

基于无监督学习的偏振图像去噪方法

胡浩丰^{1,2,3}, 金慧烽^{1,2}, 李校博^{3*}, 翟京生³, 刘铁根^{1,2}¹天津大学精密仪器与光电子工程学院, 天津 300072;²天津大学光电信息技术教育部重点实验室, 天津 300072;³天津大学海洋科学与技术学院, 天津 300072

摘要 提出一种基于无监督学习的偏振图像去噪方法,该方法打破了监督学习下深度学习需要严格配对图像的限制,使用非配对偏振图像训练一个偏振特化的循环生成对抗网络。该网络通过所提出的基于偏振信息的损失函数统计学习噪声图像和清晰图像的映射。实验结果表明,该网络可以有效地抑制室内外不同环境下偏振图像的噪声,同时能较好地恢复线性偏振度和偏振角。所提方法对复杂噪声环境下的偏振成像应用具有重要意义。

关键词 图像处理; 偏振成像; 偏振图像去噪; 无监督学习; 生成对抗网络

中图分类号 O436.3

文献标志码 A

DOI: 10.3788/AOS221645

1 引言

偏振作为光波的基本物理特性之一,可以表征光波和目标物体的物理特性。通过偏振成像获得的偏振参数被广泛地应用于工程实践中,包括生物成像^[1]、目标检测^[2-3]和 水下图像恢复^[4-6]等领域。特别是在光照不足等复杂环境中,非线性操作使得偏振度(DoP)和偏振角(AoP)等偏振参数更易受到噪声的影响。到目前为止,已经有许多方法被提出并用于偏振图像去噪,例如主成分分析(PCA)、 K 次奇异值分解(K -SVD)、块匹配三维滤波(BM3D)^[7-10]等。PCA方法利用降维技术获取图像特征实现去噪,但该算法会丢失图像细节信息。 K -SVD方法使用稀疏表示获取图像的特征信息,从而过滤噪声信息,但该算法将损失一部分纹理信息。BM3D方法利用分块和协同滤波实现了图像去噪,可以抑制图像的大部分噪声,但针对噪声敏感的偏振参数图像的去噪效果不理想。

相比于传统的仅基于物理模型的方法,基于深度学习技术的方法因其数据驱动特性在偏振图像去噪领域取得了巨大的成功。2020年, Li等^[11]首次将深度学习技术应用到偏振图像去噪,提出了一种基于监督学习的偏振去噪残差密集网络(PDRDN),该方法对于偏振噪声图像和对应的偏振参数图,均获得了接近标签图像的效果,特别是对于噪声敏感的DoP和AoP。相比于传统方法,该方法的去噪效果明显提升。然而,用于监督学习的严格配对的偏振图像获取成本高,且

在某些环境中难以获得,例如,对于夜间捕获的室外噪声图像,无法获取其严格配对的清晰图像。此外,监督学习网络的泛化性不强,在实际应用中效果易出现退化。因此使用非配对图像进行无监督学习的偏振图像去噪技术,切合偏振成像技术实际的应用需求,具有重要意义。近来,实现图像到图像转换的无监督学习方法得到了极大的发展,如CycleGan^[12]和WESPE^[13]。在CycleGan的工作中,网络使用未配对的图像数据实现域迁移,即图像在两个风格间相互转换,而噪声图像与清晰图像的转换与此类似,因此CycleGan被广泛应用于卫星图像去噪、CT图像去噪和地震记录图像去噪^[14-16]等配对图像难以获取的领域。同样地,配对的偏振图像数据集的获取存在难度大、成本高的问题,因此使用CycleGan实现无监督学习下的偏振图像去噪具有重要意义。

本文提出一种基于CycleGan的用于偏振图像去噪的网络结构。考虑到图像去噪主要关注网络的去噪性能,即重点关注单向的转换,本文删除了针对噪声图像的判别器,同时添加了两个针对偏振参数的判别器。此外,通过引入线性DoP(DoLP)和AoP的梯度损失来恢复偏振信息的变化趋势。因为分焦平面(DoFP)偏振相机在系统体积和成像速度方面具有明显优势,本文以DoFP图像为例来验证所提方法的去噪能力。特别地,本文构建了一个由室内偏振噪声图像和无噪图像组成的室内数据集和一个由夜间捕获的室外噪声图像组成的室外数据集。对于室内图像,所提方法能

收稿日期: 2022-08-26; 修回日期: 2022-09-17; 录用日期: 2022-09-22; 网络首发日期: 2022-09-30

基金项目: 广西创新驱动发展专项(桂科AA21077008)

通信作者: *lixiaobo@tju.edu.cn

够有效去除光强和偏振参数的噪声;而对于室外图像,所提方法可以在监督学习方法效果退化的情况下,有效抑制图像噪声,同时恢复偏振信息。实验结果表明,相比现有方法,所提方法具有更好的去噪效果和较强的泛化性。

2 基本原理

2.1 偏振图像噪声模型

在偏振成像系统中,理想环境下获得的清晰图像可以使用测量矩阵和斯托克斯矢量的乘积表示,但在真实环境下,偏振图像的获取容易受到多种噪声的影响,如散粒噪声、暗电流噪声、短噪声和热噪声等^[17],因此偏振噪声图像在 (x, y) 处的光强^[18]可以表示为

$$I_n(x, y) = f_n[\mathbf{W}\mathbf{S}(x, y)], \quad (1)$$

式中: (x, y) 表示图像内的像素点; $I_n(x, y)$ 为对应 (x, y) 点处的含噪光强图; $f_n(\cdot)$ 为真实环境下清晰图像到噪声图像的降质映射; \mathbf{W} 为偏振成像系统对应的测量矩阵; $\mathbf{S}(x, y) = [S_0(x, y), S_1(x, y), S_2(x, y), S_3(x, y)]^T$ 为清晰图像对应的斯托克斯矢量图。

从式(1)可以看出,偏振图像去噪的目的是通过求得 $f_n(\cdot)$,以及利用已知的测量矩阵 \mathbf{W} ,计算得到清晰图像对应的斯托克斯矢量图,从而计算对应的偏振参数。以本文使用的DoFP相机为例,相机同时获得4个偏振通道(0° 、 45° 、 90° 、 135°)的光强信息,因而可以使用4个偏振通道计算斯托克斯矢量,因此计算偏振参数,即

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

$$d_{\text{DoLP}} = \frac{2\sqrt{[(I_{0^\circ} - I_{90^\circ})^2 + (I_{45^\circ} - I_{135^\circ})^2]}}{(I_{0^\circ} + I_{45^\circ} + I_{90^\circ} + I_{135^\circ})^2}, \quad (3)$$

$$\theta_{\text{AoP}} = \frac{1}{2} \arctan\left[\frac{(I_{45^\circ} - I_{135^\circ})}{(I_{0^\circ} - I_{90^\circ})}\right], \quad (4)$$

式中: I_{0° 、 I_{45° 、 I_{90° 、 I_{135° 为4个偏振通道对应的光强图。

对于偏振噪声图像,4个偏振通道均受到噪声的影响,从式(3)和式(4)可以看出,偏振参数DoLP和AoP与4个偏振通道都相关,而非线性操作放大了4个偏振通道的噪声,导致DoLP和AoP对于噪声更为敏感,且噪声的分布更为复杂。如果去噪模型仅仅抑制了光强图噪声,偏振信息可能依旧存在严重的信息被噪声淹没甚至扭曲的现象。因此,理想的偏振图像去噪方法需要同时去除光强和偏振信息的噪声。

2.2 无监督偏振图像去噪模型

由于受到不同类型噪声的影响, $f_n(\cdot)$ 变得复杂,难以通过计算求得。深度学习等数据驱动算法是一种十分有效的拟合方法。受CycleGan的启发,本文提出一种基于无监督生成对抗网络的偏振图像去噪算法,使用非配对数据在统计上学习噪声图像和清晰图像之间的转换,而不是一对一的关联,从而克服监督学习的过拟合问题。

图1展示了无监督偏振去噪网络的框架。整个网络由两个生成器(G_D 、 G_N)和三个判别器(D_I 、 D_{DoLP} 和 D_{AoP})组成。当输入噪声图像,生成器 G_D 将生成清晰图像。为了防止生成器 G_D 将任何噪声图像映射为相同的输出图像,并保持输入和输出图像之间的结构一致性,同时为在非配对训练中引入偏振梯度损失,本文使用生成器 G_N 将生成的清晰图像转换为噪声图像。3个判别器分别接收4个通道的偏振图像、DoLP图和AoP

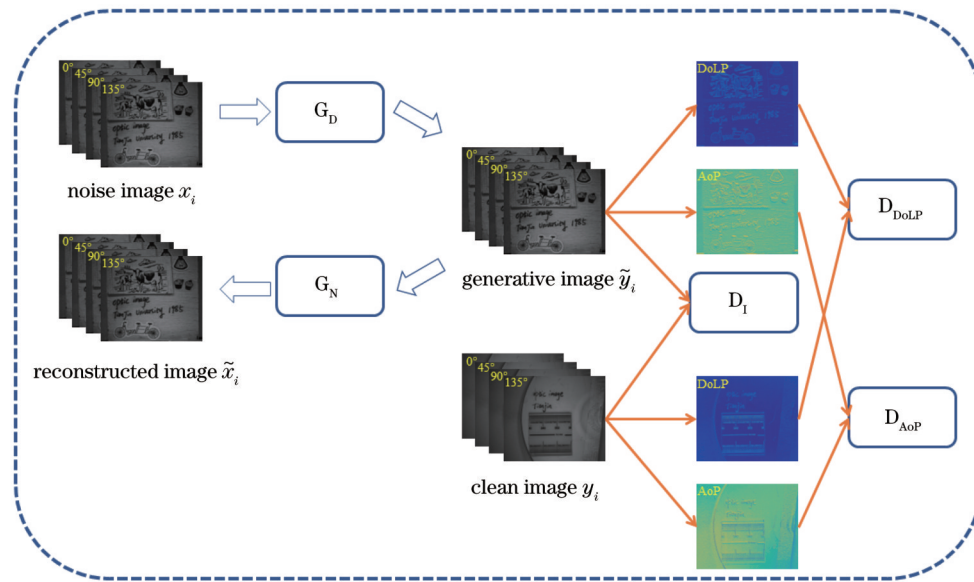


图1 无监督偏振图像去噪生成对抗网络框架

Fig. 1 Framework of unsupervised polarimetric-image denoising generative adversarial network

图,并给出输入图像为真实清晰图像的概率。对于生成器 G_D ,训练的目标是尽可能地使用生成的清晰图像“骗过”判别器。同时,使用真实清晰图像和生成的清晰图像训练判别器区分生成图像和真实图像。特别地,该网络添加了额外的两个偏振参数判别器 D_{DoLP} 和 D_{AoP} ,以引导生成器 G_D 恢复偏振信息。

生成器 G_D 和 G_N 的结构如图 2(a) 所示,其中 Conv 表示卷积层,符号 k, n 和 s 后的数字分别表示卷积核大小、卷积核的数量以及卷积步长。在生成器中,引入残差密集模块(RDB),通过密集连接学习更多的局部特征^[19]。每个 RDB 由 7 个卷积层组成:前 6 层为 3×3 卷积层,主要用于特征提取;最后一层为 1×1 卷积层,作用为特征融合。密集连接和残差连接使得 RDB 中的每个卷积层都可以连接后续所有层,并传递需要保留

的信息^[20]。对于生成器 G_D 和 G_N ,网络的头部为 2 层卷积层,尾部为 6 层卷积层,中间则是 6 个 RDB。此外,在生成器每个卷积层之后都添加了批标准化层(BN),以加速网络训练^[21]。BN 层后使用线性整流函数(ReLU)作为非线性激活函数。

判别器 D_I 、 D_{DoLP} 和 D_{AoP} 的结构如图 2(b) 所示。采用 70×70 PatchGANs^[22] 作为判别器。它由 5 个卷积层组成,核大小均为 4×4 ,中间的 3 个卷积层后均有 BN 层。PatchGANs 的输出为一张特征图,特征图中的每个像素值表示判别为真的概率。这种判别器结构不限制输入图像的大小,并且在每个卷积层之后感受野都会增长,因此最后输出的特征图中的每个像素的感受野可以对应于输入图像的一个 70×70 图像块。

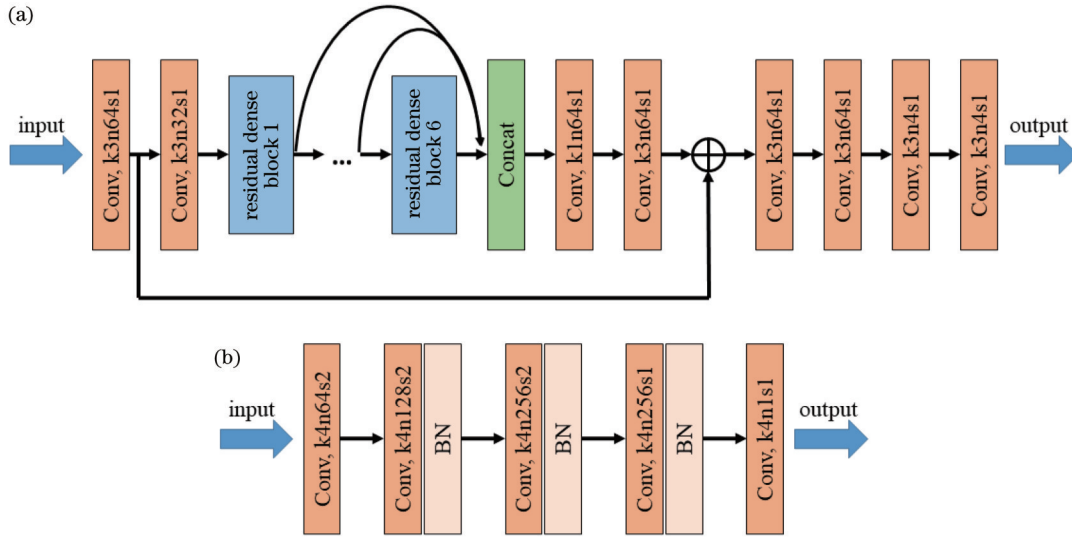


图 2 生成对抗网络结构图。(a)生成器网络结构;(b)判别器网络结构

Fig. 2 Structure of generative adversarial network. (a) Structure of generators; (b) structure of discriminators

2.3 损失函数

生成对抗损失在生成对抗网络中引导生成器与判别器的对抗,考虑到鲁棒性,采用最小二乘损失代替负对数似然损失^[23]。判别器的生成对抗损失可以描述为

$$\mathcal{L}_D^{\text{Noise}} = \frac{1}{2N} \sum_{i=1}^N \|D_I[G_D(X_i)]\|_2 + \frac{1}{2N} \sum_{i=1}^N \|D_I(Y_i) - 1\|_2, \quad (5)$$

$$\mathcal{L}_D^{\text{DoLP}} = \frac{1}{2N} \sum_{i=1}^N \|D_{DoLP}\{F_{DoLP}[G_D(X_i)]\}\|_2 + \frac{1}{2N} \sum_{i=1}^N \|D_{DoLP}[F_{DoLP}(Y_i)] - 1\|_2, \quad (6)$$

$$\mathcal{L}_D^{\text{AoP}} = \frac{1}{2N} \sum_{i=1}^N \|D_{AoP}\{F_{AoP}[G_D(X_i)]\}\|_2 + \frac{1}{2N} \sum_{i=1}^N \|D_{AoP}[F_{AoP}(Y_i)] - 1\|_2, \quad (7)$$

式中: X_i , Y_i 分别为一组非配对偏振噪声图像和偏振清晰图像; N 为一次训练输入的图像组数; $G_D(\bullet)$ 为生

成器 G_D 的输出图像; $F_{DoLP}(\bullet)$ 、 $F_{AoP}(\bullet)$ 分别为计算 DoLP 和 AoP 的函数; $D_I(\bullet)$ 、 $D_{DoLP}(\bullet)$ 、 $D_{AoP}(\bullet)$ 为 3 个判别器对图像的判别结果; $\|\bullet\|_2$ 表示 L2 范数。

判别器的训练目标是 minimized 损失函数,即尽可能地准确地区分生成图像和真实图像。相反地,对生成器 G_D 进行训练以欺骗 3 个判别器并最大化 3 个判别器的生成对抗损失。因此,生成器的生成对抗损失由下面 3 个损失组成:

$$\mathcal{L}_G^{\text{Noise}} = \frac{1}{N} \sum_{i=1}^N \|D_I[G_D(X_i)] - 1\|_2, \quad (8)$$

$$\mathcal{L}_G^{\text{DoLP}} = \frac{1}{N} \sum_{i=1}^N \|D_{DoLP}\{F_{DoLP}[G_D(X_i)]\} - 1\|_2, \quad (9)$$

$$\mathcal{L}_G^{\text{AoP}} = \frac{1}{N} \sum_{i=1}^N \|D_{AoP}\{F_{AoP}[G_D(X_i)]\} - 1\|_2. \quad (10)$$

3 个子损失的大小在不同的训练阶段是不同的,在训练的前期,3 种损失的大小相近,但在训练后期,因为 DoLP 和 AoP 图像对噪声具有更强的敏感性,因

此 D_{DoLP} 和 D_{AoP} 判别器对应的损失要大于 D_1 判别器对应的损失。为了平衡 3 个子损失的值,使用自适应权重。因此,总的生成对抗损失为

$$\omega_{\text{sub}} = \mathcal{L}_G^{\text{sub}} / (\mathcal{L}_G^{\text{Noise}} + \mathcal{L}_G^{\text{DoLP}} + \mathcal{L}_G^{\text{AoP}}), \quad (11)$$

$$\mathcal{L}_G = \sum_{\text{sub}} \omega_{\text{sub}} \mathcal{L}_G^{\text{sub}}. \quad (12)$$

为了防止生成器将任何输入图像映射到相同的输出分布,并保持输入和输出图像之间的一致性,添加了另一个生成器 G_N ,将 G_D 输出的清晰图像映射到噪声图像。使用循环一致性损失使 G_N 的输出噪声图像接近原输入噪声图像,即

$$\mathcal{L}_{\text{Grad}}^{\text{DoLP}} = \frac{1}{N} \sum_{i=1}^N \left\{ \left\| \nabla_h \langle F_{\text{DoLP}} \{ G_N [G_D (X_i)] \} \rangle - \nabla_h [F_{\text{DoLP}} (X_i)] \right\|_2 + \left\| \nabla_w \langle F_{\text{DoLP}} \{ G_N [G_D (X_i)] \} \rangle - \nabla_w [F_{\text{DoLP}} (X_i)] \right\|_2 \right\}, \quad (14)$$

$$\mathcal{L}_{\text{Grad}}^{\text{AoP}} = \frac{1}{N} \sum_{i=1}^N \left\{ \left\| \nabla_h \langle F_{\text{AoP}} \{ G_N [G_D (X_i)] \} \rangle - \nabla_h [F_{\text{AoP}} (X_i)] \right\|_2 + \left\| \nabla_w \langle F_{\text{AoP}} \{ G_N [G_D (X_i)] \} \rangle - \nabla_w [F_{\text{AoP}} (X_i)] \right\|_2 \right\}, \quad (15)$$

式中: $\nabla_h(\cdot)$ 、 $\nabla_w(\cdot)$ 分别为计算水平方向梯度和竖直方向梯度的函数。

在参考文献[12]中,身份损失用于避免在不需要时更改输入图像的色调,并且文献[12]的实验证明身份损失有助于保持输入图像的颜色。由于偏振图像由 4 个通道组成,对应于 4 个偏振方向,因此本文添加了一个身份损失以保留输入图像的偏振信息:

$$\mathcal{L}_{\text{idt}} = \frac{1}{N} \sum_{i=1}^N \left\| G_D (Y_i) - Y_i \right\|_2. \quad (16)$$

总之,生成器的最终损失是 5 种损失的加权和,可以描述为

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_G + \lambda_2 \mathcal{L}_{\text{cyc}} + \lambda_3 \mathcal{L}_{\text{Grad}}^{\text{DoLP}} + \lambda_4 \mathcal{L}_{\text{Grad}}^{\text{AoP}} + \lambda_5 \mathcal{L}_{\text{idt}}, \quad (17)$$

式中: λ_1 、 λ_2 、 λ_3 、 λ_4 、 λ_5 为各个损失的权重。

生成对抗损失、循环损失以及身份损失的权重与 CycleGan 设置^[12]相同,即 $\lambda_1 = 1$ 、 $\lambda_2 = 10$ 、 $\lambda_5 = 5$ 。对于两个梯度损失,考虑到两个偏振参数对于噪声的敏感性,将其对应权重调整设置为 $\lambda_3 = 0.05$ 、 $\lambda_4 = 0.005$ 。值得注意的是,当梯度损失权重过大时,会出现网络过于偏重偏振参数,而出现光强图严重失真问题,因此将梯度损失的权重设为一个较小值。同时考虑到实际中 AoP 相较于 DoLP 对于噪声更为敏感,将 AoP 对应的权重设为 DoLP 对应权重的 1/10。

3 实验结果与讨论

3.1 数据集构建

本文中的数据集包含室内图像和室外图像。对于室内图像,实验设置与参考文献[9]相同。对于室内偏振图像,噪声图像是在高增益和短曝光时间条件下得到,而清晰图像则是在零增益和长曝光时间的条件下,同一场景下连续拍摄 50 幅图像并取平均值得到。由

$$\mathcal{L}_{\text{cyc}} = \frac{1}{N} \sum_{i=1}^N \left\| G_N [G_D (X_i)] - X_i \right\|_2, \quad (13)$$

式中: $G_N(\cdot)$ 为生成器 G_N 的输出图像。

此外,在 DoLP 和 AoP 图像中,不同物体之间强度的变化更能反映物体的偏振特性,而相比于强度损失,梯度损失更注重图像细节的变化以及偏振信息的变化,因此为保证偏振信息的真实性,本文提出了 DoLP 和 AoP 的梯度损失。梯度损失引导网络恢复 DoLP 和 AoP 的变化趋势,以此更好恢复偏振信息,两个梯度损失可以描述为

于无法在夜间获得成对的噪声-清晰图像,只能在夜间拍摄固定高增益和长曝光时间下的噪声图像。

DoFP 偏振相机具有紧凑性高、稳定性好以及时间同步的特性^[24],因此本实验使用商用 DoFP 偏振相机 (LUCID, PHX050S-PC) 在自然光下拍摄实验图像。对于室内图像,以相同的增益和曝光时间拍摄了 110 组偏振图像,空间分辨率为 2448 pixel × 2048 pixel。其中 70 组图像被划分到训练集,其余 40 组被划分到验证集和测试集。此外,对于室外图像,45 幅噪声图像作为测试集的补充,以评估所提出方法的泛化性。训练时,每一次训练输入的噪声图像和清晰图像完全不配对。此外,使用 64 pixel × 64 pixel 的窗口,并以 32 pixel 的步长将每一幅图像切割为多幅小图,以此增加数据集的规模。

3.2 网络训练参数设置

利用峰值信噪比 (PSNR) 和结构相似性 (SSIM)^[25] 来评估所提方法的性能。PSNR 表征信号的最大可能值与使其失真的噪声功率之间的比率。PSNR 越高,生成的图像越接近标签图像 (GT)。SSIM 更关注结构信息,这可以更好地评估偏振参数的恢复情况。类似地,SSIM 越高,生成的图像与 GT 越相似。

训练的批大小 (batch size) 设置为 36,在训练前对每组未配对图像进行了随机水平翻转、随机垂直翻转以及随机顺时针旋转 90° 的操作,以增强训练效果、提高模型的泛化能力。本文使用 Nvidia RTX 3090 GPU 实现 Pytorch 框架。卷积层的所有权重均使用正态分布随机初始化,标准差设为 0.01,同时使用 Adam 优化器更新网络参数。所有层的学习率初始化为 10^{-4} ,网络一共训练 90 个周期,如果验证集的平均 PSNR 在 5 个周期内没有降低,学习率将降低到 1/5。

3.3 消融实验

为验证所提出的两种偏振判别器的有效性,本文进行消融实验,比较了不同判别器对于模型恢复噪声图像效果的影响。在消融实验中分别训练了使用光强判别器、使用光强和 DoLP 判别器、使用光强和 AoP 判别器以及使用 3 种判别器的 4 种网络,所有网络的损失函数和超参数设置保持一致。图 3 展示了不同结构网络的比较结果,其中 S_0 为光强。

从图 3 第二列可以看出,当没有额外偏振判别器

时,网络对于 DoLP 和 AoP 图像的恢复效果较差,与 GT 的差距较大。当网络加入 D_{DoLP} 后,网络对于 DoLP 图像的恢复效果有所提升,但 AoP 图像依旧存在背景亮度过低的问题。当网络加入 D_{AoP} 后,DoLP 恢复效果较好,但 AoP 图像文字部分亮度过高。仅有当两个偏振判别器都加入时,网络才能很好地恢复 DoLP 和 AoP 图像。图 3 第四列图像的细节信息被很好地恢复,纹理和背景处的亮度也与 GT 接近,说明在两个偏振判别器的帮助下,网络可以更好地恢复偏振信息。



图 3 不同网络结构对噪声图像的处理结果

Fig. 3 Processing results of noisy images with different network structures

消融实验中各个网络对测试集处理的平均 PSNR 和 SSIM 值如表 1 所示。表 1 第 3 列 Net-1 在没有任何偏振判别器的情况下,光强图像和 DoLP 图像的两项指标相比其他网络出现了大幅度下降;表 1 第 4 列的 Net-2 加入了 DoLP 判别器,可以看到光强图像和 DoLP 图像的指标数值均有上升,光强图像的 PSNR 提升到 26 dB,DoLP 图像的 PSNR 提高了 5 dB,但

AoP 图像的 PSNR 下降了 0.5 dB 左右;Net-3 在 AoP 判别器的帮助下,光强、DoLP 和 AoP 图像指标数值均得到了提升,3 种图像的指标数值均达到了较高的数值;最后一列的网络使用了所有的偏振判别器,相比于仅使用 AoP 判别器的 Net-3,Net-4 在 DoLP 判别器的帮助下,DoLP 和 AoP 图像的 PSNR 均提升了 1 dB 左右,SSIM 均提升了 0.2 左右。

表 1 不同网络结构去噪效果的光强、DoLP 和 AoP 图像指标的比较

Table 1 Indicator comparison of intensity, DoLP, and AoP images for different network structures

Network		Net-1	Net-2	Net-3	Net-4
Intensity image	PSNR /dB	23.7057	26.3142	31.2859	31.3139
	SSIM	0.7104	0.8874	0.8970	0.8877
DoLP image	PSNR /dB	19.2122	24.5586	26.2351	27.6377
	SSIM	0.5517	0.6536	0.6631	0.6849
AoP image	PSNR /dB	13.3240	12.7801	13.9141	15.4566
	SSIM	0.1912	0.1912	0.1979	0.2123

3.4 室内图像的处理结果

对于室内图像,拍摄了配对的偏振噪声图像和偏振清晰图像,因此可以对该方法进行定量评估。将所提方法与 BM3D^[10] 和 PDRDN^[11] 这两个有代表性的方

法进行比较。BM3D 是一种传统的去噪算法,它可以同时利用所有不同的偏振通道,并很好地保留图像细节。PDRDN 是一种基于监督学习的数据驱动算法,可以较好地恢复偏振参数。3 种方法的去噪结果如

图 4 和表 2 所示。本文中 BM3D 的参考块的步长设为 3, 参考块的大小设为 3×3 。PDRDN 的学习率设为

10^{-4} , 使用 MSRA 方法初始化权重, 训练周期设为 60, 使用 Adam 优化器进行训练, 批大小设为 32。

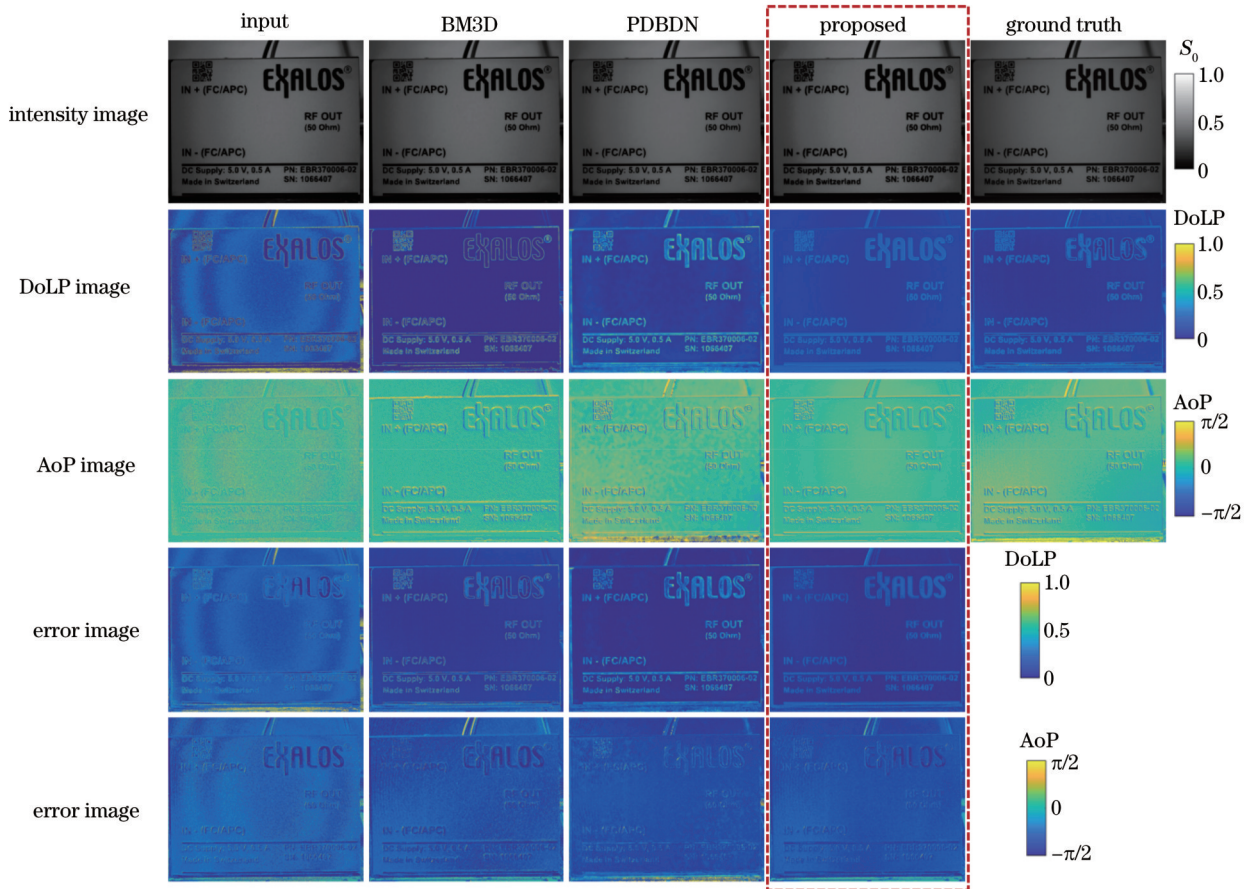


图 4 不同方法对室内噪声图像的处理结果

Fig. 4 Processing results of indoor noise images by different methods

从图 4 可以看出:BM3D 方法处理结果的 DoLP 图像纹理部分的亮度过高,但背景部分的亮度过低,因此整体风格与 GT 差距较大,而 AoP 图像细节基本恢复,但整体亮度偏低;监督学习的 PDRDN 可以较好地恢复偏振信息,DoLP 和 AoP 图像的细节均得到了较好的恢复,但是 DoLP 图像依旧会出现亮度偏高的现象,且 AoP 图像依旧有部分噪声没有去除;所提的无监督学习方法较好地恢复了偏振信息,且 DoLP 图像细节处的亮度与 GT 基本一致,AoP 图像较为平滑,噪声基本被去除。从 DoLP 以及 AoP 的误差图像(error image)可以看出,无监督学习方法的平均亮度较低,特别是对于 DoLP 误差图像,PDRDN 方法和 BM3D 方法对应的误

差图像平均亮度较大,而无监督方法对应的平均亮度较小,表明无监督方法处理结果与 GT 更为接近。

表 2 所示为测试集处理结果的 SSIM 和 PSNR 值的比较。从表 2 可以看出,传统的 BM3D 方法在恢复光强图像时有一定的优势,但在恢复偏振信息方面有着不足。特别是对于 DoLP 图像的恢复,虽然 BM3D 方法可以获得较高的 PSNR 值,但去噪结果的 SSIM 显著低于数据驱动方法。有监督和无监督的深度学习方法在两个指标上均取得了较好的结果,其中无监督方法在处理 DoLP 图像时,可以在 SSIM 接近的情况下,获得更高的 PSNR 值。

上述测试的图像与训练用的数据在同一条件下拍

表 2 不同网络方法下光强、DoLP 和 AoP 图像指标比较

Table 2 Indicator comparison of intensity, DoLP, and AoP images for different methods

Method	Intensity image		DoLP image		AoP image	
	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB
BM3D	0.8863	33.3374	0.2940	23.4488	0.1942	13.6127
PDRDN	0.8928	31.4548	0.6843	25.2243	0.2049	15.8199
Proposed	0.8877	31.3139	0.6849	27.6377	0.2123	15.4566

摄,因此监督学习和无监督学习均取得了较好的效果,尤其是对于偏振参数的恢复,数据驱动方法显著优于传统方法,其中无监督学习方法可以更好地恢复 DoLP 图像,很好地保持了图像的细节和亮度信息,优于现有的图像去噪方法。

此外,为验证无监督去噪方法对不同材料噪声图像的处理效果,本文对几种常见的材料进行了测试,包括树脂、布、木材和塑料。图 5 所示为无监督方法对不同材料噪声图像的去噪效果。在木材对应的 DoLP 图像中,输入的噪声图像和对应的 GT 部分区域 DoLP 差距较大,这是因为偏振噪声图像受到噪声影响,从式 (3) 可以看出,DoLP 高的区域是由垂直偏振方向的光强差异导致的,而噪声以及平方的操作放大了这一差异,使得噪声图像与 GT 对应的 DoLP 出现差异,但是对于 DoLP 的趋势,噪声图像和偏振图像是相同的。从图 5 可以看出,所提方法可以较好地抑制不同材料

图像的噪声。对于 AoP 图像,噪声图像完全被噪声淹没,例如塑料尺上的刻度、乳液包装上的文字和玩偶兔子的轮廓。而无监督学习可以较好地恢复 AoP 图像,图像的轮廓可以清晰地被辨认,这进一步说明了无监督方法的有效性和泛化性。从图 5 的第 2~4 行的 AoP 去噪图像和 AoP 图像的真值图像的对比可看出,真值图像的噪声水平反而比去噪图像更高,这主要是因为真值图像是在高曝光度条件下连续拍摄 50 帧图像并取平均值得到的,因此可以很好地抑制零均值的高斯白噪声(AWGN),但无法完全抑制脉冲噪声等非零均值噪声。高曝光度条件下图像的噪声很小,只有在对噪声十分敏感的 AoP 图像中才能观察到。类似地,PDRDN^[11]也出现了这种现象,而本文使用深度学习方法同样基于数据驱动,使用的网络注重于提取图像的特征,因此可以获得噪声水平更低的去噪图像。

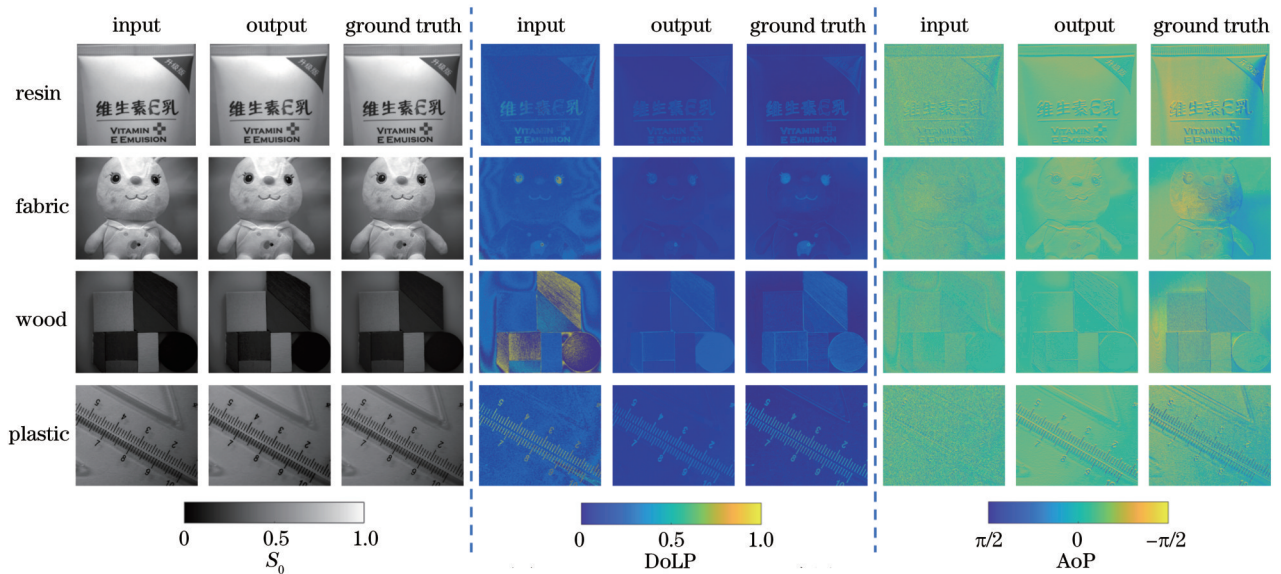


图 5 4 种不同材料的去噪结果

Fig. 5 Denoising results of four different materials

为进一步说明本文方法对不同场景的适应性,对不同材料的物体的去噪结果进行了定量的分析,结果如表 3 所示。可以看到,利用本文方法去噪后不同材料的偏振图像均可以取得较高的 PSNR 值和 SSIM

值,虽然不同材料偏振图像的各项指标有所波动,但各项指标均在表 2 所示的均值左右。总体来说,本文方法对于不同材料的物体的去噪效果有一定的差异,但都能得到较好的去噪效果。

表 3 不同材料的光强、DoLP 和 AoP 图像指标比较

Table 3 Indicator comparison of intensity, DoLP, and AoP images for different materials

Material	Intensity image		DoLP image		AoP image	
	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB
Resin	0.9007	32.5596	0.6703	27.3557	0.2071	15.7295
Fabric	0.9207	33.4452	0.7776	30.1627	0.2462	15.8806
Wood	0.8755	31.0895	0.6748	30.2264	0.2319	16.6111
Plastic	0.9181	32.8964	0.7755	27.9045	0.2409	16.4649

3.5 室外图像的处理结果

对于夜间暗光下拍摄的室外噪声图像,无法获取相应的清晰图像,因此只能定性地检验所提方法的去噪效果。图 6 展示了不同方法处理室外偏振噪声图像的效果。虽然受到噪声的干扰,但是噪声图像对应的 DoLP 的变化趋势仍然可以反映清晰图像偏振度的变化。因此可以通过观察 DoLP 的变化趋势评估不同方法的去噪效果。从图 6 可以看出,对于 DoLP 图像,3 种方法均可以将汽车的轮廓复原,但从噪声 DoLP 图像可以看出,车窗部分的偏振度高于汽车前盖。对于

BM3D 方法,DoLP 和 AoP 图像仅有轮廓得到了较好的恢复,汽车整体的偏振度基本一致,没有随着材质的变化而变化。监督学习的 PDRDN 方法基本恢复了图像的轮廓,但 DoLP 图像只有部分车窗区域的偏振度较高,且车灯区域的高偏振度没有得到很好还原。此外,PDRDN 处理后的 AoP 图像出现了非常严重的扭曲。从图 6 最后一列图像可以看到,无监督学习方法在去除噪声的同时保持了偏振度的变化趋势,较好地恢复了偏振信息。

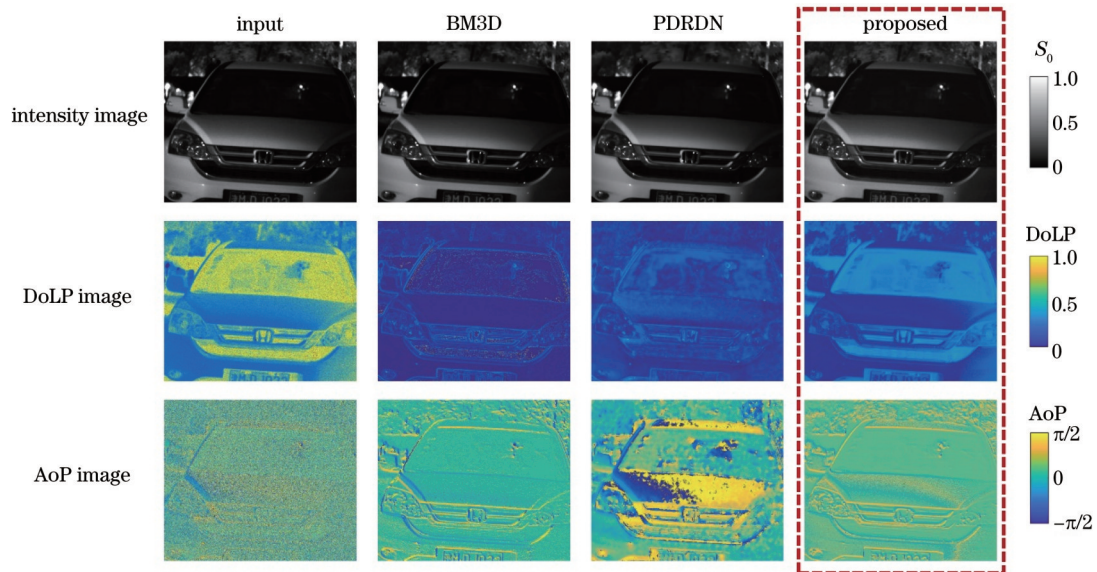


图 6 不同方法对室外噪声图像的处理结果

Fig. 6 Processing results of outdoor noise images by different methods

从图 6 可以看出,监督学习方法应用于不同环境时,去噪效果易出现较大的退化,特别是对于偏振信息的恢复。传统的 BM3D 方法在恢复不同环境时的表现较为稳定,但与处理室内图像时一样,BM3D 无法较好地恢复偏振信息。无监督方法通过学习大量的数据,取得了较好的去噪效果,同时非配对的训练方式又使得无监督方法在统计上学习噪声图像和清晰图像之间的转换,因此应用于不同环境时,无监督方法有着较好的稳定性。

4 结 论

提出一种基于无监督学习的偏振图像去噪方法,在基础 CycleGan 模型上,设计了适用于偏振图像去噪的生成对抗网络。通过非配对数据无监督的训练网络,得到一个能够有效去除偏振图像噪声并恢复偏振信息的去噪网络模型。针对室内图像测试了所提方法,并同时给出了定性和定量的评价。实验结果表明,所提方法在室内图像去噪方面可以达到与监督学习相似的效果,可以有效地恢复偏振信息,特别是在 DoLP 图像恢复上有着一定的优势。进一步地,本文亦针对

不同材料的偏振图像进行了测试。结果均表明,所提方法有着较好的泛化性,可以有效恢复不同材料的偏振信息。此外,还对室外图像进行了测试,并给出了定性的评价。实验结果表明,所提方法可以在训练集为室内图像的情况下,有效地去除室内外图像的噪声并恢复真实的偏振信息。所提模型和方法可以延拓到其他应用,例如可以用来研究极端环境(例如夜晚、微光等)下的偏振图像去噪效果以及偏振信息的恢复情况。可以看到,所提方法虽然恢复了输入偏振噪声图像偏振信息的变化趋势,但偏振信息与标签图像依旧有着差距,同时标签图像的获取方法仅能完全抑制零均值噪声,而无法完全实现无噪的标签图像获取,因此去噪后如何获得更精确的偏振信息以及如何获得完全无噪的标签图像是我们接下来的重要研究方向。

参 考 文 献

- [1] Li D Z, He H H, Zeng N, et al. Polarization imaging and scattering model of cancerous liver tissues[J]. Journal of Innovative Optical Health Sciences, 2013, 6(3): 1350025.
- [2] Garcia-Caurel E, Ossikovski R, Foldyna M, et al. Advanced Mueller ellipsometry instrumentation and data analysis[M]// Losurdo M, Hingerl K. Ellipsometry at the nanoscale.

- Heidelberg: Springer, 2013: 31-143.
- [3] Liu Z D, Liao R, Ma H, et al. Classification of marine microalgae using low-resolution Mueller matrix images and convolutional neural network[J]. *Applied Optics*, 2020, 59(31): 9698-9709.
- [4] Xu Q, Guo Z Y, Tao Q Q, et al. Transmitting characteristics of polarization information under seawater[J]. *Applied Optics*, 2015, 54(21): 6584-6588.
- [5] 杨力铭, 梁健, 张文飞, 等. 基于非偏振光照明的水下偏振成像目标增强技术[J]. *光学学报*, 2018, 38(6): 0611003.
Yang L M, Liang J, Zhang W F, et al. Underwater polarimetric imaging target enhancement technology based on unpolarized illumination[J]. *Acta Optica Sinica*, 2018, 38(6): 0611003.
- [6] 汪杰君, 梁磊, 李树, 等. 水下目标偏振差分成像模型修正与实现[J]. *光学学报*, 2019, 39(11): 1111003.
Wang J J, Liang L, Li S, et al. Correction and implementation of polarization-difference imaging model for underwater target[J]. *Acta Optica Sinica*, 2019, 39(11): 1111003.
- [7] Zhang J C, Luo H B, Liang R G, et al. PCA-based denoising method for division of focal plane polarimeters[J]. *Optics Express*, 2017, 25(3): 2391-2400.
- [8] 尹佳琪, 王世勇, 李范鸣. 基于改进主成分分析的分焦平面偏振图像去噪算法[J]. *光学学报*, 2021, 41(7): 0710002.
Yin J Q, Wang S Y, Li F M. Division-of-focal-plane polarization image denoising algorithm based on improved principal component analysis[J]. *Acta Optica Sinica*, 2021, 41(7): 0710002.
- [9] Ye W B, Li S T, Zhao X J, et al. A K times singular value decomposition based image denoising algorithm for DoFP polarization image sensors with Gaussian noise[J]. *IEEE Sensors Journal*, 2018, 18(15): 6138-6144.
- [10] Abubakar A, Zhao X J, Li S T, et al. A block-matching and 3-D filtering algorithm for Gaussian noise in DoFP polarization images[J]. *IEEE Sensors Journal*, 2018, 18(18): 7429-7435.
- [11] Li X B, Li H Y, Lin Y, et al. Learning-based denoising for polarimetric images[J]. *Optics Express*, 2020, 28(11): 16309-16321.
- [12] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2242-2251.
- [13] Ignatov A, Kobyshev N, Timofte R, et al. WESPE: weakly supervised photo enhancer for digital cameras[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 804-813.
- [14] Song J, Jeong J H, Park D S, et al. Unsupervised denoising for satellite imagery using wavelet directional CycleGAN[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(8): 6823-6839.
- [15] Gu J, Yang T S, Ye J C, et al. CycleGAN denoising of extreme low-dose cardiac CT using wavelet-assisted noise disentanglement[J]. *Medical Image Analysis*, 2021, 74: 102209.
- [16] Li W D, Wang J. Residual learning of cycle-GAN for seismic data denoising[J]. *IEEE Access*, 2021, 9: 11585-11597.
- [17] Guo S, Yan Z F, Zhang K, et al. Toward convolutional blind denoising of real photographs[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 16-21, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 1712-1722.
- [18] Hu H F, Lin Y, Li X B, et al. IPLNet: a neural network for intensity-polarization imaging in low light[J]. *Optics Letters*, 2020, 45(22): 6162-6165.
- [19] Zhang Y L, Tian Y P, Kong Y, et al. Residual dense network for image super-resolution[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-22, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 2472-2481.
- [20] Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 8-10, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 5197-5206.
- [21] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//ICML'15: Proceedings of the 32nd International Conference on International Conference on Machine Learning-Volume 37, July 6-11, 2015, Lille, France. New York: ACM Press, 2015: 448-456.
- [22] Li C, Wand M. Precomputed real-time texture synthesis with Markovian generative adversarial networks[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9907: 702-716.
- [23] Mao X D, Li Q, Xie H R, et al. Least Squares generative adversarial networks[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2813-2821.
- [24] Zhang J C, Luo H B, Hui B, et al. Image interpolation for division of focal plane polarimeters with intensity correlation[J]. *Optics Express*, 2016, 24(18): 20799-20807.
- [25] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.

Polarization Image Denoising Based on Unsupervised Learning

Hu Haofeng^{1,2,3}, Jin Huifeng^{1,2}, Li Xiaobo^{3*}, Zhai Jingsheng³, Liu Tiegeng^{1,2}

¹School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, Tianjin 300072, China;

²Key Laboratory of Optoelectronic Information Technology, Ministry of Education, Tianjin University, Tianjin 300072, China;

³School of Marine Science and Technology, Tianjin University, Tianjin 300072, China

Abstract

Objective The imaging process of polarization images in the natural environment is easily affected by noise, which not only causes the acquired relevant polarimetric parameters to deviate from their real values but also affects the further

processing of subsequent polarization information. Due to nonlinear operations, polarimetric parameters such as the degree of polarization (DoP) and the angle of polarization (AoP) are easily distorted by noise, especially in photon-starved environments. Therefore, effective denoising is crucial to polarimetric imaging. The denoising method based on deep learning can significantly remove the influence of noise on polarization images. However, the performance of current supervised algorithms is highly dependent on the labeled dataset, and high-quality polarization labels are difficult to obtain in practical applications, which limits the application of the existing methods. Therefore, this paper proposes a polarization image denoising method based on unsupervised learning. This method breaks the restriction that supervised learning-based deep learning requires strictly paired images and uses unpaired polarization images to train a polarization-specialized cycle generative adversarial network (CycleGan). The method in this paper are of great significance to the application of polarimetric imaging in complex noise environments.

Methods In the proposed CycleGan structure, the discriminator for the input domain is removed, and two discriminators for polarimetric parameters are added. In the structure of generators, the residual dense block (RDB) is introduced to extract abundant local features via densely connected convolutional layers, and PatchGANs are adopted for discriminators, which can work on arbitrarily sized images and grow the receptive field after each convolution layer. In addition, a batch normalization (BN) layer and a ReLU layer are added right after each convolutional layer to accelerate network training. Furthermore, a cycle consistency loss is maintained to keep the consistency between input and output, and two cycle gradient losses are introduced for the degree of linear polarization (DoLP) and AoP to preserve the variations of polarization information. With the help of the designed network structure and the polarization-based loss function, the network trained by unpaired polarization images can statistically learn the mapping between noisy and clean images.

Results and Discussions Experiments show that the network can effectively suppress the noise of polarization images in different indoor and outdoor environments and recover DoLP and AoP. The ablation experiment proves the effectiveness of additional polarization discriminators. With two discriminators, the network accurately recovers both DoLP and AoP images (Fig. 3) and achieves the highest PSNR/SSIM value among different network structures (Table 1). Compared with other methods, the unsupervised method has the best performance in terms of intensity, DoLP, and AoP images (Fig. 4). The average PSNR and SSIM of indoor images illustrate that the method has advantages in the reconstruction of DoLP images (Table 2). Several groups of experiments on different materials, including resin, fabric, wood, and plastic, are conducted to verify the universality of the proposed method. The denoised results reveal that the proposed method can suppress the noise of these materials for polarization information (Fig. 5). Finally, experiments with outdoor noise polarization images are carried out to verify the robustness of the method. Compared to the supervised method, the unsupervised method does not see dramatical performance degradation when applied to different environments (Fig. 6), which is important for the application of polarization imaging in realistic environments.

Conclusions This paper proposes a polarization image denoising method based on unsupervised learning. On the basis of the CycleGan model, a structure of generative adversarial network suitable for polarization image denoising is designed. Through an unsupervised training network with unpaired images, a denoising network model that can effectively remove the noise of polarization images and restore polarization information is obtained. Experiments with indoor images are conducted to test the method, and qualitative and quantitative evaluations are given. The experimental results show that this method can achieve the same performance as the supervised learning method in indoor image denoising and can effectively restore polarization information, especially in DoLP image restoration. Furthermore, the polarization images of different materials are tested. The results reveal that this method has good generalization and can effectively recover the polarization information of different materials. In addition, the outdoor images are also tested, and a qualitative evaluation is presented. The experimental results suggest that this method can effectively remove the noise of indoor and outdoor images and restore real polarization information when indoor images are used as the training set. The models and methods proposed in this study can be extended to other applications. For example, they can be used to study polarization image denoising and polarization information recovery in extreme environments (e. g. , night, low light).

Key words image processing; polarimetric imaging; polarization image denoising; unsupervised learning; generative adversarial network