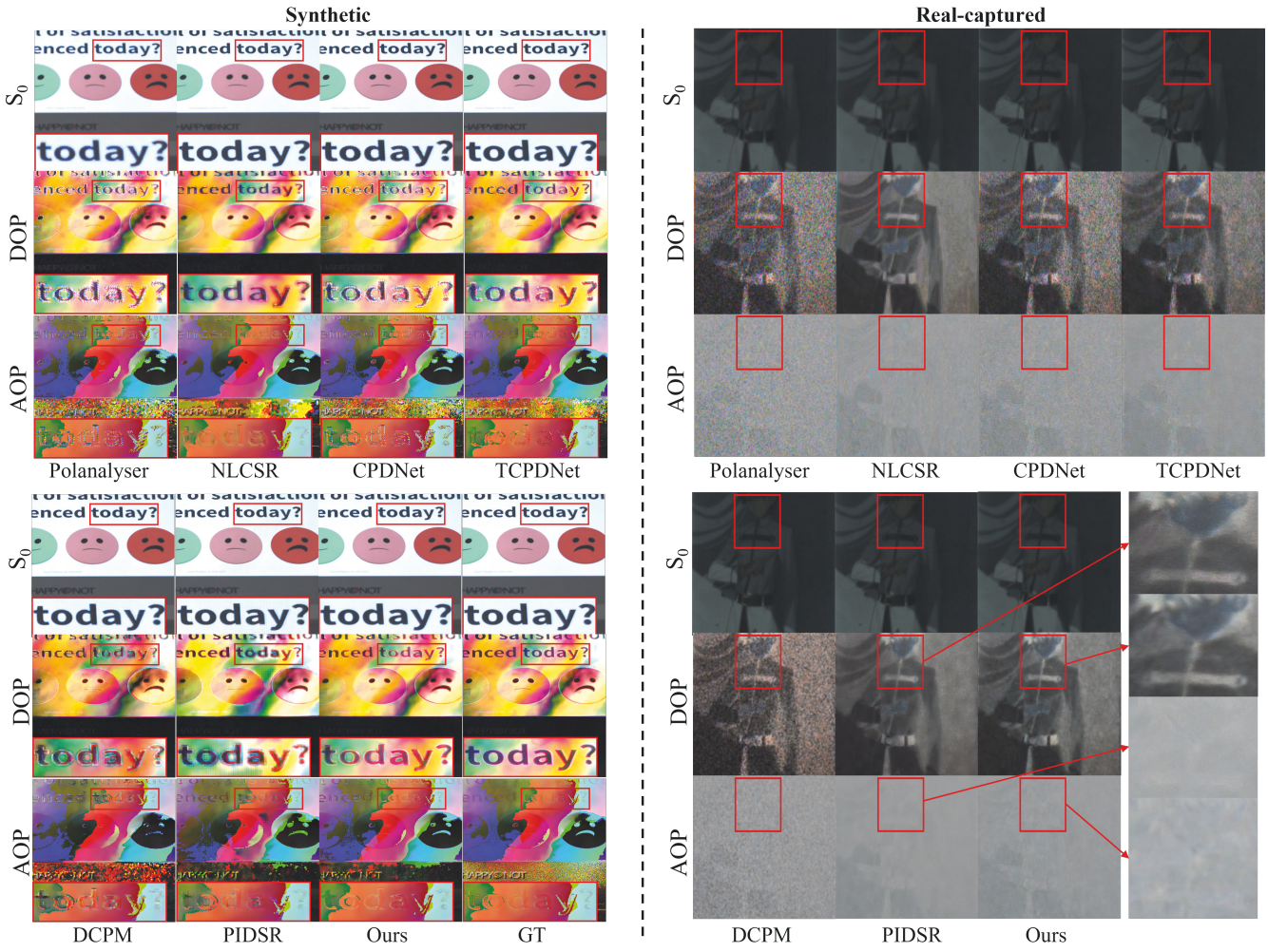

Figure 4: Visual comparisons for CPDM of different methods. Both on synthetic and real-captured images, our method achieves the most outstanding polarization reconstruction performance in terms of AOP and DOP.

4.4 Ablation Studies

Importance of polarization uncertainty in the dual-branch framework. The network adjusts the contribution of the base branch and the SD branch to the final result under the guidance of polarization uncertainty. In Fig. 5, we visualize the outputs of both branches along with the corresponding polarization uncertainty maps. The polarization uncertainty in the text regions is relatively high, which is also reflected in the AOP and DOP results of the base branch. Consequently, the network assigns greater weight of the SD branch in these areas. In contrast, for regions with low uncertainty (as highlighted in the green box), the base branch is capable of accurately reconstructing the polarization characteristics. In such areas, the SD branch tends to produce over-smoothed results, degrading demosaicking accuracy. Therefore, the network gives higher confidence to the base branch, enhancing overall fidelity.

Superiority of modelling uncertainty from the polarization perspective. In the main text, we explicitly model polarization uncertainty based on the PDF of DOP. To demonstrate its effectiveness in recovering polarization information, we also derive uncertainty for the four polarization directions and for S_0 using Eqs. (8) and (10)[4], respectively. The uncertainty maps are shown in Fig. 6. The uncertainty derived from intensity images reflects errors in reconstruct-

Method	$PSNR_{S_0}$	$PSNR_{DOP}$	$PSNR_{AOP}$
Intensity	**42.1747**	40.2787	15.8751
Intensity (S_0)	42.1165	40.3096	16.4514
Polarization	41.9823	**40.6696**	**16.8207**

Table 2: Ablation study on the forms of uncertainty. "Intensity", "Intensity (S_0)" and "Polarization" denote the uncertainty computed from the four direction polarization images, S_0, and DOP, respectively. The bold black indicates the best result.

[4]The derivation can be found in the supplementary material.

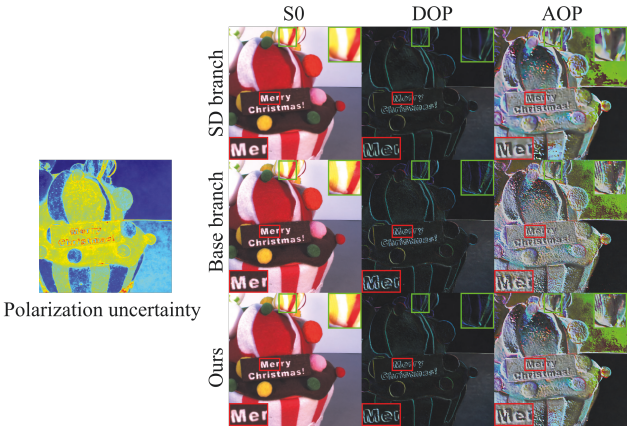

Figure 5: Visualization of results from different branches. Our method fuses the two branches to achieve complementary strengths: using the SD branch in high-uncertainty regions to improve reconstruction, and the base branch in low-uncertainty regions to preserve fidelity.

Method	$PSNR_{S_0}$	$PSNR_{DOP}$	$PSNR_{AOP}$
All weight	–	–	–
Only U-Net	40.8976	40.2579	16.0951
rank 2	–	–	–
rank 4 (Ours)	41.9823	**40.6696**	**16.8207**
rank 8	**42.0145**	40.6652	16.5832

Table 3: Ablation study on the SD configuration. "All weight" denotes full-parameter fine-tuning, "Only U-Net" indicates that only the diffusion U-Net is updated, "rank" is a hyperparameter of LoRA and "–" indicates training failure. The bold black indicates the best result.

ing intensity information. However, in the context of CPDM, our primary concern is the error in polarization characteristics. Polarization uncertainty provides information correlated with polarization (DOP) reconstruction error, enabling more accurate DOP recovery compared to the other two uncertainty measures, and Table 2 further confirms this observation.

Impact of SD configuration. We enable the SD model to adapt to the CPDM task while preserving its diffusion prior by injecting LoRA layers into both the VAE and U-Net components of the SD. Table 3 shows several alternative configurations we explored for the SD model. The "All weight" setting, in which all parameters are fine-tuned, leads to training failure due to the limited data distribution. When LoRA is injected only into the U-Net, performance degrades, indicating that updating the VAE is also crucial. Therefore, we choose to apply LoRA to both the U-Net and VAE. We also investigate the effect of LoRA rank on performance, when the rank is set to 2, training becomes unstable; when the rank is 4 or higher, performance plateaus. Since larger ranks introduce higher computational and memory overhead, we ultimately set the rank to 4.

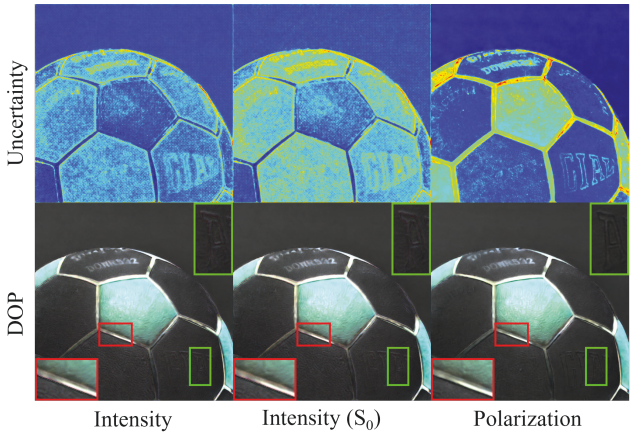

Figure 6: Visualization of different forms of uncertainty. Polarization uncertainty provides more direct information about polarization reconstruction, thus leading to the most significant improvement in DOP.

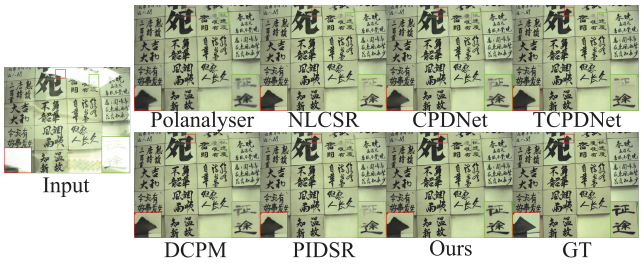

Figure 7: Visual comparisons for PRR of different methods. Our method achieves the clearest reflection removal results.

4.5 Applications

To fully demonstrate the demosaicking performance of our method, we conduct the polarization-based reflection removal (PRR) experiment shown in Fig. 7. After demosaicking, all images are processed using the method provided by PolarFree (Yao et al. 2025) for reflection removal. Our method exhibits the fewest artifacts and the clearest text, demonstrating its strong performance in the PRR task.

5 Conclusion

In this paper, we transfer the powerful diffusion prior to the CPDM task via LoRA. Derived from large-scale natural image distributions, the diffusion prior helps overcome the performance bottleneck imposed by the scarcity of demosaicking training data. We propose a dual-branch demosaicking architecture, the outputs of the two branches are adaptively fused through explicitly modelled polarization uncertainty. The uncertainty guides the diffusion branch to refine regions with large polarization errors, while using the base branch in low error regions to maintain fidelity. The proposed method achieves state-of-the-art performance across multiple benchmarks, significantly improving the accuracy of CPDM.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (Grant No. 62105372), the National Natural Science Foundation of China (Grant No. 62471497), the Fundamental Research Foundation of National Key Laboratory of Automatic Target Recognition (Grant Number: WDZC20255290209), Hunan Provincial Research and Development Project (Grant No. 2025QK3019). This work was supported in part by the High Performance Computing Center of Central South University.

References

Bowman, F. 2012. *Introduction to Bessel functions*. Courier Corporation.

Chen, B.; Li, G.; Wu, R.; Zhang, X.; Chen, J.; Zhang, J.; and Zhang, L. 2025a. Adversarial diffusion compression for real-world image super-resolution. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 28208–28220.

Chen, B.; Zhang, Z.; Li, W.; Zhao, C.; Yu, J.; Zhao, S.; Chen, J.; and Zhang, J. 2025b. Invertible diffusion models for compressed sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(5): 3992–4006.

Chihaoui, H.; Lemkhenter, A.; and Favaro, P. 2024. Blind image restoration via fast diffusion inversion. *arXiv preprint arXiv:2405.19572*.

Chung, H.; Kim, J.; Mccann, M. T.; Klasky, M. L.; and Ye, J. C. 2022. Diffusion posterior sampling for general noisy inverse problems. *arXiv preprint arXiv:2209.14687*.

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

Guo, Y.; Dai, X.; Wang, S.; Jin, G.; and Zhang, X. 2024. Attention-based progressive discrimination generative adversarial networks for polarimetric image demosaicing. *IEEE Transactions on Computational Imaging*, 10: 713–725.

Hu, E. J.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W.; et al. 2022. Lora: Low-rank adaptation of large language models. *ICLR*, 1(2): 3.

Hwang, I.; Choi, K.; Ha, H.; and Kim, M. H. 2025. Benchmarking Burst Super-Resolution for Polarization Images: Noise Dataset and Analysis. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 24899–24909.

Kawar, B.; Elad, M.; Ermon, S.; and Song, J. 2022. Denoising diffusion restoration models. *Advances in Neural Information Processing Systems*, 35: 23593–23606.

Li, C.; Luo, Y.; Wu, C.; Zhang, J.; Yang, D.; and Zhao, D. 2025. Demosaicking customized diffusion model for snapshot polarization imaging. *Optics & Laser Technology*, 188: 112868.

Li, N.; Zhao, Y.; Pan, Q.; and Kong, S. G. 2019. Demosaicking DoFP images using Newton's polynomial interpolation and polarization difference model. *Optics Express*, 27(2): 1376–1391.

Luo, Y.; Zhang, J.; and Li, C. 2025. CPIFuse: Toward realistic color and enhanced textures in color polarization image fusion. *Information Fusion*, 120: 103111.

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

Luo, Y.; Zhang, J.; and Tian, D. 2023. Sparse representation-based demosaicking method for joint chromatic and polarimetric imagery. *Optics and Lasers in Engineering*, 164: 107526.

Maeda, R. 2019. Polanalyser: Polarization Image Analysis Tool.

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

Nguyen, V.; Tanaka, M.; Monno, Y.; and Okutomi, M. 2022. Two-step color-polarization demosaicking network. In *2022 IEEE International Conference on Image Processing (ICIP)*, 1011–1015. IEEE.

Qiu, S.; Fu, Q.; Wang, C.; and Heidrich, W. 2021. Linear polarization demosaicking for monochrome and colour polarization focal plane arrays. In *Computer Graphics Forum*, volume 40, 77–89. Wiley Online Library.

Rebhan, D.; Rosenberger, M.; and Notni, G. 2019. Principle investigations on polarization image sensors. In *Photonics and Education in Measurement Science 2019*, volume 11144, 50–54. SPIE.

Sun, L.; Wu, R.; Liang, J.; Zhang, Z.; Yong, H.; and Zhang, L. 2023. Improving the Stability and Efficiency of Diffusion Models for Content Consistent Super-Resolution. *arXiv preprint arXiv:2401.00877*.

Sun, L.; Wu, R.; Ma, Z.; Liu, S.; Yi, Q.; and Zhang, L. 2025. Pixel-level and semantic-level adjustable super-resolution: A dual-lora approach. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 2333–2343.

Sun, Y.; Zhang, J.; and Liang, R. 2021. Color polarization demosaicking by a convolutional neural network. *Optics Letters*, 46(17): 4338–4341.

Wang, Y.; Yu, J.; and Zhang, J. 2022. Zero-shot image restoration using denoising diffusion null-space model. *arXiv preprint arXiv:2212.00490*.

Wen, S.; Zheng, Y.; and Lu, F. 2021. A sparse representation based joint demosaicing method for single-chip polarized color sensor. *IEEE Transactions on Image Processing*, 30: 4171–4182.

Wen, S.; Zheng, Y.; Lu, F.; and Zhao, Q. 2019. Convolutional demosaicing network for joint chromatic and polarimetric imagery. *Optics Letters*, 44(22): 5646–5649.

Wu, B.; Peng, Y.; Hu, R.; and Zhou, X. 2025. Glossy Object Reconstruction with Cost-effective Polarized Acquisition. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 422–431.

Wu, R.; Sun, L.; Ma, Z.; and Zhang, L. 2024a. One-step effective diffusion network for real-world image super-resolution. *Advances in Neural Information Processing Systems*, 37: 92529–92553.

Wu, R.; Yang, T.; Sun, L.; Zhang, Z.; Li, S.; and Zhang, L. 2024b. Seesr: Towards semantics-aware real-world image super-resolution. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 25456–25467.

Wu, R.; Zhao, Y.; Li, N.; and Kong, S. G. 2021. Polarization image demosaicking using polarization channel difference prior. *Optics Express*, 29(14): 22066–22079.

Yao, M.; Wang, M.; Tam, K.-M.; Li, L.; Xue, T.; and Gu, J. 2025. PolarFree: Polarization-based Reflection-Free Imaging. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 10890–10899.

Yue, Z.; Liao, K.; and Loy, C. C. 2025. Arbitrary-steps image super-resolution via diffusion inversion. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 23153–23163.

Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; and Yang, M.-H. 2022. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 5728–5739.

Zhang, J.; Shao, J.; Luo, H.; Zhang, X.; Hui, B.; Chang, Z.; and Liang, R. 2018. Learning a convolutional demosaicing network for microgrid polarimeter imagery. *Optics Letters*, 43(18): 4534–4537.

Zhou, S.; Zhou, C.; Lyu, Y.; Guo, H.; Ma, Z.; Shi, B.; and Sato, I. 2025. PIDSR: Complementary Polarized Image Demosaicing and Super-Resolution. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 16081–16090.