

引用格式: XU Guoming, WANG Jie, MA Jian, et al. Polarization Image Super-resolution Reconstruction Based on Dual Attention Residual Network[J]. Acta Photonica Sinica, 2022, 51(4):0410001

徐国明,王杰,马健,等. 基于双重注意力残差网络的偏振图像超分辨率重建[J]. 光子学报, 2022, 51(4):0410001

基于双重注意力残差网络的偏振图像超分辨率重建

徐国明^{1,2,3}, 王杰¹, 马健^{1,2}, 王勇³, 刘佳庆¹, 李毅⁴

(1 安徽大学 互联网学院, 合肥 230039)

(2 安徽大学 农业生态大数据分析与应用技术国家地方联合工程研究中心, 合肥 230601)

(3 陆军炮兵防空兵学院 偏振光成像探测技术安徽省重点实验室, 合肥 230031)

(4 安徽文达信息工程学院 智能技术研究所, 合肥 231201)

摘要:在利用深度学习进行偏振图像计算成像过程中,图像映射函数的解空间极大、空间分辨率一般较低,难以生成清晰的纹理细节且存在高频信息缺失等问题。为解决该问题,提出一种结合双注意力机制的深度残差偏振图像超分辨率网络。该网络由一个具有全局跳跃连接的残差网络组成,包含 10 个残差组,每个残差组包含 20 个具有局部跳跃连接的双重注意力块级联的残差块;同时考虑通道间的相互依赖性,设计自适应通道特征调整机制;引入级联的空间注意力块,将残差的特征更集中于关键的空间内容。将所提方法与 Bicubic、SRCNN、FSRCNN、EDSR 等方法进行对照实验与成像系统对比校正实验,结果表明该方法重建图像纹理细节更加丰富,亮度均匀,较为接近成像系统的高清图像,同时峰值信噪比和结构相似性指标优于其他方法但参数量仅约为 EDSR 的 2/5。

关键词: 计算成像;超分辨率;深度残差网络;偏振图像;双重注意力块

中图分类号: TP391

文献标识码: A

doi: 10.3788/gzxb20225104.0410001

0 引言

在偏振成像探测中,气溶胶或者探测目标的物理属性差异和变化由偏振特性来表征,高维度的偏振特性能有效提高目标与背景的对比度,从而为实现目标空间结构的反演奠定基础,并能够在杂乱的背景中增强对目标的识别效果^[1]。这些特性使得偏振成像探测被广泛应用于目标识别与跟踪、机器视觉、地理信息分析、遥感图像处理等领域^[2]。空间调制型全偏振成像技术是继传统的分时和同时偏振成像技术后发展起来的新体制偏振成像技术^[3]。曹奇志等^[4]提出了基于改进型 Savart 偏光镜的微型快拍成像测偏技术。杨敏等^[5]利用单像元非均匀性差分图像校正方法和迭代排序方式,设计了面向运动目标探测的分时型红外偏振成像系统。在实际应用中,由于受成像距离远及大气扰动的影响使得投影在焦平面上图像的极限分辨率严重下降^[6](远小于光学系统衍射极限分辨率),从而造成获得的偏振图像空间分辨率较低,而高分辨率的图像对目标探测精度具有重要的意义和价值。另一方面,图像的空间分辨率受限于探测器像元数目,为此,在不替换硬件成像系统的情况下,通常采用超分辨率(Super-resolution Reconstruction, SR)方法,通过利用信息融合和信号处理技术来提高图像的采样频率^[7],从而获取高于成像系统的高分辨率图像。该方法是图像处理和实际工程应用中常用的技术手段,也是底层计算机视觉的热点研究问题。

图像超分辨率重建是指采用图像处理和机器学习技术,从同一场景中已有的一张或多张低分辨率

基金项目: 国家自然科学基金(No.61906118),安徽省自然科学基金(Nos.1908085MF208, 2108085MF230),陆军装备部十三五预研子课题,安徽省高校自然科学研究重点项目(No.KJ2019A0906)

第一作者: 徐国明(1979—),男,教授,博士,主要研究方向为偏振成像探测、图像表示与超分辨率。Email: xgm121@163.com

通讯作者: 王杰(1996—),男,硕士研究生,主要研究方向为图像超分辨率、机器学习。Email: 1197545193@qq.com。

收稿日期: 2021-09-18; **录用日期:** 2021-11-17

<http://www.photon.ac.cn>

(Low-Resolution, LR)图像构建高分辨率(High-Resolution, HR)图像的技术^[8]。徐国明等^[9]针对空间调制型全偏振计算成像系统,提出基于场景特征迁移学习的超分辨率方法。孟祥超等^[10]基于我国2018年5月发射的高分五号卫星上搭载的多角度偏振成像仪展开研究,通过融合HR的全色影像和LR的高光谱影像,重建得到同时具有高空间分辨率和高光谱分辨率的遥感影像。近年来,基于学习^[11-12]的图像超分辨率重建方法因为良好的性能而受到广泛关注。DONG C等^[13]提出了基于卷积神经网络的图像超分辨率(Super-resolution Convolutional Neural Network, SRCNN),但其网络层次太浅,无法充分提取图像特征。DONG C等^[14]进一步提出了一种快速的卷积神经网络的结构(Accelerating the Super-resolution Convolutional Neural Network, FSRCNN)。该网络选择相比SRCNN中更小尺寸的滤波器,增加更多的映射层并在网络的最后引入了反卷积层。LIM B等^[15]通过使用简化的残差块,建立了一个非常宽的增强深度超分辨率网络(Enhanced Deep residual Network Super-resolution, EDSR)。注意力机制在底层计算机视觉问题中表现出了显著的性能^[16-18]。HU J等^[19]提出了通道注意力块可以提高网络的分辨能力,不同的特征通道自适应地赋予不同的权值,使网络集中于更有信息量的特征。LIU J等^[20]提出了一个新的残差特征聚合框架,结合增强的空间注意力块,更有效地提取图像特征。

上述方法直接应用于实际偏振成像系统的图像超分辨率,无法充分考虑偏振成像特性,偏振图像的先验信息不能有效表达,从而导致模型的性能有限。这些方法忽略了跨特征通道区分的学习能力、未能充分利用残留特征以及缺失关键位置的特征提取,从而导致性能下降。偏振图像超分辨率难以产生纹理细腻、边缘清晰的细节信息。针对该问题,本文提出了一种双重注意力残差网络(Dual Attention Residual Network, DARN)模型。DARN网络结构由一个具有全局跳跃连接的残差网络组成,实现底层网络与顶层网络的连接,以稳定深层网络的训练。其中残差网络又包含若干残差组,每个残差组包含若干具有局部跳跃连接的双重注意力块级联的残差块,可以充分利用特征通道之间的相关性,同时将残差的特征集中在关键的空间内容上。网络末端的上采样模块采用亚像素卷积层来重建高分辨率图像。

1 相关工作

1.1 残差学习

经典的卷积层或全连接层在信息传递时,会存在信息丢失、损耗等问题。残差网络^[21](Residual Network, ResNet)在某种程度上解决了这个问题,通过直接将输入信息绕道传到输出,保护信息的完整性。整个网络只需要学习输入、输出差别的那一部分,简化学习目标和难度。残差块(Residual Block, RB)基于前馈卷积网络与一个恒等跳跃连接构成,如图1所示。在形式上,用 $H(x)$ 表示所期望得到的实际映射,堆叠的非线性多层网络去拟合另一个残差映射函数 $F(x)$, x 为第 i 层残差块的输入,实际的映射关系可表示为

$$H(x) = F(x) + x \quad (1)$$

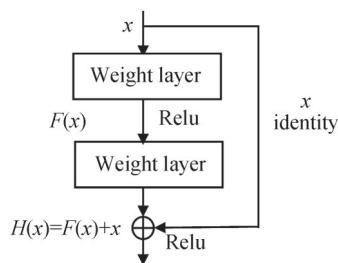


图1 残差块的基本单元

Fig.1 Basic structure of residual block

残差单元通过恒等映射的引入在输入和输出之间建立了一条直接的关联通道,从而使得强大的有参层集中学习输入和输出之间的残差。残差网络更容易优化,并且能够通过增加相当的深度来提高准确率。深度残差网络解决了网络结构增加深度带来的退化问题,能够通过单纯地增加网络深度^[22],来提高网络性能。本文设计的网络结构基于深度残差网络。

1.2 通道注意力机制

通道注意力机制被视为一种引导,使可用的处理资源的分配偏向于输入的最有信息量的部分^[19]。以往基于CNN的图像超分辨率方法平等地对待低分辨率图像通道特征。为了使网络集中于更有信息量的特征,可以利用特征通道之间的相关性,网络结构图如图2所示。通过使用全局平均池化层将输入特征的通道全局空间信息转换成通道描述符。

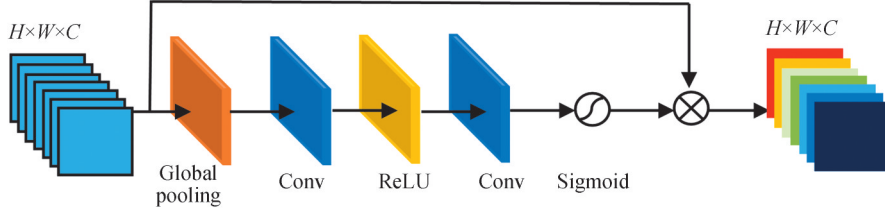


图2 通道注意力机制网络结构

Fig.2 Network structure of channel attention mechanism

在图2网络结构中,设置 $X=[x_1, \dots, x_i, \dots, x_c]$ 为输入,具有大小为 $H \times W$ 的 c 个特征映射,其中Conv表示卷积层的功能, \otimes 表示Hadamard相乘^[23],ReLU与sigmoid分别代表神经网络不同的激活函数^[24]。通过将 X 的空间维度 $H \times W$ 收缩,可以得到通道统计 $z \in \mathbb{R}^c$ 。 z 的第 c 个元素的关系表达式为

$$z_c = H_{GP}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (2)$$

式中, $x_c(i, j)$ 表示第 c 个特征映射 x_c 在位置 (i, j) 处的通道特征值, $H_{GP}(\cdot)$ 表示全局池化功能。这种通道统计看作是局部描述符的集合,其统计有助于表达整个图像。

通过全局平均池化层后从聚合信息中完全捕获通道的关联性,引入sigmoid函数作为简单门控机制。这种机制学习通道间的非线性相互作用并可以强调多通道特性。非线性相互作用关系可表示为

$$s = f(\tau_U \delta(\tau_D)) \quad (3)$$

式中, $f(\cdot)$ 和 $\delta(\cdot)$ 分别表示sigmoid门控和ReLU函数功能。 τ_D 是卷积层的权重,作用是减少通道数目,比率为 r 。经过ReLU函数激活后,低维特征映射再通过卷积层,通道数以比率 r 增加,其权重为 τ_U 。最终获得通道统计量 s ,重新缩放输入 x_c 。通道注意力机制的特征表征通过原输入特征映射 x_c 与第 c 个通道缩放描述符 s_c 逐通道Hadamard相乘,可表示为

$$\hat{x}_c = s_c \cdot x_c \quad (4)$$

考虑到低分辨率图像的输入和特征包含丰富的低频信息,在不同通道中被同等对待,缺乏跨特征通道的区分学习能力,并阻碍了深层网络的表征能力。本文方法基于深度残差网络结构,加入了通道注意力机制。

1.3 空间注意力机制

空间注意力机制在深层网络中广泛存在,但对空间注意力机制的分析相对较少。相关方向主要是通过可视化或分析整个注意力模块在图像处理中的学习注意力权重来进行研究。以往的工作表明,编码器-解码器注意力模块中的注意力权重分配起着类似于传统方法中特征配对的作用。给定一个目标特征和一组关键特征,注意力函数衡量目标特征与关键特征配对的相关性得出注意力权重然后自适应地聚合关键内容。为了使模型能够关注来自不同表示子空间和不同位置的关键内容,多个注意函数的输出被线性地聚集成可学习的权重。假定 q 表示一个目标特征,其内容表示为 Z_q , k 表示一个关键特征,其内容表示为 x_k 。多个注意力函数的输出特征 y_q ,可表示为

$$y_q = \sum_{m=1}^M W_m \left[\sum_{k \in \Omega_q} A_m(q, k, z_q, x_k) \odot W'_m x_k \right] \quad (5)$$

式中, m 表示第 m 个注意力函数, Ω_q 表示目标特征关键区域, $A_m(q, k, Z_q, x_k)$ 表示第 m 个注意力函数中的注意

力权重。 W_m 和 W'_m 表示可学习权重。通常,注意力权重在 Ω_q 内归一化,即 $\sum_{k \in \Omega_q} A_m(q, k, z_q, x_k) = 1$ 。

设计空间注意力网络结构^[20]时,考虑到以下因素:1)注意力块必须足够轻,因为它将被插入到网络的每个残差模块中;2)注意力块需要一个大的感受野才能很好地完成图像超分辨率的任务。基于此,设计注意力机制网络如图3所示,输入的图像特征映射通过一个卷积核大小为 1×1 的卷积层,降低通道维度,整个残差块结构轻量化。同时,使用跳过连接将空间降维前的高分辨率特征直接连接到空间注意力块的末端。然后输入特征通过一个步长为2的卷积层,扩大感受野。使用最大池化层将输入特征的全局空间信息变换成空间描述符。常规的卷积核大小为 2×2 的最大池化层带来的感受野扩大有限,引入了具有更大窗口的卷积核大小为 7×7 的卷积层和步长为3的最大池化层。通过增加上采样层恢复输入特征的空间维度,并且使用大小为 1×1 的卷积核来恢复通道维度。最后,引入sigmoid函数作为简单门控机制。空间注意力机制的特征表征通过原输入特征映射与空间缩放描述符对应Hadamard相乘。

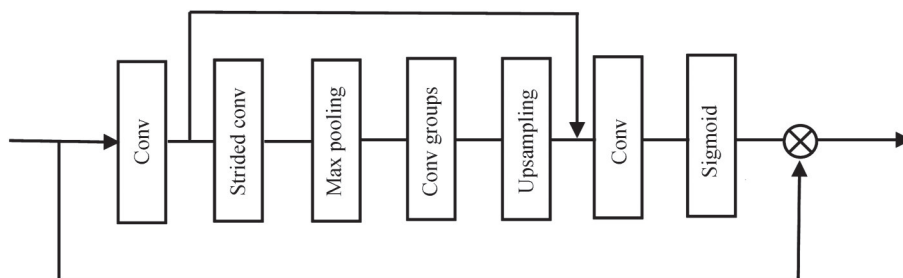


图3 空间注意力机制网络结构
Fig.3 Spatial attention mechanism network structure

2 本文算法

DARN方法主要由浅层特征提取模块、残差中深层特征提取模块、特征映射模块和重建模块四个部分组成。

2.1 网络结构

图4所示为设计的深度残差双重注意力网络基本结构,用 I_{LR} 和 I_{SR} 分别表示DARN网络的输入和输出的偏振图像。从 I_{LR} 的输入中提取浅层特征 F_0 ,使用一个卷积核大小为 3×3 的卷积层完成浅层特征获取。该过程可表示为

$$F_0 = H(I_{LR}) \quad (6)$$

式中, $H(\cdot)$ 表示卷积层(Conv)的浅层特征提取功能。

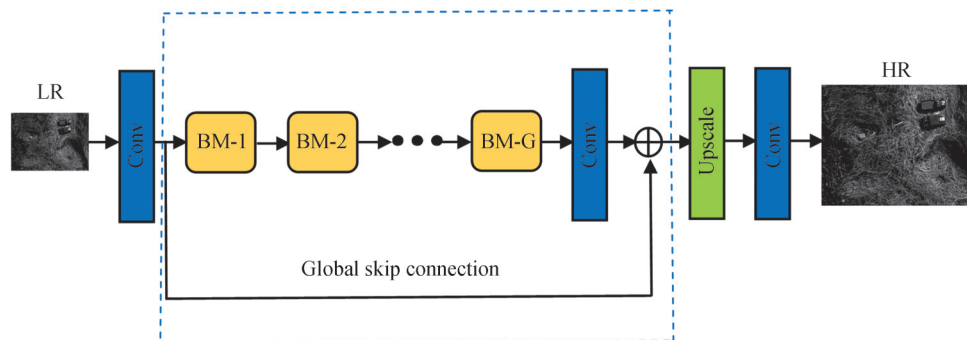


图4 深度残差双重注意力网络基本结构
Fig.4 Network architecture of dual attention deep residual network

将提取的特征 F_0 发送到主干模块进行残差中深度特征学习。主干模块由 G 个基础模块(Base Modules, BM)的残差组网络构成。其中深度特征 F_1 的学习过程可表示为

$$F_1 = B_R(F_0) \quad (7)$$

式中, B_R 表示网络结构中全局残差块包含的非常深的残差组功能。它包含基础模块 BM 和全局跳跃连接 (Global Skip Connection, GSC)。如图 5 所示, 每个基础模块 BM 还包含局部跳跃连接 (Local Skip Connection, LSC) 和 B 个通道注意力机制 (Channel Attention Mechanism, CA) 与空间注意力机制 (Spatial Attention Mechanism, SA) 级联的残差块 (Dual Attention Residual Block, DARB)。提取的深度特征经过一个卷积核大小为 3×3 的卷积层 (Conv) 再通过残差结构与浅层特征 F_0 逐像素相加。此时非线性特征映射 F_2 可表示为

$$F_2 = H(F_1) \oplus F_0 \quad (8)$$

式中, $H(\cdot)$ 表示卷积层 (Conv) 的特征提取功能, \oplus 表示特征逐像素相加。

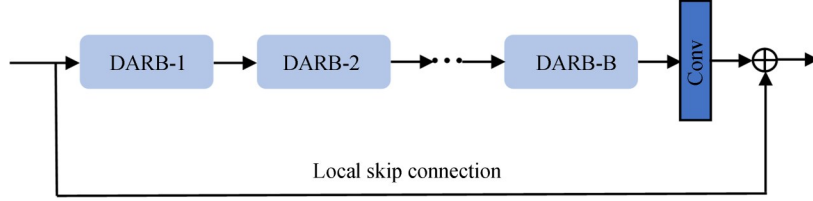


图5 基础模块的网络结构
Fig.5 Network structure of base modules

上采样部分通常可选择反卷积 (也称转置卷积)、最近邻上采样卷积和亚像素卷积^[25] (Efficient Subpixel Convolutional Neural Network, ESPCN) 的方法。本文顶层网络使用亚像素卷积, 这种网络末端上采样方法已经被证明在计算复杂度方面更有效, 并且比前端上采样的 SR 方法获得更高的性能。最后, 通过一个卷积核大小为 3×3 的卷积层 (Conv) 重建放大的特征图实现图像的超分辨率。重建过程可表示为

$$I_{SR} = H_{RE}(F'_2) = f(I_{LR}) \quad (9)$$

式中, I_{SR} 表示超分辨率图像, $H_{RE}(\cdot)$ 表示重构函数, $f(\cdot)$ 表示超分辨率网络的函数。 F'_2 表示 F_2 经过上采样后的输出特征。

2.2 残差中深层特征提取

在图像超分辨率中, 通过堆叠 RB 构建深层网络会遭受训练困难, 并且很难获得更多的性能增益。受 EDSR^[15] 的启发, 使用残差组作为更深层次网络的基本模块, 采用全局残差学习来减轻训练难度。设第 t 个基础模块 BM 的输出深度特征为 F_t , 则 F_t 可表示为

$$F_t = B_t(F_{t-1}) = B_t(B_{t-1}(\dots(B_0(F_0))\dots)) \quad (10)$$

式中, B_t 表示网络结构中第 t 个基础模块 BM 的残差组函数, F_t 与 F_{t-1} 分别表示第 t 个基础模块 BM 的输入特征映射和输出特征映射, F_0 表示浅层特征。残差块堆叠在一起, 形成 SR 网络的主干部分 (如图 5)。在图像重建任务中, 残差块可以产生一些有用的分层特征, 这些特征集中在原始图像的不同方面。第一个残差块的特征必须经过很长的路径并且重复加法和卷积运算才能到达最后一个模块。因此, 残差特征很难被充分利用, 并且在整个网络的学习过程中起着非常局部的作用。为此, 引入跳跃连接来实现底层网络与顶层网络的连接, 提高深层网络训练的稳定性, 实现更好的性能。

若干个基础模块 BM 和跳跃连接允许网络的主干部分专注于 LR 特征更具信息性的组成部分。每个基础模块 BM 包含了 LSC 和 CA 与 SA 级联的残差块 DARB。如图 6 所示, CA 提取通道间的通道统计量, 使可用的处理资源的分配偏向于输入的最有信息量的部分。为了最大化深度残差通道注意力网络框架的有效性, 可与 SA 结合使用。该结构将残差的特征集中在关键的空间内容上, 从而获得更具代表性的特征, 如图 7 所示。第 t 个基础模块 BM 的第 b 个 DARB 输出深度特征 $F_{t,b}$ 可表示为

$$F_{t,b} = F_{t,b-1} + \alpha_{t,b} [\eta_{t,b}(X_{t,b}) \cdot X_{t,b}] \quad (11)$$

式中, $\eta_{t,b}(\cdot)$ 表示通道注意力机制的功能函数, $\alpha_{t,b}(\cdot)$ 表示空间注意力机制的功能函数, $F_{t,b-1}$ 与 $F_{t,b}$ 分别表示残差块 DARB 的输入特征与输出特征, $X_{t,b}$ 表示从输入特征 $F_{t,b-1}$ 中学习的残差特征。

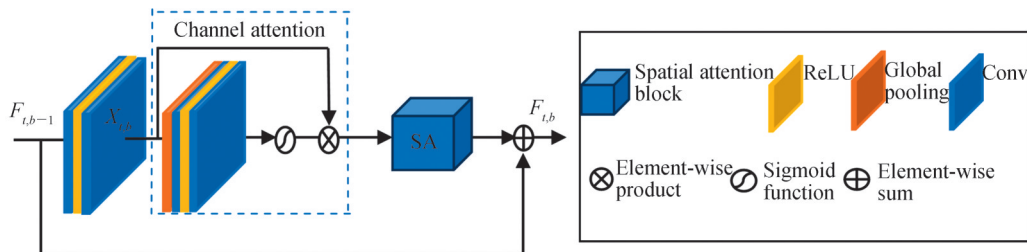


图6 双重注意力机制残差块的网络结构
Fig.6 Network structure of dual attention residual block

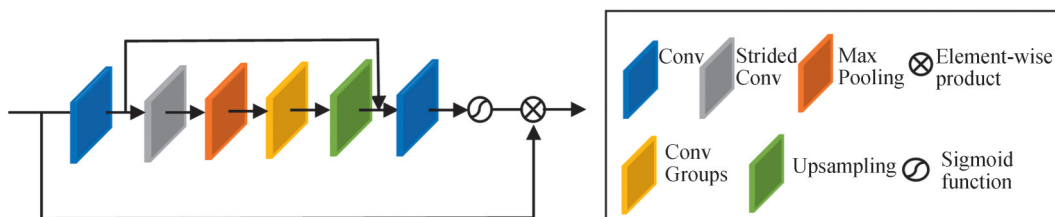


图7 空间注意力的网络结构
Fig.7 Network structure of spatial attention block

基于双重注意力残差块(DARB)和全局残差学习的结构为高精度图像SR构建一个DARN深层网络,并实现性能改进。

2.3 损失函数

常见的损失函数包括 L_2 、 L_1 、感知损失和对抗损失。为了验证深度残差网络DARN的有效性,选择与之前残差网络算法相同的 L_1 损失函数对DARN网络模型进行优化。设 x 为输入LR图像, θ 为待优化的网络参数集。目标是学习一个映射函数 $f(\cdot)$, 用于生成高分辨率图像 $\hat{y} = f(x; \theta)$ 。假设给定一组训练集, 首先通过双三次插值下采样来缩放高清图像HR, 研究每个不同下采样级别的数据集HR-LR配对的映射关系。使用稳定的 L_1 损失函数来处理损失值, 而不是最小化 y 和 \hat{y} 之间的均方误差。训练DARN网络的目标是最小化 L 损失函数, 损失函数可表示为

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \| f(I_{LR}^i) - I_{HR}^i \|_1 \quad (12)$$

式中, $f(\cdot)$ 表示超分辨率重建网络的函数, θ 表示提出网络的参数集, I_{LR}^i 和 I_{HR}^i 分别表示训练数据中第 i 个低分辨率图像及与其对应的真实高分辨率图像, $\| \cdot \|_1$ 为1范数, 采用随机梯度下降法优化损失函数。

3 实验结果与分析

3.1 实验环境与数据

实验环境为: 训练与测试图像集采用处理器为 Intel(R) Core(TM) i9-10900X CPU @ 3.70GHz, 系统运行内存 128 GB, GPU 显卡为 24 GB 显存容量的 NVIDIA Quadro RTX 6000 的服务器。训练软件运行环境为 Python3.6, 编程框架为 Pytorch 0.4.1, 搭建 Cuda11.2 用于实验加速。测试评估样本的软件为 Matlab R2019。

偏振图像由自行研制的一套基于 Savart 偏光镜的空间调制型红外与可见光双通道偏振相机, 实现对目标给定谱段下完整偏振状态的快速测量采集, 如图8所示。空间调制器模块由一对 Savart 偏光镜、一块半波片和一块偏振片组成。该成像系统采用 Stokes 矢量的空间调制原理^[9], 同时将 4 个 Stokes 矢量 ($S_0 \sim S_3$) 调制在同一幅图像中, 一次采集即可得到包含目标 4 个 Stokes 矢量的调制信息, 据此可以解析出多个偏振参量图像。本实验采用 470 幅自建的偏振图像数据集以及公开数据集 DIV2K (800 幅) 进行混合训练实验。再采用多组场景下采集的偏振图像作为测试数据集。同时对训练图像随机旋转 90° 、 180° 、 270° , 并水平翻转进行数据增强。实验主要包括两个部分: 1) 使用双三次退化模型对偏振相机采集的高清训练样本集进行不同倍

数的下采样处理并添加噪声和模糊,得到与之对应的低分辨率图像,完成网络模型的训练,并与其他方法进行测试对比以验证算法的有效性;2)将本文方法重建的图像与偏振成像系统采集的高分辨率同一场景图像进行比较,为系统定标校正提供数据参考。

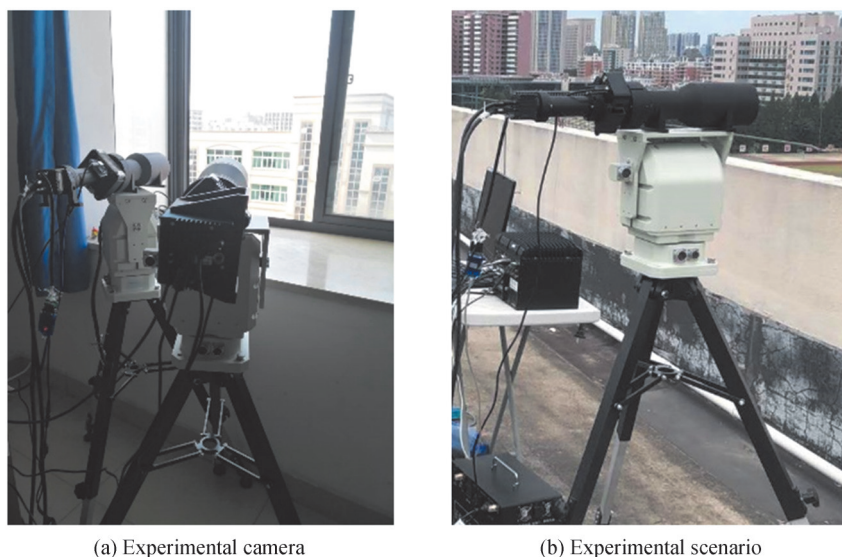


图8 光谱偏振相机

Fig.8 Spectral polarization camera

3.2 评价指标与实验参数设置

实验通过定性分析、定量分析、与成像系统对比校正三个方面进行评价。评价标准采用峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)和结构相似性(Structural Similarity, SSIM)在变换的YCbCr空间^[26]的Y通道(亮度分量)上评估不同方法的超分辨率效果以及信息熵、清晰度指标来评价图像的质量。

网络结构中,基础模块BM的数量 G 设置为10。每个BM模块中,残差块DARB的数量 B 设置为20,通道数设置为64。通道降维度和升维度卷积层的卷积核大小为 1×1 ,其余所有卷积层的卷积核大小为 3×3 。网络每个批次训练时,提取16张大小为 48×48 的LR样本作为输入,每600个批次进行一次测试。本文算法的模型由Adam optimizer^[27]训练,其中一阶矩估计的指数衰减速率 $\beta_1 = 0.9$,二阶矩估计的指数衰减速率 $\beta_2 = 0.999$,一个很小的常数 $\epsilon = 10^{-8}$,网络共迭代训练600 epoch。初始学习率为 10^{-4} ,每训练200 epoch以后(即,迭代 1.2×10^5 次)反向传播迭代学习率衰减一半。

3.3 实验结果分析

为测试本文方法的性能,选取一组植绒布图像,如图9所示。实验使用PSNR和SSIM评价图像的重建效果,PSNR是基于对应像素点间的误差,其值越大说明重建质量越好。SSIM分别从亮度、对比度、结构三

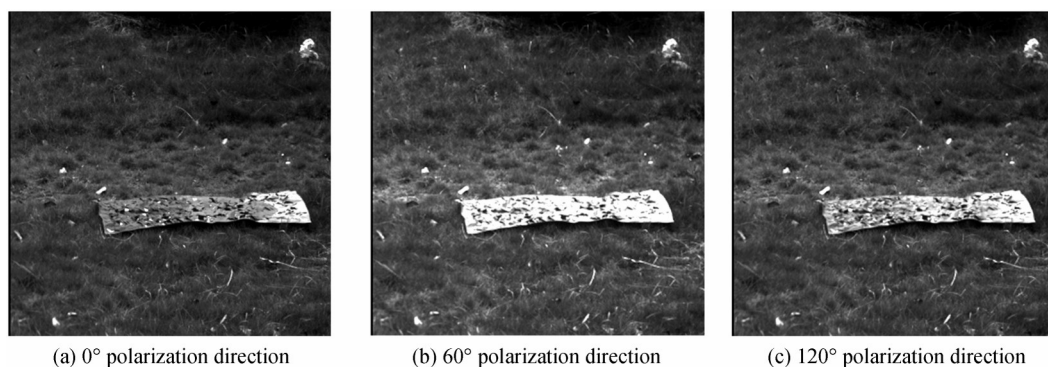


图9 不同偏振方向植绒布图像

Fig.9 Different polarization direction of fabric image

方面度量图像相似性,其值越高说明重建图像与原始高清图像越接近。将本文方法与Bicubic、SRCNN、FSRCNN、EDSR方法进行对照实验,分别在缩放因子为2、3、4时进行重建,得到PSNR值和SSIM值如表1所示。

表1 植绒布偏振图像在不同算法下的指标对比
Table 1 Comparison of indicators of fabric image polarization by different algorithms

Scale	Method	Params	0°		60°		120°	
			PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
2	Bicubic	—	37.626 0	0.939 3	34.775 4	0.925 4	36.116 4	0.929 1
	SRCNN	57K	39.808 2	0.959 0	37.164 2	0.950 7	38.393 5	0.952 4
	FSRCNN	12K	39.858 2	0.959 9	37.520 4	0.952 4	38.583 1	0.953 8
	EDSR	40.7M	42.277 4	0.968 6	39.892 0	0.962 2	40.974 9	0.963 6
	Ours	17.7M	42.294 5	0.968 8	39.919 2	0.962 6	40.988 0	0.963 9
3	Bicubic	—	34.056 6	0.876 1	31.197 2	0.846 1	32.629 7	0.857 4
	SRCNN	57K	35.727 5	0.902 9	32.984 9	0.881 0	34.356 1	0.888 8
	FSRCNN	12K	35.572 2	0.902 7	33.047 5	0.881 8	34.280 4	0.889 0
	EDSR	43.6M	37.783 2	0.923 6	35.145 5	0.905 0	36.403 2	0.911 5
	Ours	17.9M	37.834 8	0.924 6	35.245 3	0.906 3	36.494 2	0.912 8
4	Bicubic	—	31.966 8	0.824 3	29.106 6	0.780 0	30.555 3	0.798 3
	SRCNN	57K	33.382 4	0.848 7	30.478 2	0.811 4	31.910 6	0.826 3
	FSRCNN	12K	33.471 3	0.849 3	30.665 4	0.813 7	32.031 8	0.827 7
	EDSR	43M	35.272 3	0.881 6	32.495 8	0.850 5	33.794 3	0.862 4
	Ours	17.9M	35.356 8	0.882 5	32.563 5	0.851 9	33.863 4	0.863 5

从表1的实验数据可直观看出,本文方法的客观评价指标显著优于其他方法。本文方法的参数量明显比EDSR方法少,同时实现了更高的性能,在模型大小和性能之间取得了更好的权衡。实验结果还表明,较深的网络可能比较宽的网络更容易获得更好的性能。在缩放因子为2时,本文方法的PSNR值相对Bicubic、SRCNN、FSRCNN、EDSR方法平均分别提升4.894 dB、2.611 dB、2.413 dB、0.019 dB,SSIM值平均分别提升0.033、0.011、0.009、0.000 3;在缩放因子为3时,PSNR值平均分别提升3.896 dB、2.168 dB、2.224 dB、0.080 dB,SSIM值平均分别提升0.054、0.023、0.023、0.001;在缩放因子为4时,PSNR值平均分别提升3.385 dB、2.004 dB、1.871 dB、0.073 dB,SSIM值平均分别提升0.065、0.037、0.035、0.001。

为进一步验证本文方法重建图像的效果,从采集的偏振图像中选择两张细节丰富、纹理清晰的图像进行测试。从各超分辨率算法放大2倍后的重建图像细节展示,直观看出本文方法重建图像具有较好的视觉效果,如图10、11所示。由图中标记区域的字母“UN”和飞机可以看出,Bicubic方法重建的图像出现模糊、振铃效应、纹理不清晰等问题,超分辨率效果有限;基于神经网络的方法SRCNN、FSRCNN重建的图像视觉上比较平滑,图像锐度有所提升,但背景明显被虚化,仍存在边缘模糊、分辨率低等问题;深度网络EDSR重建图像进一步提高图像锐度、改善视觉效果,但图像标记区域出现了模糊伪影、图像边缘模糊等问题;本文算法重建图像效果明显优于其他算法重建图像,纹理细节更加清晰,图10(f)中的字母“UN”和图11(f)中的飞机机翼尾灯,都取得较好的效果,亮度均匀,接近原始的高清图像。本文方法的客观评价指标PSNR值与SSIM值也均优于其他方法。实验数据充分说明本文方法相比其他算法客观评价指标有所提高,视觉效果较好。

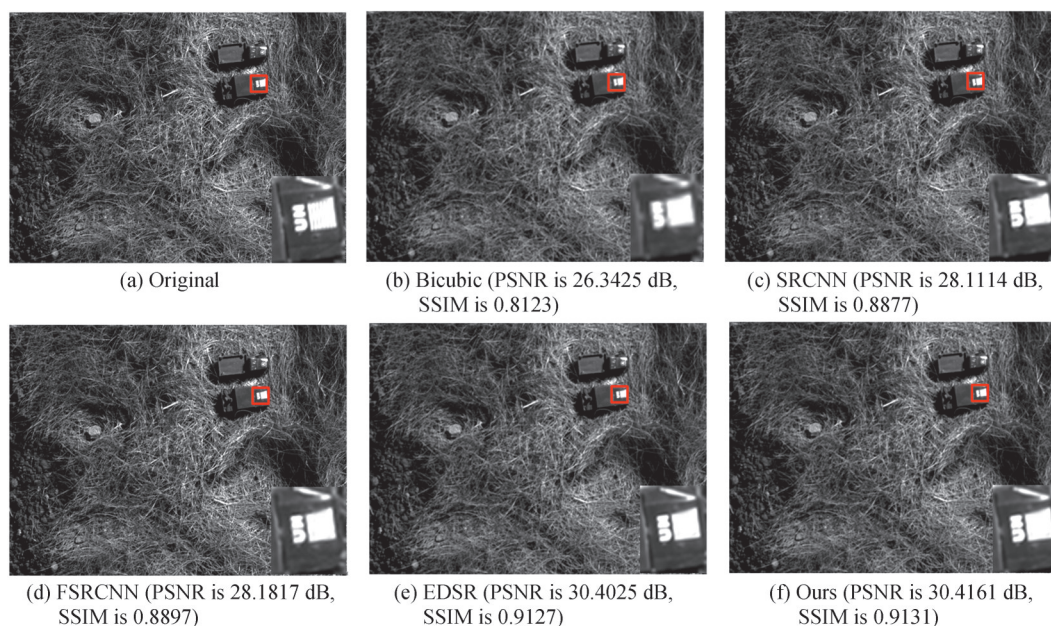


图10 卡车缩放模型的不同方法重建效果

Fig.10 Reconstruction effect of truck contraction model by different methods

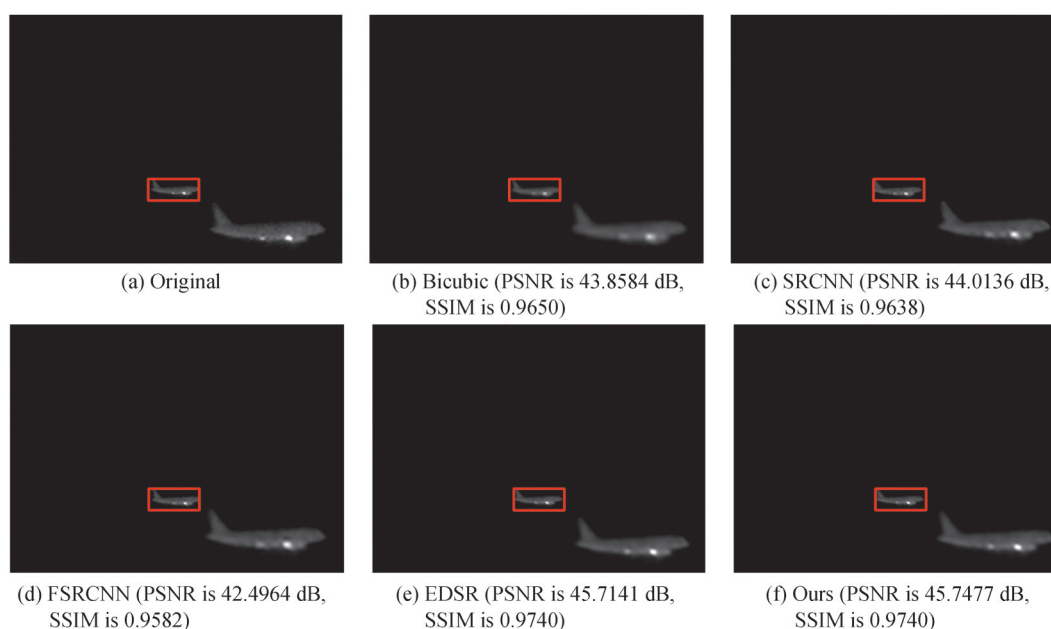


图11 飞机的不同方法重建效果

Fig.11 Reconstruction effect of airplane image by different methods

3.4 全偏振图像解析

全偏振成像系统通过调制解析获取探测物体的光谱特征、特定空间特征以及反映目标材质的相关本征偏振信息。它从不同角度反映出场景以及探测目标信息、目标全偏振参量实现同步探测。为了进一步体现本文算法的有效性并为系统定标校正,选取清晰度高的石碑图像通过调制解析得到原始的全偏振图像解析结果作为实验的测试样本。然后随机选取若干张纹理丰富、细节复杂的全偏振解析图像通过双三次插值下采样生成低分辨率解析图像,再用训练好的网络模型进行图像超分辨率重建。实验使用PSNR和SSIM评价图像的重建效果。

随机选取了直升机对地观测下的5张细节、纹理清晰的合肥渡江战役纪念碑的全偏振图像解析图构成测试图像集,如图12所示,分别为V图、 E_y 图、 S_0 图、Q图、U图各一张。将本文方法与Bicubic、SRCNN、

FSRCNN、EDSR方法进行对比实验,各网络模型在 $\times 2$ 的缩放因子下,得到重建图像的PSNR值和SSIM值如表2所示。

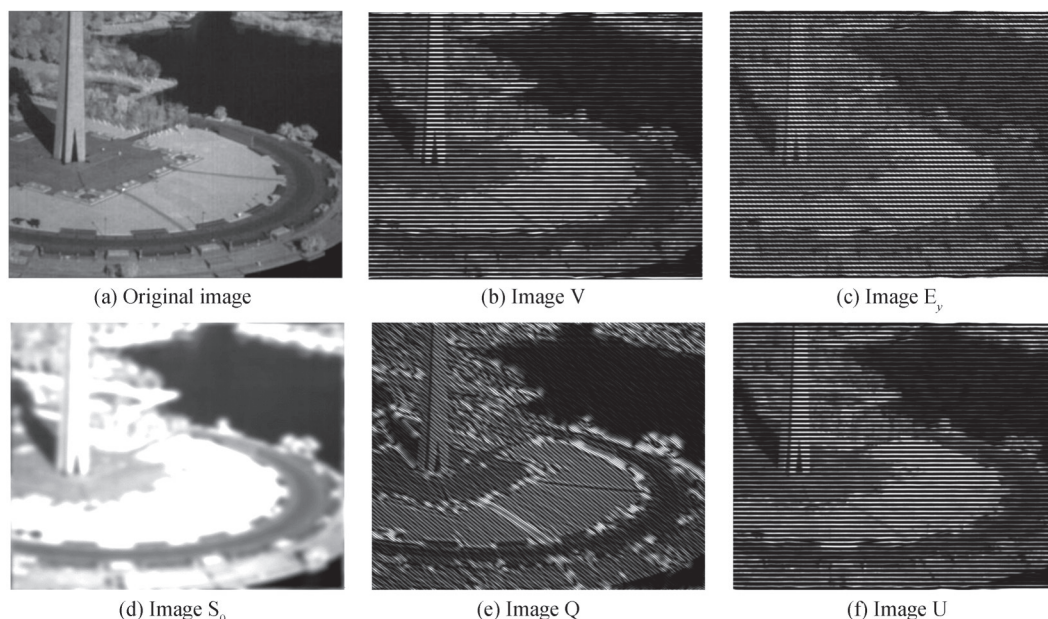


图12 纪念碑全偏振图像解析
Fig.12 Analysis of monument image full polarization

表2 纪念碑全偏振图像解析图在不同算法下的指标对比

Table 2 Comparison of indicators on the analysis of monument image full polarization by different algorithms

Method	Time/s	Image V		Image E_y		Image S_0		Image Q		Image U	
		PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM	PSNR/dB	SSIM
Bicubic	0.05	22.406 9	0.907 0	22.043 2	0.909 2	52.669 6	0.997 5	20.953 3	0.783 8	22.587 6	0.909 6
SRCNN	0.24	27.944 3	0.963 6	26.216 7	0.965 9	51.989 4	0.997 6	23.089 1	0.870 9	28.221 5	0.964 3
FSRCNN	0.25	27.567 7	0.960 4	25.850 3	0.958 8	51.167 9	0.997 3	24.194 1	0.903 8	27.786 7	0.961 9
EDSR	0.49	32.761 5	0.990 8	29.315 7	0.982 1	53.533 4	0.997 6	29.268 5	0.958 5	32.860 4	0.990 9
Ours	0.48	35.331 9	0.993 9	30.940 4	0.987 1	53.538 1	0.997 6	29.141 3	0.956 1	35.052 7	0.993 7

从表2的实验数据可以看出,本文方法的客观评价指标显著优于其他方法。在图像测试耗时方面,Bicubic方法较快,其他基于学习的方法耗时均比Bicubic方法高。本文方法在测试时间和性能之间进行了更好的权衡,牺牲少量时间,改善了重建图像质量且测试时间在0.5 s之内。本文方法的PSNR值相对Bicubic、SRCNN、FSRCNN、EDSR方法平均分别提升8.668 dB、5.308 dB、5.487 dB、1.253 dB,SSIM值平均分别提升0.084、0.033、0.029、0.001。由表1可知,本文方法的参数数量明显比EDSR方法少,同时EDSR网络的通道数为本文方法的4倍,提取更多特征信息,仅个别情况下与本文方法性能上相当。本文方法在合肥渡江战役纪念碑全偏振图像解析结果图上进行SR重建的PSNR值和SSIM值明显优于对比方法。

为进一步验证本文方法重建图像的效果,从测试的全偏振图像解析图中选择一张细节丰富、纹理清晰的V图进行测试。从各超分辨率算法在2倍的缩放因子下的重建图像细节展示,可直观看出本文方法重建图像具有较好的视觉效果,如图13所示。图13标记区域的条纹可以看出,Bicubic方法重建的图像出现模糊、背景明显被虚化、条纹不清晰等问题,重建图像效果有限;基于卷积网络的方法SRCNN、FSRCNN重建的图像视觉上比较平滑,图像锐度有所提升,但仍存在边缘模糊,分辨率低等问题;深度网络EDSR重建图像进一步提高图像锐度、改善视觉效果,但图像标记区域出现了亮度不均匀、图像边缘模糊等问题;本文算法重建图像效果明显优于其他算法重建图像,纹理细节更加清晰,亮度均匀,较为接近原始的高清图像。本文方法的客观评价指标PSNR值与SSIM值也均优于其他方法,充分验证了所提方法的优越性。

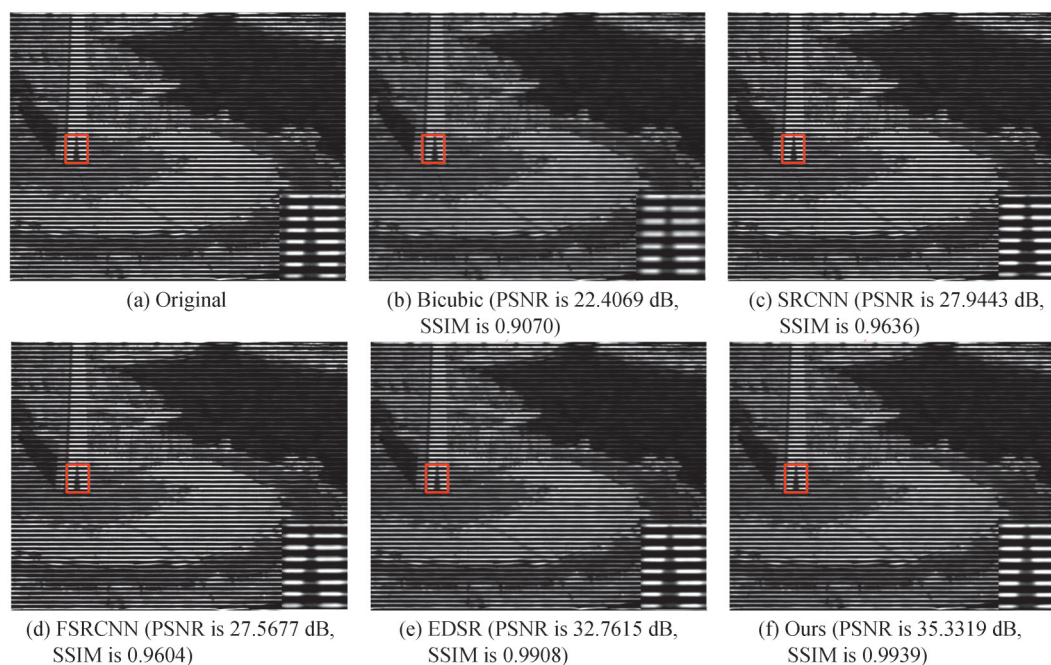


图 13 V 图的不同方法重建效果

Fig.13 Reconstruction effect of image V by different methods

3.5 系统对比较正

研制的光谱偏振相机包含两个通道,一个为可见光近红外光谱偏振相机,另一个为短波红外光谱偏振相机。偏振成像系统^[7]根据探测目标以及图像处理的要求,可实现成像空间分辨率的调节。如分辨率为 2048×2048 的可见光偏振图像与 640×512 的短波红外偏振图像,设置系统可分别采集 HR 图像和 LR (1024×1024 和 320×256)图像。为了对比本文方法与成像系统的图像效果,将 LR 图像进行 SR 重建。然后与系统采集的 HR 图像进行评价。实验采用信息熵、清晰度来无参考评价图像的质量。信息熵表示图像包含平均信息量的多少,熵值越大则所含信息量越多。清晰度表示图像中微小细节反差与纹理变化特征,值越大则图像清晰程度越高。

表 3 比较了本文方法 SR 重建图像与系统采集的 HR 和 LR 图像的无参考评价指标。从表 3 可以看出,在信息熵指标上,SR 重建结果普遍略低于系统 HR 图像,但高于系统 LR 图像。对于清晰度指标,SR 重建结果基本与系统 HR 图像质量保持一致,但相较于 LR 图像清晰度提高。图 14 所示为系统采集的 HR 和 LR 图像与 SR 重建图像效果对比,可以看出相机采集的 LR 图像出现“凹凸条纹”模糊、背景明显被虚化、纹理不清晰等问题,而 2 倍 SR 重建的图像“凹凸条纹”纹理细节清晰、亮度均匀、接近系统采集的 HR 图像,相较于 LR 图像质量有明显的改善。不同倍数超分辨率重构下的信息熵、清晰度与超像素倍数的关系曲线如图 15 所示。从图 15 的折线图可直观看出,随着超分辨率倍数的增加,对于信息熵指标,系统 HR 图像曲线趋于平缓持平,SR 重建图像曲线呈现下降趋势但与系统 HR 图像曲线保持较近的距离。对于清晰度指标,系统 HR 图像曲线与 SR 重建图像曲线趋于平缓基本持平,数据波动小。实验数据充分说明本文方法的有效性,重建图像较为接近系统成像的效果。

表 3 伪装网偏振图像无参考评价结果

Table 3 No reference comparison of camouflage net polarization images results

Scale	Information entropy			Definition		
	LR	HR	SR	LR	HR	SR
×2	6.83	6.87	6.85	0.91	0.93	0.93
	6.57	6.61	6.59	0.91	0.93	0.94
	6.72	6.76	6.74	0.91	0.93	0.94

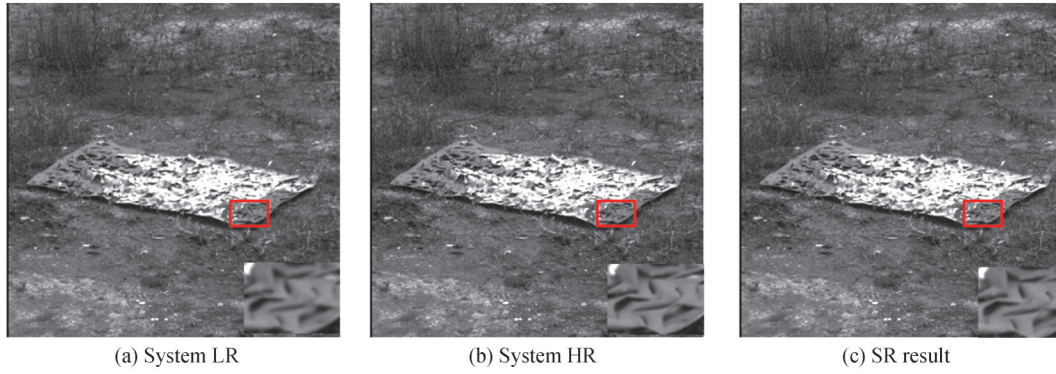


图 14 伪装网偏振图像重建结果与系统采集结果对比

Fig.14 Comparison between polarization image reconstruction results of camouflage net and the system-acquired results

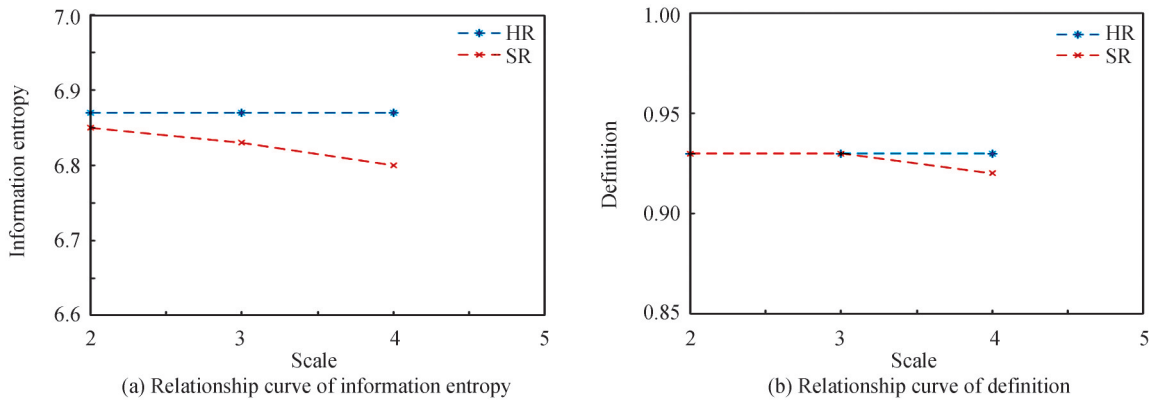


图 15 无参考评价指标与超像元倍数关系曲线

Fig.15 Curve of relationship between no reference evaluation index and supercell multiple

4 结论

本文提出了一种基于双重注意力残差网络DARN的偏振图像超分辨率重建方法。通过实验在主观视觉效果、客观评价指标与成像系统对比校正三个方面对算法进行评价。所提方法的重建图像纹理细节更加丰富、清晰,较为接近系统成像的效果,充分说明该方法实现了更好的重建准确性和视觉改善效果,能够很好地抑制输入偏振图像中的平滑区域,生成更加高质量的偏振图像,为偏振图像的超分辨率以及全偏振相机系统校正的理论研究提供了一种数据支撑。本文主要考虑多通道与空间特征的提取,网络深度与网络规模有待改善。下一阶段的主要任务是减少网络模型参数量,降低网络深度的同时尽可能地优化训练网络的效率以及进一步提高重建图像的质量,在此基础上考虑融入多个残差块聚合网络,简化网络模型。

参考文献

- [1] WANG Xiaolong, WANG Feng, LIU Xiao, et al. Hyperspectral polarization characteristics of typical camouflage target under desert background[J]. Laser&Optoelectronics Progress, 2018, 55(5): 051101.
王小龙, 王峰, 刘晓, 等. 荒漠背景下典型伪装目标的高光谱偏振特性[J]. 激光与电子学进展, 2018, 55(5): 051101.
- [2] ZHAO Y, YI C, KONG S G, et al. Multi-band polarization imaging and applications[M]. Berlin: Springer Publishing Company, Incorporated, 2016: 47-71.
- [3] YE Song, YAN Haofang, SUN Xiaobing, et al. Modified image demodulation algorithm for spatially modulated full polarization imaging systems[J]. Acta Optica Sinica, 2019, 39(6): 0607001.
叶松, 严浩方, 孙晓兵, 等. 空间调制型全偏振成像系统的图像解调算法修正[J]. 光学学报, 2019, 39(6): 0607001.
- [4] CAO Qizhi, ZHANG Jing, DEHOOG E, et al. Static subminiature snapshot imaging polarimeter using spatial modulation [J]. Acta Physica Sinica, 2016, 65(5): 31-42.
曹奇志, 张晶, DEHOOG E, 等. 空间调制稳态微型快拍成像测偏技术研究[J]. 物理学报, 2016, 65(5): 31-42.
- [5] YANG Min, XU Wenbin, TIAN Yuze, et al. Time-sharing infrared polarization imaging system for moving target detection[J]. Acta Optica Sinica, 2020, 40(15): 1511001.

- 杨敏,徐文斌,田禹泽,等.面向运动目标探测的分时型红外偏振成像系统[J].光学学报,2020,40(15):1511001.
- [6] WU Yan, CHEN Fansheng, CHEN Guilin, Applicable conditions and improving spatial resolution upper limits of image restoration and super-resolution reconstruction[J]. Journal of Infrared and Millimeter Waves, 2010, 29(5): 351-356.
吴艳,陈凡胜,陈桂林.图像复原与超分辨率重构基本适用条件及提高空间分辨率上限的研究[J].红外与毫米波学报,2010,29(5):351-356.
- [7] SHAO Xiaopeng, SU Yun, LIU Jinpeng, et al. Connotation and system of computational imaging (Invited)[J]. Acta Photonica Sinica, 2021, 50(5): 0511001.
邵晓鹏,苏云,刘金鹏,等.计算成像内涵与体系(特邀)[J].光子学报,2021,50(5):0511001.
- [8] TANG Yanqiu, PAN Hong, ZHU Yaping, et al. A survey of image super-resolution reconstruction[J]. Acta Electronica Sinica, 2020, 48(7): 1407-1420.
唐艳秋,潘泓,朱亚平,等.图像超分辨率重建研究综述[J].电子学报,2020,48(7):1407-1420.
- [9] XU Guoming, YUAN Hongwu, XUE Mogen, et al. Spatial modulation full polarization computing imaging super-resolution via scene transfer[J]. Journal of Computer-Aided Design&Computer Graphics, 2021, 33(9): 1440-1449.
徐国明,袁宏武,薛模根,等.空间调制全偏振计算成像场景迁移超分辨率方法[J].计算机辅助设计与图形学学报,2021,33(9):1440-1449.
- [10] MENG Xiangchao, SUN Weiwei, REN Kai, et al. Spatial-spectral fusion of GF-5/GF-1 remote sensing images based on multi-resolution analysis[J]. Journal of Remote Sensing, 2020, 24(4): 379-387.
孟祥超,孙伟伟,任凯,等.基于多分辨率分析的GF-5和GF-1遥感影像空-谱融合[J].遥感学报,2020,24(4):379-387.
- [11] HU X C, MU H Y, ZHANG X Y, et al. Meta-SR: a magnification-arbitrary network for super-resolution [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos: IEEE Computer Society Press, 2019: 1575-1584.
- [12] SOH J W, CHO S W, CHO N I. Meta-transfer learning for zero-shot super-resolution[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 3513-3522.
- [13] DONG C, LOY C C, HE K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [14] DONG C, LOY C C, TANG X O. Accelerating the super-resolution convolutional neural network[C]. Proceedings of the 14th European Conference on Computer Vision (ECCV), Amsterdam, the Netherlands: Springer, 2016: 391-407.
- [15] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017: 1132-1140.
- [16] GAO P, ZHANG Q, WANG F, et al. Learning reinforced attentional representation for end-to-end visual tracking[J]. Information Sciences, 2020, 517: 52-67.
- [17] ZHANG Y, LI K, LI K, et al. Image super-resolution using very deep residual channel attention networks [C]. Proceedings of the European conference on computer vision (ECCV), Cham: Springer International Publishing, 2018: 286-301.
- [18] DAI T, CAI J, ZHANG Y, et al. Second-order attention network for single image super-resolution[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Los Alamitos: IEEE Computer Society Press, 2019: 11065-11074.
- [19] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 7132-7141.
- [20] LIU J, ZHANG W, TANG Y, et al. Residual feature aggregation network for image super-resolution[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 2356-2365.
- [21] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
- [22] DU B, WANG Z M, ZHANG L F, et al. Robust and discriminative labeling for multi-label active learning based on maximum correntropy criterion[J]. IEEE Transactions on Image Processing, 2017, 26(4): 1694-1707.
- [23] HIAI F, LIN M H. On an eigenvalue inequality involving the Hadamard product[J]. Linear Algebra and its Applications, 2017, 515: 313-320.
- [24] CLEVERT D A, UNTERTHINER T, HOCHREITER S. Fast and accurate deep network learning by exponential linear units (ELUs)[J/OL]. (2016-02-22)[2021-08-05]. <https://arxiv.org/abs/1511.07289>.
- [25] SHI W Z, CABALLERO J, HUSZAR F, et al. Real-time single image and video super-resolution using an efficient subpixel convolutional neural network [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 1874-1883.
- [26] POMI A, SLUSALLEK P. Interactive ray tracing for virtual TV studio applications[J]. Journal of Virtual Reality and Broadcasting, 2005, 2(1): 1-10.

- [27] SHEN H. Towards a mathematical understanding of the difficulty in learning with feedforward neural networks [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 811–820.

Polarization Image Super-resolution Reconstruction Based on Dual Attention Residual Network

XU Guoming^{1,2,3}, WANG Jie¹, MA Jian^{1,2}, WANG Yong³, LIU Jiaqing¹, LI Yi⁴

(1 School of Internet, Anhui University, Hefei 230039, China)

(2 National Engineering Research Center for Agro-Ecological Big Data Analysis & Application, Anhui University, Hefei 230601, China)

(3 Anhui Province Key Laboratory of Polarized Imaging Detecting Technology, Army Artillery and Air Defense Forces Academy of PLA, Hefei 230031, China)

(4 Institute of Intelligent Technology, Anhui Wenda University of Information Engineering, Hefei 231201, China)

Abstract: In polarization imaging detection, the difference and change of physical properties of aerosols or detection targets are reflected by polarization characteristics. The high-dimensional polarization characteristics effectively improve the contrast between the target and the background, thereby laying the foundation for realizing the inversion of the target's spatial structure. This feature can enhance the recognition effect of the target in the cluttered background. Affected by the imaging distance and atmospheric interference, the limit resolution of the image projected on the focal plane is greatly reduced (much smaller than the optical system diffraction limit resolution), resulting in a lower spatial resolution of the polarized image. On the other hand, the spatial resolution of the polarized image is limited by the number of detector pixel. High resolution images are of great significance and value to the accuracy of target detection. For this reason, without replacing the hardware imaging system, the super-resolution reconstruction method is usually adopted. This method is a common technical means in image processing and practical engineering applications, and it is also a hot research issue of underlying computer vision. The existing image super-resolution algorithms have some problems, such as low utilization of feature information, large amount of parameters, blurred image reconstruction details and so on. The input features of low-resolution images contain rich low-frequency information, which is treated equally in different channels. In computational imaging process using deep learning, image mapping function solution space is very large, it is difficult to generate detailed texture and high-frequency information lack, which hinders convolutional neural network representation ability in image super-resolution. In order to solve this problem, a depth residual polarization image super-resolution network combined with double attention mechanism is proposed. This paper proposes a dual attention residual network model. The network structure is composed of a residual network with a global jump connection, which realizes the connection between the bottom network and the top network to stabilize the training of the deep network. The residual network contains 10 residual groups, and each residual group contains 20 residual blocks cascaded by dual attention blocks with local skip connections; At the same time, considering the interdependence between channels, an adaptive channel feature adjustment mechanism is designed. The channel attention mechanism is regarded as a guide, which biases the allocation of available processing resources towards the most informative part of the input; Cascaded spatial attention blocks are introduced to focus residuals characteristics on the key spatial contents. Spatial attention function measures the correlation between target features and key features, obtains the attention weight, and then aggregates the key content adaptively. The up-sampling module at the end of the network uses sub-pixel convolutional layers to reconstruct high-resolution images. The experiments mainly consist of two parts: first, use the bicubic degradation model to down-sample the training data set collected by the polarization camera, then add noise and blur to obtain the corresponding low-resolution image, and complete the training of the network model. The proposed method is compared with Bicubic, SRCNN, FSRCNN, EDSR methods to verify the effectiveness of the algorithm. In addition, the image reconstructed by this method is compared with the imaging system to provide data reference for system calibration and correction. The experimental results show that the reconstructed image with this method has richer texture details and uniform brightness,

which is closer to the high-definition image of the imaging system. The peak signal-to-noise ratio and structural similarity index of this method are better than other methods, but the amount of parameters is only about 2/5 of that of EDSR. Thus, this method has the advantages of low information redundancy and better reconstruction effect.

Key words: Computational imaging; Super-resolution; Depth residual network; Polarization images; Dual attention block

OCIS Codes: 100.6640; 100.3010; 100.3020; 110.5405