

采用上下文双自相关注意力的偏振图像超分辨率重建网络

陈广秋¹ 张赛¹ 刘枫铭¹ 贯成才¹ 刘国峰¹ 段锦¹ 姜洪艳²

(1. 长春理工大学电子信息工程学院, 130022, 长春; 2. 长春市吉海测控技术有限责任公司, 130012, 长春)

摘要: 针对彩色分焦平面偏振图像全分辨率重建中, 存在的颜色失真、边缘退化和偏振参数不一致问题, 提出一种基于上下文双自相关注意力的偏振图像超分辨率重建网络。该网络由双路特征提取模块、上下文双自相关注意力 Transformer 和偏振特征重建模块组成。其中, 双路特征提取模块用于增强浅层结构与纹理特征; 上下文双自相关注意力 Transformer 通过上下文门控融合、位置感知自相关注意力机制和局部深度自相关注意力机制, 建立不同偏振方向之间的空间结构关系与通道相关性; 偏振特征重建模块通过特征精炼和像素重排恢复全分辨率彩色偏振图像。研究表明: 在 SunPol-Real、OL 和 Monno 数据集上, 所提网络对低分辨率图像进行超分辨率重建, 0°、45°、90° 和 135° 这 4 个方向图像的峰值信噪比的平均值分别达到 38.16、39.39 和 40.02 dB, 结构相似性的平均值分别达到 0.967 6、0.965 8 和 0.981 5, 从而验证了该网络能够改善边缘模糊、纹理缺失和偏振信息不一致的问题。该研究为彩色分焦平面偏振图像的超分辨率重建提供了一种有效的技术路径。

关键词: 偏振图像; 超分辨率; 自相关注意力; 位置感知; 局部深度

中图分类号: TP391 **文献标志码:** AA

文章编号: 0253-987X(XXXX)XX-0001-15

Polarization Image Super-Resolution Reconstruction Network Using Contextual Dual Self-Correlation Attention

CHEN Guangqiu¹, ZHANG Sai¹, LIU Fengming¹, GUAN Chengcai¹, LIU Guofeng¹, DUAN Jin¹, JIANG Hongyan²

(1. School of Electronic Information Engineering, Changchun University of Science and Technology, Changchun 130022, China;

2. Changchun Genhous Technology Co., Ltd., Changchun 130012, China)

Abstract: To address the problems of color distortion, edge degradation, and inconsistency in polarization parameters during the full-resolution reconstruction of color division-of-focal-plane (DoFP) polarization images, a polarization image super-resolution reconstruction network based on contextual dual autocorrelation attention is proposed. The network consists of a dual-branch feature extraction module, a contextual dual-autocorrelation attention Transformer, and a polarization feature reconstruction module. Specifically, the dual-branch feature extraction module is employed to enhance

收稿日期: 2026-04-02。

shallow structural and texture features. The contextual dual-autocorrelation attention Transformer establishes spatial structural relationships and channel correlations among different polarization orientations through contextual gated fusion, position-aware autocorrelation attention, and local depth-wise autocorrelation attention mechanisms. The polarization feature reconstruction module restores full-resolution color polarization images through feature refinement and pixel rearrangement. Experimental results show that, on the SunPol-Real, OL, and Monno datasets, the proposed network achieves average peak signal-to-noise ratio values of 38.16, 39.39, and 40.02 dB, respectively, and average structural similarity index values of 0.967 6, 0.965 8, and 0.981 5, respectively, across the reconstructed images at the 0°, 45°, 90°, and 135° polarization orientations. These results demonstrate that the proposed network can effectively alleviate edge blurring and texture loss while improving the consistency of polarization information. This study provides an effective technical approach for the super-resolution reconstruction of color DoFP polarization images.

Keywords: Polarization image; Super-resolution; Self-correlation attention; Position awareness; Local depth

偏振成像系统能够测量目标反射或透射光的偏振态,在遥感成像、表面检测和自动驾驶目标识别等领域被广泛应用^[1]。分焦平面(DoFP)偏振成像系统^[2]通过微偏振阵列可在单次成像中获取0°、45°、90°和135°这4个方向的偏振信息。然而,由于不同偏振方向的像素以周期性方式排布,单一偏振方向的有效采样数量减少,从而导致彩色DoFP图像空间分辨率的进一步降低;同时,不同偏振方向间的空间位置差异会引入瞬时视场误差,从而影响偏振信息的测量精度。因此,偏振图像超分辨率的重建与去马赛克技术被用于缓解有效像素稀疏、图像模糊和细节丢失等问题^[3]。现有DoFP偏振图像的重建方法主要包括插值方法^[4]、稀疏表示方法^[5]和深度学习方法^[6]。其中,插值方法计算简单,但容易产生模糊、锯齿和细节缺失;稀疏表示方法可保留部分结构信

息,但字典构建复杂、计算效率较低。近年来,深度学习方法逐渐成为该领域的主要研究方向,相关研究从早期基于CNN的偏振去马赛克网络^[7],发展到彩色-偏振联合重建^[8]、双分支特征建模^[9]、偏振特性引导的超分辨率重建^[10-11]以及空间-物理一致性约束模型^[12]。尽管上述方法在重建精度上优于传统方法,但在复杂纹理、大尺度结构、低光照、强噪声和强反射等场景下,仍可能出现边缘断裂、颜色失真、马赛克伪影和偏振参数局部误差等问题。其主要原因在于,现有深度学习方法多依赖局部卷积特征提取与简单特征融合,对DoFP偏振图像中长距离空间依赖、彩色-偏振信息耦合以及不同偏振方向之间的相关性,难以充分建模,从而影响偏振度(DOP)和偏振角(AOP)的计算准确性。

针对上述问题,本文提出一种结合上下文双自相关注意力Transformer

(CDSA-Transformer)的偏振图像超分辨率重建网络,并命名为CDAT-SR。该网络通过双路特征提取模块,提取浅层结构与纹理特征,利用CDSA-Transformer空间位置相关性和通道深度相关性进行建模,并通过偏振特征重建模块恢复高分辨率彩色偏振图像。本文有如下主要工作。

(1)针对颜色失真问题,设计上下文门控融合模块和位置感知自相关注意力机制,用于增强多尺度上下文信息融合和不同偏振方向间的空间结构建模能力,从而改善多偏振方向图像重建的一致性。

(2)针对边缘与轮廓重建不足问题,在偏振特征重建模块中引入特征精炼层,通过多尺度卷积分支补偿高频边缘和纹理信息,提升细节重建质量。

(3)针对伪影和偏振参数误差问题,

引入局部深度自相关注意力机制,并结合深度卷积前反馈网络和L1-Stokes联合损失函数,增强不同偏振方向特征之间的相关性建模和偏振信息的一致性约束。

1 偏振图像超分辨率重建网络设计

本文结合双路特征提取(DPFE)模块、上下文双自相关注意力Transformer、偏振特征重建(PFR)模块三部分构建偏振图像超分辨(去马赛克)重建网络,其结构如图1所示。图中,96、192、384分别为对应网络层输出特征图的通道数。由图1可知,DPFE模块由线性、非线性特征提取支路并联组成;CDSA-Transformer由5组CDSA-Transformer模块(每组3个)组成;PFR模块由特征精炼层与特征重塑层组成。

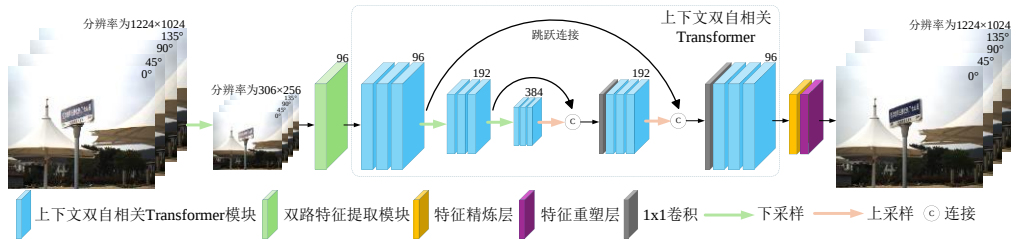


图1 偏振图像超分辨率重建网络结构

Fig. 1 Structure of polarization image super-resolution reconstruction network

1.1 双路特征提取模块

为增强网络对浅层特征的细节表达与非线性建模能力,本文设计了双路特征提取模块。该模块包含非线性特征提

取支路和线性调制支路:非线性支路采用三层卷积构成的Bottleneck结构,并使用斜率为0.2的LeakyReLU激活函数提取局部纹理特征;线性支路通过单层

卷积生成门控信号, 对非线性特征进行逐元素调制, 从而获得 DPFE 模块输出。

1.2 上下文双自相关注意力 Transformer 模块

为了增强网络的上下文感知能力, 维持偏振图像的物理属性, 本文设计了一种上下文双自相关注意力 Transformer。由图1可知, 该模块采用跳跃连接, 形成 U-net 结构, 由5层深度特征提取层组成。前3层每层包含3个 CDSA-Transformer 模块, 并在前两层中含有下采样; 后2层每层则由1个 1×1 卷积层串联3个 CDSA-Transformer 模块组成, 并含有上采样。

设 CDSA-Transformer 模块的输入特征图张量 $F \in \mathbf{R}^{B \times C \times H \times W}$, 其中, B 、 C 、 H 、 W 分别为批处理维度、通道数、高度、宽度; 深度特征提取层的输出张量为 $F_{i,j} \in \mathbf{R}^{B_{i,j} \times C_{i,j} \times H_{i,j} \times W_{i,j}}$, 其中, i, j 为第 i 层第 j 个模块。

首先, 将其按照固定窗口大小进行窗口划分得到 CDSA 机制的输入张量 $E_{i,j} \in \mathbf{R}^{B_{i,j}^w \times C_{i,j}^w \times H_w \times W_w}$, 其中, 上、下角标 w 表示划分窗口, $B_{i,j}^w = B_{i,j} \times \frac{H_{i,j} \times W_{i,j}}{H_w \times W_w}$ 。然后, 将 $E_{i,j}$ 送入到 CDSA 机制, 得到输出张量 $M_{i,j} \in \mathbf{R}^{B_{i,j}^w \times N \times C_{i,j}^w}$, 其中, $N = H_w \times W_w$ 为窗口内的像素总数, 将其与输入特征图进行逐像素相加得到残差特征张量 $M'_{i,j} \in \mathbf{R}^{B_{i,j}^w \times N \times C_{i,j}^w}$ 。随后, 将 $M'_{i,j}$ 进行层归一化, 输入到具有残差结构的深度卷积前反馈网络(DCFN)中, 得到 DCFN

的输出张量 $T_{i,j} \in \mathbf{R}^{B_{i,j}^w \times N \times C_{i,j}^w}$ 。最后, 将 $T_{i,j}$ 重塑成 CDSA-Transformer 模块的输出特征张量 $F_{i,j} \in \mathbf{R}^{B_{i,j} \times C_{i,j} \times H_{i,j} \times W_{i,j}}$ 。

1.2.1 上下文双自相关注意力机制

首先, 将 $E_{i,j} \in \mathbf{R}^{B_{i,j}^w \times C_{i,j}^w \times H_w \times W_w}$ 通过上下文门控融合 (CGF) 模块得到 $G_{i,j} \in \mathbf{R}^{B_{i,j}^w \times 3C_{i,j}^w \times H_w \times W_w}$ 。其次, 将通道维度拆分获得3个张量 $Q_{i,j}, K_{i,j}, V_{i,j} \in \mathbf{R}^{B_{i,j}^w \times C_{i,j}^w \times H_w \times W_w}$, 其中, $Q_{i,j}$ 表示查询张量, $K_{i,j}$ 表示键张量, $V_{i,j}$ 表示数值张量。然后, 将其重塑为窗口序列形式 $Q'_{i,j}, K'_{i,j}, V'_{i,j} \in \mathbf{R}^{B_{i,j}^w \times N \times C_{i,j}^w}$ 进行位置感知自相关注意力(PASCA)和局部深度自相关注意力(LDSCA)机制, 分别得到输出张量 $\phi_{i,j} \in \mathbf{R}^{B_{i,j}^w \times N \times C_{i,j}^w}$ 和 $\varphi_{i,j} \in \mathbf{R}^{B_{i,j}^w \times N \times C_{i,j}^w}$ 。最后, 将 $\phi_{i,j}$ 和 $\varphi_{i,j}$ 进行拼接和全连接层降维, 得到 CDSA 机制的输出张量 $M_{i,j} \in \mathbf{R}^{B_{i,j}^w \times N \times C_{i,j}^w}$ 。

PASCA 与 LDSCA 在计算形式上借鉴了自注意力的相关性建模思想, 但其建模目标并非自然图像中的语义增强, 而是针对彩色分焦平面偏振图像中多偏振方向之间的协同关系。

(1) 上下文门控融合模块。针对颜色信息与偏振信息耦合、特征交织的问题, 本文采用 CGF 模块替代传统线性投影层, 以实现多源特征的自适应筛选与融合, 从而增强颜色与偏振信息的协同重建能力。CGF 模块由上、中、下3条分支组成: 上分支采用三层卷积结构提取较大感受野下的上下文特征; 中分支同样由三层卷积构成, 用于建模局部区域内的上下文信息; 下分支通过轻量

级卷积生成自适应门控信号。随后,上、中分支的输出通过逐像素相加实现上下文的特征融合,并与下分支生成的门控信号逐元素相乘,从而得到 CGF 模块的输出。

(2) 位置感知自相关注意力机制。首先,将 $\mathbf{Q}'_{ij}, \mathbf{K}'_{ij}, \mathbf{V}'_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 沿向量维度拆分成 h 个部分,每个部分的维度为 $d = C_{ij}^w/h$,此时 $\mathbf{Q}''_{ij}, \mathbf{K}''_{ij}, \mathbf{V}''_{ij} \in \mathbf{R}^{B_{ij}^w \times h \times N \times d}$ 。随后,计算 $\mathbf{Q}''_{ij}$ 与 $\mathbf{K}''_{ij}$ 之间的相似度,并融合相对位置偏置进行归一化,得到 PASCA 注意力权重矩阵 $\mathbf{A}_{ij}^{\text{PASCA}} \in \mathbf{R}^{B_{ij}^w \times h \times N \times N}$ 。最后,将 $\mathbf{A}_{ij}^{\text{PASCA}}$ 与 $\mathbf{V}''_{ij}$ 进行特征聚合加权,得到 $\mathbf{X}_{ij}^{\text{PASCA}} \in \mathbf{R}^{B_{ij}^w \times h \times N \times d}$ 。上述过程可表示为

$$\mathbf{A}_{ij}^{\text{PASCA}} = \text{softmax}\left(\frac{\mathbf{Q}''_{ij}(\mathbf{K}''_{ij})^T}{\sqrt{d}} + \mathbf{B}_{\text{pos}}\right) \quad (1)$$

$$\mathbf{X}_{ij}^{\text{PASCA}} = \mathbf{A}_{ij}^{\text{PASCA}} \mathbf{V}''_{ij} \quad (2)$$

式中: $\text{softmax}(\cdot)$ 表示归一化运算; $\mathbf{B}_{\text{pos}}$ 为相对位置偏置。相对位置偏置是利用一个多层感知机网络^[13-14]根据窗口内像素对的相对坐标差值动态生成相对位置偏置。

接着将 $\mathbf{X}_{ij}^{\text{PASCA}} \in \mathbf{R}^{B_{ij}^w \times h \times N \times d}$ 进行重塑得到 PASCA 机制的输出张量 $\phi_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 。

PASCA 在局部窗口内对不同偏振方向间的空间结构相关性进行建模,并通过相对位置偏置增强边缘、纹理及弱结构区域的协同恢复能力,从而缓解方向间结构的不一致和偏振参数的失真问题。

(3) 局部深度自相关注意力机制。首先,依据窗口内像素数,将 $\mathbf{Q}'_{ij}, \mathbf{K}'_{ij}, \mathbf{V}'_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 拆分成 b 个部分,每部分的像素数为 $s = N/b$,此时 $\mathbf{Q}''_{ij}, \mathbf{K}''_{ij}, \mathbf{V}''_{ij} \in \mathbf{R}^{B_{ij}^w \times b \times s \times C_{ij}^w}$ 。随后,通过计算查询向量 $\mathbf{Q}''_{ij}$ 与键向量 $\mathbf{K}''_{ij}$ 之间的相似度,并进行归一化得到 LDSCA 注意力权重矩阵 $\mathbf{A}_{ij}^{\text{LDSCA}} \in \mathbf{R}^{B_{ij}^w \times b \times s \times C_{ij}^w}$ 。最后,将 $\mathbf{A}_{ij}^{\text{LDSCA}}$ 与值向量 $\mathbf{V}''_{ij}$ 进行特征聚合加权,得到 $\mathbf{X}_{ij}^{\text{LDSCA}} \in \mathbf{R}^{B_{ij}^w \times b \times s \times C_{ij}^w}$,上述过程可表示为

$$\mathbf{A}_{ij}^{\text{LDSCA}} = \text{Softmax}\left(\frac{(\mathbf{Q}''_{ij})^T \mathbf{K}''_{ij}}{\sqrt{s}}\right) \quad (3)$$

$$\mathbf{X}_{ij}^{\text{LDSCA}} = \mathbf{V}''_{ij} (\mathbf{A}_{ij}^{\text{LDSCA}})^T \quad (4)$$

接着将 $\mathbf{X}_{ij}^{\text{LDSCA}} \in \mathbf{R}^{B_{ij}^w \times b \times s \times C_{ij}^w}$ 进行重塑,得到 LDSCA 机制的输出结果 $\phi_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 。

LDSCA 在局部窗口内对不同偏振方向特征之间的深度相关性进行建模,增强了多方向偏振特征的联合表征能力,从而提升局部结构的偏振信息的一致性。

最后,将 $\phi_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 与 $\phi_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 在通道维度进行拼接与全连接层降维后,得到 CDSA 机制的输出张量 $\mathbf{M}_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 。

1.2.2 深度卷积前反馈网络

DCFN 的输入张量为 $\mathbf{M}'_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 。首先,利用全连接层结合 GeLU 激活函数进行通道维度扩展得到通道扩展特征张量 $\mathbf{O} \in \mathbf{R}^{B_{ij}^w \times N \times 2C_{ij}^w}$,重塑得到空间特征张量 $\mathbf{O}' \in \mathbf{R}^{B_{ij}^w \times 2C_{ij}^w \times H_w \times W_w}$;

结合 5×5 卷积层和GeLU激活函数进行局部空间特征提取。随后,进行结果重塑,得到局部空间特征张量 $P \in \mathbf{R}^{B_{ij}^w \times N \times 2C_{ij}^w}$,将其与 $O \in \mathbf{R}^{B_{ij}^w \times N \times 2C_{ij}^w}$ 逐像素相加,得到残差融合特征张量

$P' \in \mathbf{R}^{B_{ij}^w \times N \times 2C_{ij}^w}$ 。最后,通过全连接层降维,得到 DCFN 的输出张量 $T_{ij} \in \mathbf{R}^{B_{ij}^w \times N \times C_{ij}^w}$ 。上述的上下文双自相关Transformer模块整体结构如图2所示。

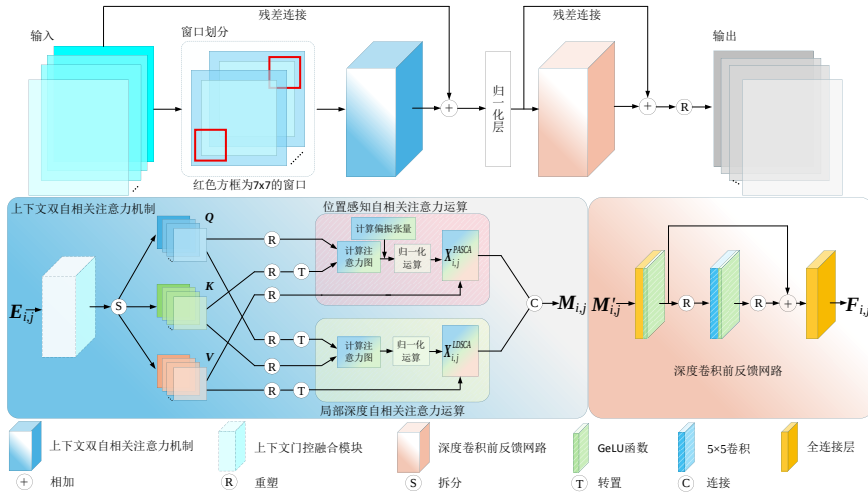


图2 上下文双自相关Transformer模块结构

Fig. 2 Structure of context dual self-correlation attention transformer module

1.3 偏振特征重建模块

为将 CDSA-Transformer 提取的深层特征转换为高分辨率图像,本文通过串联特征精炼层与特征重塑层构建偏振

特征重建(PFR)模块,其结构如图5所示。图中, $F_{5,3} \in \mathbf{R}^{B_{5,3} \times C_{5,3} \times H_{5,3} \times W_{5,3}}$ 为特征精炼层的输入, Z 为偏振特征重建模块输出,通道数为12。

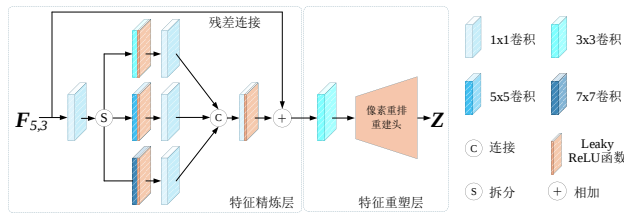


图3 偏振特征重建模块结构

Fig. 3 Structure of polarization feature reconstruction module

1.3.1 特征精炼层

首先,将 $F_{5,3}$ 通过 1×1 的卷积进行通道扩张得到扩展特征张量 $F'_{5,3} \in \mathbf{R}^{B_{5,3} \times 3C_{5,3} \times H_{5,3} \times W_{5,3}}$,再将 $F'_{5,3}$ 拆分为3个子特征张量,各子特征张量的通道数均为 $C_{5,3}$,并分别输入并行的多尺度卷积分支。3个分支分别采用 3×3 、 5×5 和 7×7 卷积提取不同感受野下的局部特征,并结合LeakyReLU激活函数和 1×1 卷积进行非线性特征映射;然后,将3个分支的输出沿通道维度进行拼接,得到多尺度融合特征张量 $U_{cat} \in \mathbf{R}^{B_{5,3} \times 3C_{5,3} \times H_{5,3} \times W_{5,3}}$;随后,通过 1×1 的卷积进行降维和LeakyReLU函数激活得到融合特征张量 $U_c \in \mathbf{R}^{B_{5,3} \times C_{5,3} \times H_{5,3} \times W_{5,3}}$;最后, U_c 与输入特征进行残差连接得到输出张量 $U_{refined} \in \mathbf{R}^{B_{5,3} \times C_{5,3} \times H_{5,3} \times W_{5,3}}$ 。

1.3.2 特征重塑层

特征重塑层是由 3×3 的卷积和像素重排重建头组成,用来重塑偏振图像的空间分辨率。特征重塑层如图3所示。像素级重排头是将通道维度重排为空间维度,从而实现高效的上采样。

1.4 L1-Stokes 损失函数

针对偏振图像重建任务,本文在损失函数中引入基于Stokes参数的物理一致性约束,对重建图像的参量 S_0 、 S_1 与 S_2 进行联合优化,以在提升重建质量的同时,保持偏振信息的一致性。最终损失函数 L_{total} 是由平均绝对误差损失函数^[15]与Stokes损失函数^[16]组成,表达式为

$$L_{total} = \frac{1}{N} \sum_{n=1}^N \|I_{SR}^{(n)} - I_{HR}^{(n)}\| + \lambda_0 \|S_0^{SR} - S_0^{HR}\| + \lambda_1 \|S_1^{SR} - S_1^{HR}\| + \lambda_2 \|S_2^{SR} - S_2^{HR}\| \quad (5)$$

$$S_0 = \frac{I_0 + I_{45} + I_{90} + I_{135}}{2}; S_1 = I_0 - I_{90}; S_2 = I_{45} - I_{135} \quad (6)$$

式中: N 为总像素数; n 为像素索引;上、下角标SR、HR分别表示超分辨率重建结果与参考图像; I_0 、 I_{45} 、 I_{90} 、 I_{135} 分别为 0° 、 45° 、 90° 和 135° 偏振方向对应的光强图像; λ_0 、 λ_1 、 λ_2 为权重系数,根据参考文献^[17]可知权重系数可分别取值为0.8、0.1、0.1。

2 实验与结果分析

本文实验数据集包括自制数据集SunPol-Real以及公开数据集OL和Monno^[18]。SunPol-Real数据集采用Lucid彩色偏振相机(PHX050S-QC)在真实场景下采集,该数据集包含海面、房屋等多类自然场景。经数据筛选与整理后,包含1000组彩色偏振图像,其中700组作为训练集,300组作为测试集。3种数据集的部分可视化结果如图4所示。

公开数据集OL和Monno主要采集于相对可控的实验室环境,适合模型的公平对比,但场景复杂度有限。为评估模型在真实开放场景的鲁棒性,本文构建了SunPol-Real自制数据集,其包含自然光照变化、复杂背景和噪声干扰等因素。

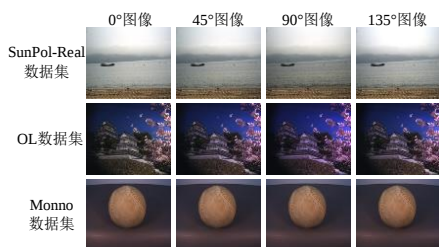


图4 SunPol-Real、OL和Monno数据集部分可视化图像

Fig. 4 Partial visualization images from the SunPol-Real, OL, and Monno datasets.

为验证 CDAT-SR 的有效性, 将其与 Bilinear、Bicubic、EARI^[19]、PCDP^[20]、NLCSR^[21]、CPDNet、PIDSr 和 Hit-SR^[22] 进行对比。实验环境配置^[23] 如表 1 所示。本文的客观评价指标采用峰值信噪比 (PSNR)^[24] 和结构相似性 (SSIM)^[25]。

表1 实验环境配置

Table 1 Configuration of experimental environment

类别	配置
操作系统	Ubuntu 20.04
深度学习框架	PyTorch 2.0.0, torchvision 0.12.0
CUDA 版本	CUDA 11.8
Python 版本	Python 3.8
GPU	NVIDIA GeForce RTX 3090
CPU	Intel(R) Xeon(R) Gold 6330

实验中 DOP 图像和 AOP 图像的计算公式为

$$I_{\text{DOP}} = \frac{\sqrt{S_1^2 + S_2^2}}{S_0}; I_{\text{AOP}} = \frac{1}{2} \tan^{-1} \left(\frac{S_1}{S_2} \right) \quad (7)$$

其中, S_1 、 S_2 、 S_3 是由式(6)获得。

在上述对比实验中, 对深度学习类超分辨率模型进行参数量、浮点运算量以及平均推理时间的对比, 结果如表 2 所示。

表2 模型性能对比实验结果

Table 2 Experimental results of model performance compariso

模型	参数量/M	浮点运算量/G	推理时间/ms
CPDNet	0.5	4.2	4.365
PIDSr	7.6	10.5	32.212
Hit-SR	0.8	3.6	24.287
CDAT-SR	13.8	9.2	8.379

从表 2 可以看出, CDAT-SR 虽然参数量较高, 但其推理时间低于 PIDSr 和 Hit-SR, 且仅高于 CPDNet, 这表明该模型在重建性能和推理效率之间取得了较好的平衡。

2.1 SunPol-Real、OL 和 Monno 数据集的对比实验

图 5 所示为不同超分(去马赛克)模型在 SunPol-Real、OL 和 Monno 数据集上得到的结果图像。图中, 红色框为放大区域。

由图 5 可知, 各模型在局部细节恢复和偏振信息保持方面表现不同。在 SunPol-Real 数据集中, Bilinear 和 Bicubic 模型重建的广告牌文字边缘较为模糊, 局部轮廓不够清晰。EARI、PCDP 和 NLCSR 模型能够恢复主要结构, 但部分细节仍有过平滑现象, AOP 图像中也出现了不同程度的颜色

偏差。对于 OL 数据集, Bilinear 和 Bicubic 模型对橙色水果内外边缘的恢复较为模糊,表面高光和局部纹理有所弱化;EARI 和 PCDP 模型虽然保留了基本轮廓,但细节恢复仍不充分。NLCSR、CPDNet 和 Hit-SR 模型的 AOP 图像与参考图像存在较明显的颜色差异, PIDSR 模型在部分区域还表现出一定的亮度偏差。在 Monno 数据集中, Bilinear 和 Bicubic 模型重建的瓶身边缘及局部文字较为模糊。EARI 和 PCDP 模型改善了主体轮廓,但对细小纹理的恢复仍有限;NLCSR 和 CPDNet 模型的 AOP 图像存在局部颜色偏移, Hit-SR 模型在 AOP 图像和 DOP 图像中则出现较明显的亮度下降。总体来看, CDAT-SR 模型在目标边缘和局部纹理恢复方面表现更稳定,其 AOP 图像和 DOP 图像的颜色及亮度分布也更接近参考图像。

为了验证模型的稳定性,在 SunPol-Real、OL 和 Monno 数据集中随机抽取 20 组测试图像,将其测试结果 S0 图像、AOP 图像和 DOP 图像的平均结构相似性和平均峰值信噪比制作成柱状图进行比较,如图 6 所示。图中,下标为各模型名称, A_{S0} 、 A_{AOP} 、 A_{DOP} 分别为基于 S0 图像、AOP 图像和 DOP 图像计算得到的评价指标。

由图 6 可知,本文模型 CDAT-SR 的超分辨重建结果的 S0 图像、AOP 图像和 DOP 图像的客观评价指标均为最高,这说明本文模型在 SunPol-Real、

OL 和 Monno 数据集中表现稳健。

为综合评价不同模型在各偏振方向上的重建精度,本文在 SunPol-Real、OL 和 Monno 这 3 个数据集上,分别计算各模型在生成 4 个偏振方向(0° 、 45° 、 90° 和 135°)图像的平均 PSNR 和平均 SSIM,表 3 为不同模型在 3 个数据集上的 4 偏振方向平均重建结果对比。

结合表 3 和图 5 可知,客观评价结果与主观视觉评价结果基本一致,本文提出的 CDAT-SR 模型在各项评价指标和视觉重建质量上均表现最优。

2.2 消融实验

为验证各模块的有效性,本文设置 5 组消融实验,分别考察 DPFE、CGF、PASCA 和 LDSCA 对模型性能的影响。其中,第 1 组以单路特征提取替换 DPFE,第 2 组以线性投影层替换 CGF,第 3 组和第 4 组分别仅保留 PASCA 或 LDSCA,第 5 组为完整 CDAT-SR 模型。

表 4 为消融实验客观评价数据,评价数据为在自制数据集 SunPol-Real 中,计算 0° 、 45° 、 90° 和 135° 偏振超分图像的平均峰值信噪比和平均结构相似性。

由表 4 可知,DPFE、CGF、PASCA 和 LDSCA 均对模型性能的提升具有贡献。其中,引入 DPFE 后,平均 PSNR 和 SSIM 分别提升了 5.30 dB 和 0.069 8;将线性投影层替换为 CGF 后,模型性能进一步提高。相比单独使用 PASCA 或 LDSCA,二者联合使用时取得的结果最优,这说明空间结构建模与

通道相关性建模具有互补作用。



图5 不同模型在SunPol-Real、OL和Monno数据集中的结果图像

Fig. 5 Result images of different methods on the SunPol-Real, OL, and Monno test sets

本文模型是针对偏振图像进行设计的，但为了验证模型对普通RGB图像的有效性，使用公开数据集DIV2K对本文模型进行泛化性实验，图7为泛化性实验结果。

从图7中可以看出，本文模型对于通用的RGB图像也能得到较好的结果。

2.3 Stokes 损失函数权重合理性实验

为分析Stokes损失函数中，各分量权重对模型性能的影响，本文在保持其他超参数一致，且 $\lambda_0 + \lambda_1 + \lambda_2 = 1$ 的条件下，通过调整 λ_0 、 λ_1 、 λ_2 的取值，进行权重敏感性实验，实验结果如表5和表

6所示。实验分为2组：实验一是固定 $\lambda_1=\lambda_2$ ，调整 λ_0 的取值，分析整体光强约束与偏振方向约束之间的权衡关系；实验二是固定 λ_0 ，调整 λ_1 与 λ_2 的取值，分析不同偏振方向分量对模型重建性能

的影响。考虑到AOP在低偏振区域易受噪声影响，本文采用偏振度误差作为偏振一致性的辅助评价指标 $A_{\text{DOPTerror}}$ 的表达式为

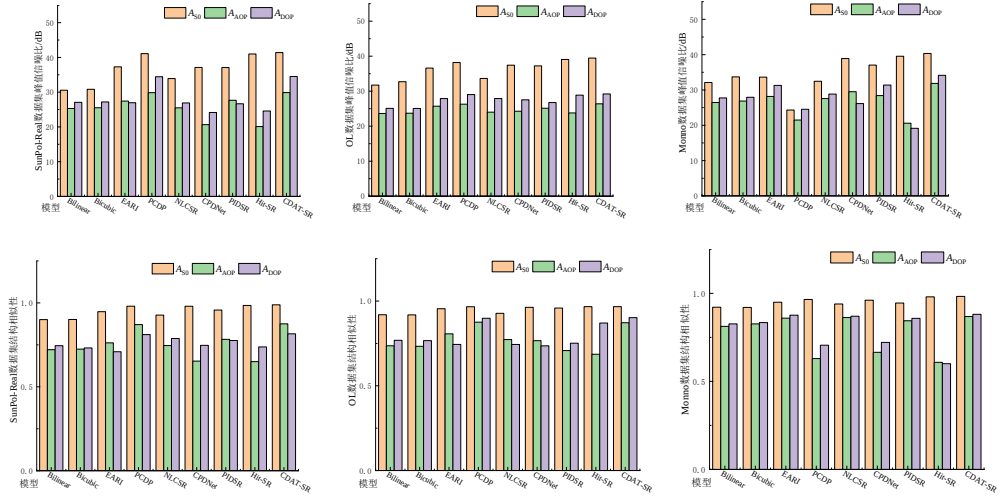


图6 不同数据集中20组测试图像的评价指标对比

Fig. 6 Comparison of evaluation metrics for 20 test images across different datasets

表3 不同模型在数据集的平均重建结果对比

Table 3 Comparison of average reconstruction results of different models on the dataset

模型	平均峰值信噪比/dB			平均结构相似性		
	SunPol-Real	OL	Monno	SunPol-Real	OL	Monno
Bilinear	29.46	31.83	31.78	0.875 8	0.903 1	0.939 3
Bicubic	29.52	31.62	32.33	0.880 1	0.898 9	0.941 9
EARI	35.04	35.46	33.24	0.922 8	0.939 3	0.947 1
PCDP	38.12	39.03	23.27	0.967 7	0.963 8	0.957 8
NLCSR	32.77	33.26	32.04	0.902 2	0.916 3	0.936 7
CPDNet	35.10	33.48	36.47	0.957 7	0.945 8	0.945 8
PIDSr	35.41	36.42	36.37	0.934 9	0.947 8	0.942 0
Hit-SR	38.05	39.27	39.24	0.964 4	0.963 7	0.979 8
CDAT-SR	38.16	39.39	40.02	0.967 6	0.965 8	0.981 5

表4 消融实验结果

Table 4 Results of ablation experiment

组号	是否加入模型				平均峰值信噪比/dB	平均结构相似性
	DPFE	CGF	PASCA	LDSCA		
第1组	否	是	是	是	32.86	0.897 7
第2组	是	否	是	是	34.16	0.913 4
第3组	是	是	是	否	35.34	0.924 5
第4组	是	是	否	是	36.86	0.946 2
第5组	是	是	是	是	38.16	0.967 5



图7 泛化性实验结果

Fig. 7 Results of generalization experiments

$$A_{\text{DOPerror}} = \frac{1}{N} \sum |I_{\text{DOP}}^{\text{SR}} - I_{\text{DOP}}^{\text{HR}}| \quad (8)$$

式中： $I_{\text{DOP}}^{\text{SR}}$ 、 $I_{\text{DOP}}^{\text{HR}}$ 分别为超分辨率重建结果与参考高分辨率图像计算得到的DOP图像。

由表5和表6可知， λ_0 取0.8， λ_1 、 λ_2 取0.10时为最合理的权重参数。

2.4 无参考真实场景偏振图像超分

为验证本文模型在真实无参考场景下的适用性，本文对彩色分焦平面相机采集的原始马赛克图像进行超分辨率重建，结果如图8所示。原始马赛克图像包含混叠的颜色分量与偏振分量，经解

混后得到 0° 、 45° 、 90° 和 135° 这4个方向的彩色偏振图像，将解混后的低分辨率偏振图像输入CDAT-SR网络，在行、列方向分别进行4倍超分辨率重建。图9为超分辨重建后计算得到的S0图像、DOP图像和AOP图像及其局部放大图像（红色方框为放大区域）。

表5 权重敏感性实验一

Table 5 Experiment I on weight sensitivity

λ_0	λ_1	λ_2	峰值信噪比/dB	结构相似性	A_{DOPerror}
0	0.50	0.50	37.11	0.953 4	0.023 4
0.2	0.40	0.40	37.38	0.954 1	0.022 1
0.4	0.30	0.30	37.87	0.956 3	0.020 8
0.6	0.20	0.20	37.93	0.957 7	0.019 6
0.7	0.15	0.15	38.14	0.960 2	0.018 9
0.8	0.10	0.10	38.16	0.967 5	0.018 5
0.9	0.05	0.05	37.95	0.964 7	0.020 7
1.0	0	0	37.76	0.957 9	0.023 9

由图9可知，本文所提模型在无参考真实场景偏振图像的重建过程中，同样具备良好的有效性。

表6 权重敏感性实验二

Table 6 Experiment II on weight sensitivity

λ_0	λ_1	λ_2	峰值 信噪 比/dB	结构相 似性	A_{DOPerror}
0.8	0	0.20	37.84	0.963 7	0.022 6
	0.05	0.15	38.02	0.965 8	0.020 1
	0.10	0.10	38.16	0.967 5	0.018 5
	0.15	0.05	38.07	0.966 3	0.019 4
	0.20	0	37.91	0.964 5	0.021 3

3 结论

本文提出了一种偏振图像超分辨率重建网络。该模型通过双路特征提取、

上下文双自相关注意力 Transformer 和偏振特征重建模块，分别增强浅层结构表征、多偏振方向协同建模和高分辨率图像恢复能力。实验结果表明：CDAT-SR 在 SunPol-Real、OL 和 Monno 数据集上均取得较好的效果，在 0° 、 45° 、 90° 和 135° 这 4 个方向偏振图像上的平均峰值信噪比分别为 38.16、39.39 和 40.02 dB，平均结构相似性分别为 0.967 6、0.965 8 和 0.981 5。同时，AOP 和 DOP 图像的客观指标也整体优于对比模型。上述结果表明，CDAT-SR 能够有效改善边缘模糊、纹理细节缺失和偏振信息不一致等问题。

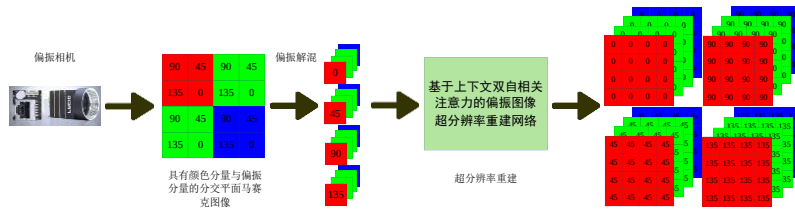


图8 无参考真实场景偏振图像超分辨率的重建

Fig. 8 Reconstruction of polarimetric image super-resolution in reference-free real-scene

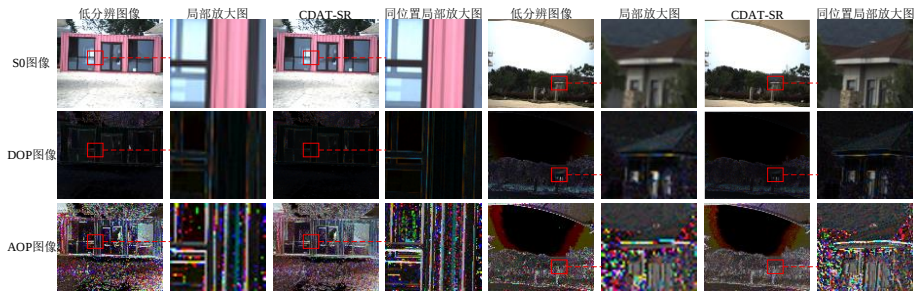


图9 无参考真实场景的S0、DOP和AOP图像

Fig. 9 S0, DOP, and AOP images of reference-free real scenes

参考文献

- [1] YU Dabing, LI Qingwu, ZHANG Zhiliang, et al. Color polarization image super-resolution reconstruction via a cross-branch supervised learning strategy [J]. Optics and Lasers in Engineering, 2023, 165: 107469.
- [2] ZHANG Jingjing, BAO Kangsheng, ZHANG Xin, et al. Conditional generative adversarial defogging algorithm based on polarization characteristics [C]//Proceedings of the 2021 33rd Chinese Control and Decision Conference (CCDC). Piscataway, NJ, USA: IEEE, 2021: 2148 – 2155.
- [3] 吴洪伍, 盖绍彦, 达飞鹏. 采用感受野优化与渐进特征融合的图像超分辨率算法[J]. 西安交通大学学报, 2025, 59(01): 136-147. WU Hongwu, GAI Shaoyan, DA Feipeng. Image super-resolution algorithm based on receptive field optimization and progressive feature fusion [J]. Journal of Xi'an Jiaotong University, 2025, 59(1): 136 – 147.
- [4] SU Yiming, HAO Chunhui, CHEN Wen, et al. Universal polarization demosaicking using Newton polynomial and residual interpolation [J]. IEEE Sensors Journal, 2025, 25(7): 11508 – 11520.
- [5] SI Weijian, WANG Yan, ZHANG Chunjie. Three-parallel co-prime polarization sensitive array for 2-D DOA and polarization estimation via sparse representation [J]. IEEE Access, 2019, 7: 15404 – 15413.
- [6] CAO Yin, ZHANG Heng, FAN Rong, et al. Underwater image restoration based on color polarization imaging and deep learning [C]//Proceedings of the 2025 6th International Conference on Computer Vision, Image and Deep Learning (CVIDL). Piscataway, NJ, USA: IEEE, 2025: 612 – 616.
- [7] ZHANG Junchao, SHAO Jianbo, LUO Haibo, et al. Learning a convolutional demosaicing network for microgrid polarimeter imagery [J]. Optics Letters, 2018, 43(18): 4534 – 4537.
- [8] WEN Sijia, ZHENG Yinqiang, LU Feng, et al. Convolutional demosaicing network for joint chromatic and polarimetric imagery [J]. Optics Letters, 2019, 44(22): 5646 – 5649.
- [9] SUN Yuanyuan, ZHANG Junchao, LIANG Rongguang. Color polarization demosaicking by a convolutional neural network [J]. Optics Letters, 2021, 46(17): 4338 – 4341.
- [10] GUAN Chao, LI Jiangtao, TIAN Yaping, et al. Polarization-enhanced super-resolution imaging reconstruction based on a denoising diffusion probabilistic model [J]. Optics Express, 2025, 33(19): 39238 – 39248.
- [11] LI Shuoshi, ZHOU Yuan, KUNG S Y. PSRNet: A progressive self-refine network for lightweight optical remote sensing image dehazing [J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 1 – 13, Art. no. 5648413.
- [12] ZHOU Shuangfan, ZHOU Chu, LYU Youwei, et al. PIDSr: Complementary polarized image demosaicing and super-resolution [C]//Proceedings of the 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ, USA: IEEE, 2025: 16081 – 16090.
- [13] WANG Wenxiao, CHEN Wei, QIU Qibo, et al. CrossFormer++: A versatile vision transformer hinging on cross-scale atten-

- tion [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(5): 3123 – 3136.
- [14] 陈广秋, 刘枫铭, 段锦, 等. 基于行列栅格感知 Transformer 的车道线检测方法 [J]. 光学精密工程, 2026, 34(6): 953 – 972.
- CHEN Guangqiu, LIU Fengming, DUAN Jin, et al. Row-column grid-aware Transformer based lane detection method [J]. Optics and Precision Engineering, 2026, 34(6): 953 – 972.
- [15] 左龙, 张鹏, 荆树旭, 等. 用于图像超分辨率重建的双通道残差网络 [J]. 西安交通大学学报, 2022, 56(1): 158 – 164.
- ZUO Long, ZHANG Peng, JING Shuxu, et al. Dual-channel residual network for image super-resolution reconstruction [J]. Journal of Xi'an Jiaotong University, 2022, 56(1): 158 – 164.
- [16] GIL J J, OSSIKOVSKI R. Polarized light and the Mueller matrix approach [M]. Boca Raton, FL, USA: CRC Press, 2022.
- [17] HU Haofeng, YANG Shiyao, LI Xiaobo, et al. Polarized image super-resolution via a deep convolutional neural network [J]. Optics Express, 2023, 31(5): 8535 – 8547.
- [18] MORIMATSU M, MONNO Y, TANAKA M, et al. Monochrome and color polarization demosaicking based on intensity-guided residual interpolation [J]. IEEE Sensors Journal, 2021, 21(23): 26985 – 26996.
- [19] MORIMATSU M, MONNO Y, TANAKA M, et al. Monochrome and color polarization demosaicking using edge-aware residual interpolation [C]//Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP). Piscataway, NJ, USA: IEEE, 2020: 2571 – 2575.
- [20] WU Rongyuan, ZHAO Yongqiang, LI Ning, et al. Polarization image demosaicking using polarization channel difference prior [J]. Optics Express, 2021, 29(14): 22066 – 22079.
- [21] LUO Yidong, ZHANG Junchao, SHAO Jianbo, et al. Learning a non-locally regularized convolutional sparse representation for joint chromatic and polarimetric demosaicking [J]. IEEE Transactions on Image Processing, 2024, 33: 5029 – 5044.
- [22] ZHANG Xiang, ZHANG Yulun, YU Fisher. HiT-SR: Hierarchical transformer for efficient image super-resolution [C]//Computer Vision—ECCV 2024. Cham, Switzerland: Springer, 2024: 483 – 500.
- [23] 徐胜军, 邓博文, 史亚, 等. 一种编解码结构的车牌图像超分辨率网络 [J]. 西安交通大学学报, 2022, 56(10): 101 – 110.
- XU Shengjun, DENG Bowen, SHI Ya, et al. An encoder-decoder structured network for license plate image super-resolution [J]. Journal of Xi'an Jiaotong University, 2022, 56(10): 101 – 110.
- [24] 高峰, 汲胜昌, 郭洁, 等. 采用对比学习的多阶段 Transformer 图像去雾方法 [J]. 西安交通大学学报, 2023, 57(1): 195 – 210.
- GAO Feng, JI Shengchang, GUO Jie, et al. A multi-stage Transformer network for image dehazing based on contrastive learning [J]. Journal of Xi'an Jiaotong University, 2023, 57(1): 195 – 210.
- [25] LIU Yibin, WANG Shengbin, LUO WU Guoqing, et al. Microwave image super-resolution using polarization [C]//Proceedings of the 2025 10th International Conference on Signal and Image Processing (ICSIP). Piscataway, NJ, USA: IEEE, 2025: 258 – 262.