

## 基于多尺度特征增强的合成孔径光学图像复原\*

张银胜<sup>1)2)</sup> 童俊毅<sup>2)</sup> 陈戈<sup>2)</sup> 单梦姣<sup>2)</sup> 王硕洋<sup>2)</sup> 单慧琳<sup>1)2)†</sup>

1) (无锡学院江苏省集成电路可靠性技术及检测系统工程研究中心, 无锡 214105)

2) (南京信息工程大学电子信息工程学院, 南京 210044)

(2023年11月6日收到; 2024年1月17日收到修改稿)

受物理孔径大小和光线散射等影响, 合成孔径光学系统成像因通光面积不足和相位失真而出现降质模糊. 传统合成孔径光学系统成像复原算法对噪声敏感, 过于依赖退化模型, 自适应性差. 对此提出一种基于生成对抗网络的光学图像复原方法, 采用 U-Net 结构获取图像多级尺度特征, 利用基于自注意力的混合域注意力提高网络在空间、通道上的特征提取能力, 构造多尺度特征融合模块和特征增强模块, 融合不同尺度特征间的信息, 优化了编解码层的信息交互方式, 增强了整体网络对原始图像真实结构的关注度, 避免在复原过程中被振铃现象产生的伪影干扰. 实验结果表明, 与其他现有方法相比, 该方法在峰值信噪比、结构相似性和感知相似性评估指标上分别提高了 1.51%, 4.42% 和 5.22%, 有效解决合成孔径光学系统成像结果模糊退化的问题.

**关键词:** 图像处理, 合成孔径, 多尺度, 特征融合**PACS:** 42.30.Tz, 07.05.Pj, 95.75.Mn**DOI:** 10.7498/aps.73.20231761

## 1 引言

随着高分辨率成像技术在对地观测、环境监测和军事侦察等领域的广泛应用<sup>[1]</sup>, 对成像技术分辨率的要求越来越高. 光学系统成像的分辨率往往与系统孔径大小成正比, 但受光学元件的材料和设计、加工难度等影响, 传统光学系统的物理孔径大小受到限制, 且系统的分辨率通常无法随意改变, 并随着观测距离增大而下降, 灵活性较低. 合成孔径光学成像技术通过一定的空间方式将多个独立的子孔径以一定方式排列, 获得和单一大孔径成像等效的分辨率. 加工难度小, 成本低且更易实现的优点使其广泛应用于各成像系统中<sup>[2]</sup>.

伴随合成孔径光学系统的广泛应用, 其自身问题也逐渐显现, 其中影响最大的是多孔成像带来的降质模糊问题, 原因主要有两方面: 1) 多个子孔

径组合会减小光线通过的有效面积, 引起系统中点扩散函数 (point spread function, PSF)<sup>[3]</sup> 变化, 造成调制传递函数 (modulation transfer function, MTF) 的中低频下降, 最终影响成像的细节和分辨率. 2) 由于光波传播特性, 不同子孔径可能具有细微的相位差异, 多个子孔径组合时, 相位间会产生共相误差<sup>[4]</sup>, 导致图像模糊并产生伪影, 严重时甚至无法成像.

为了获得高分辨率图像, 通常在合成孔径光学系统成像后结合图像复原技术获得清晰的图像, 但传统图像复原技术需要针对光学系统成像中的干扰特性进行分析, 过度依赖人工设计模型且难以适用于不同的场景. 深度学习的发展有力促进了图像复原技术的提升, 针对普通模糊类型图像, Sun 等<sup>[5]</sup> 利用卷积神经网络预测模糊的概率分布, 对受损图像进行去模糊操作, 耗时较长且在不同场景下需要重新进行预测. Tao 等<sup>[6]</sup>、Zamir 等<sup>[7]</sup> 和 Chong 等<sup>[8]</sup>

\* 国家自然科学基金 (批准号: 62071240, 62106111) 资助的课题.

† 通信作者. E-mail: shanhuilin@nuist.edu.cn

采用多尺度网络, 直接对受损图像进行复原, 并对不同阶段的特征进行交叉融合, 优化了特征信息在不同尺度间的传递方式. Li 等<sup>[9]</sup> 关注受损图像的多尺度特征, 将退化表征引入去模糊的过程, 使网络能很好地处理空间变化且复杂的模糊类型. Kupyn 等<sup>[10,11]</sup> 提出基于深度残差结构的去模糊网络, 并引入特征金字塔结构和轻量化的特征提取网络, 利用残差结构提取图像的深层信息, 在多个数据集上取得了良好的去模糊效果, 但对图像暗部细节处理较差. 江泽涛和覃露露<sup>[12]</sup> 采用基于 U-Net 的生成对抗网络来处理暗部细节缺失的图像, 在过暗环境下能够恢复图像的原有细节, 但对图像去模糊、去噪效果不佳. 陈炳权等<sup>[13]</sup>、王山豹等<sup>[14]</sup> 和刘杰等<sup>[15]</sup> 针对图像结构特征的提取, 提出多种尺度特征增强融合方法, 提高了网络特征提取能力和结构锐化表达能力. 在图像复原任务中, 同时关注图像的空间和通道信息尤为重要<sup>[16]</sup>, 王向军和欧阳文森<sup>[17]</sup> 针对空间注意力和通道注意力之间的串、并等多种结合方式, 提出一种最优连接结构. Zamir 等<sup>[18]</sup> 提出基于改进多头注意力机制的 Transformer 网络, 缺点是网络模型较大难以训练. Li 等<sup>[19]</sup> 进一步完善自注意力机制, 通过窗口自注意力和通道注意力增强卷积, 在全局、区域和局部范围内明确地模拟图像层次结构, 在特征提取过程中更重视原始图像结构信息. Tsai 等<sup>[20]</sup> 在图像水平和垂直方向重新加权图像特征, 降低了 Transformer 结构的参数量. Chen 等<sup>[21]</sup> 在 Zamir 等<sup>[18]</sup> 基础上与简单 U-Net 架构进行融合, 进一步降低了网络复杂度, 在去噪和去模糊任务上取得了显著提升.

Tang 等<sup>[22,23]</sup> 提出训练神经网络对合成孔径光学系统图像进行处理, 采用 U-Net 模型对退化的合成孔径图像进行复原, 但其采用的网络结构较为传统, 对原始图像多尺度信息的利用停留在简单跳连接上, 且引入的残差注意力机制未能同时关注空间和通道上的信息, 在合成孔径光学图像复原任务上仍有提升空间. 与传统图像中的运动模糊和抖动模糊不同, 合成孔径光学系统的成像模糊一般为点扩散函数所描述的特定情况, 模糊形状也呈现非均匀分布. 此外, 受子孔径相位排列影响, 合成孔径光学系统成像容易产生振铃现象, 进行复原时易受振铃现象影响而产生伪影.

针对以上问题, 本文提出一种合成孔径光学图像复原方法, 基于多路特征增强的生成对抗网络

架构 (multiple feature enhancement generative adversarial network, MFE-GAN), 能高效、准确恢复受损的合成孔径光学图像. 针对去模糊任务中图像多尺度信息利用不足的问题, 构造了一种多尺度特征聚合模块 (multi-scale feature aggregation module, MSFA), 更好融合不同尺度特征间信息; 针对合成孔径光学退化图像中存在的振铃现象, 构造了一种特征增强模块 (feature enhancement module, FEM), 加强关注退化图像中边缘信息, 锐化图像的真实结构; 针对传统网络对关键特征信息关注不足的情况, 在自注意力和文献<sup>[17]</sup> 的基础上提出一种基于自注意力的混合域注意力机制 (mixed domain attention based on self-attention, MDBS), 关注图像关键信息; 采用多尺度判别器对生成结果在不同尺度上进行综合判别, 加强了鉴别器对生成结果的监督作用. 本文在遥感数据集上进行实验, 和其他主流算法对比, 实验结果证明了本方法在合成孔径光学系统图像复原任务中的有效性和优越性.

## 2 网络模型

### 2.1 合成孔径光学系统退化模型

合成孔径光学系统由多个子孔径按一定方式在像面排列, 在满足相位条件后在焦平面上通过干涉成像. 以环形八孔阵列的光学系统为例, 8 个子孔径环形排列共同作用来代替单一孔径, 但整体系统的 PSF 会发生明显弥散现象, 导致最终成像发生受损退化, 产生振铃现象. 图 1(a), (b) 分别为合成孔径系统八孔径环形阵列结构示意图和发生弥散后的光学系统 PSF.

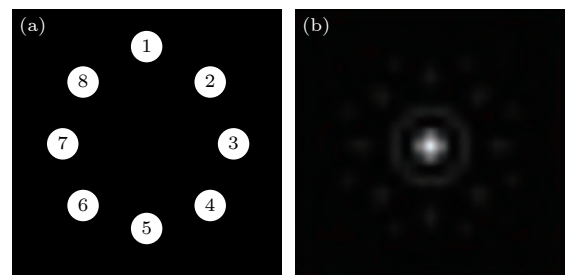


图 1 (a) 合成孔径系统八孔径环形阵列结构示意图; (b) 发生弥散后的光学系统 PSF

Fig. 1. (a) Structure diagram of eight-aperture ring array of synthetic aperture system; (b) optical system PSF after dispersion.

所退化模型的成像过程可简单描述为

$$M(x, y) = I(x, y) * p(x, y) + n(x, y), \quad (1)$$

式中,  $M(x, y)$  表示退化图像,  $I(x, y)$  表示原始图像,  $n(x, y)$  表示成像过程中受到的噪声,  $p(x, y)$  表示光学系统的 PSF,  $*$  表示卷积过程, 即利用实际测得的光学系统点扩散函数与原始图像进行卷积, 模拟实际成像中的受损退化过程, 并加入噪声模拟系统固有噪声. 光学系统 PSF 的具体计算式为

$$F(x, y) = F_{\text{sub}}(x, y) \times \left( N + 2 \sum_{n=1}^N \sum_{m=n+1}^N \cos \left\{ \frac{2\pi}{\lambda f} [(a_n - a_m)x + (b_n - b_m)y] \right\} \right), \quad (2)$$

式中,  $F$  为光学系统的 PSF;  $(x, y)$  为图像的平面坐标;  $\lambda$  为中心波长;  $f$  为焦距;  $a_m$  和  $b_m$  分别表示子孔径圆心在坐标轴上的位置;  $F_{\text{sub}}$  为光学系统中各个子孔径的点扩散函数,

$$F_{\text{sub}} = \left[ DJ_1 \left( \frac{\pi D r}{\lambda f} \right) / 2r \right]^2, \quad (3)$$

其中,  $J_1$  为一阶 Bessel 函数,  $D$  为光瞳的直径,  $r = \sqrt{x^2 + y^2}$ .

## 2.2 MFE-GAN 复原模型

本文提出的 MFE-GAN 复原网络总体框架图如图 2 所示, 由一个生成器和一个多尺度鉴别器共同组成, 其中生成器由编码区、特征融合区和解码区组成. 生成器对输入的模糊图像进行信息采集,

根据关键特征生成新图像后, 传递给判别器在多个尺度上进行综合判别.

## 2.3 生成器模型

对于合成孔径光学图像的复原而言, 网络的特征提取能力和对原始图像多尺度信息的利用程度, 会直接影响复原的结果. MFE-GAN 网络的生成器由编码区、解码区和特征融合区组成. 其中编码区由多层卷积组成, 深层提取图像的底层信息, 利用步长卷积的特性来代替传统 U-Net 网络中的下采样, 获取图像不同尺度特征, 一定程度上减轻了下采样过程中产生的信息丢失问题. 解码区采用解卷积达到上采样的效果, 每层在输入之前与经过特征融合区域后的编码区对应特征相结合, 再进行解卷积操作, 提高了网络的特征提取能力, 使得特征信息在编解码区更好交互, 并在编码区和解码区之间添加隐藏层保证了整体生成器的流通性. 为了更好地关注原始图像的局部细节信息, 分别在编解码区域中部提出基于自注意力的混合域注意力, 让整体网络在编解码的过程中更好关注图像关键信息.

## 2.4 多尺度特征聚合模块

传统 U-Net 网络结构采用简单跳跃连接方式将原始图像信息从编码区传递到解码区域, 未充分利用图像不同尺度的特征. 本文提出一种多尺度特征聚合模块 MSFA, 具体结构如图 3 所示. MSFA 共有 4 个输入, 分别对不同尺度上的特征信息进行融合后再进行跳跃连接. 由于深层特征是由浅层特征卷积而来, 因此采用从下至上的融合方式, 分别让编码区域的深层特征与上一级特征信息融合,

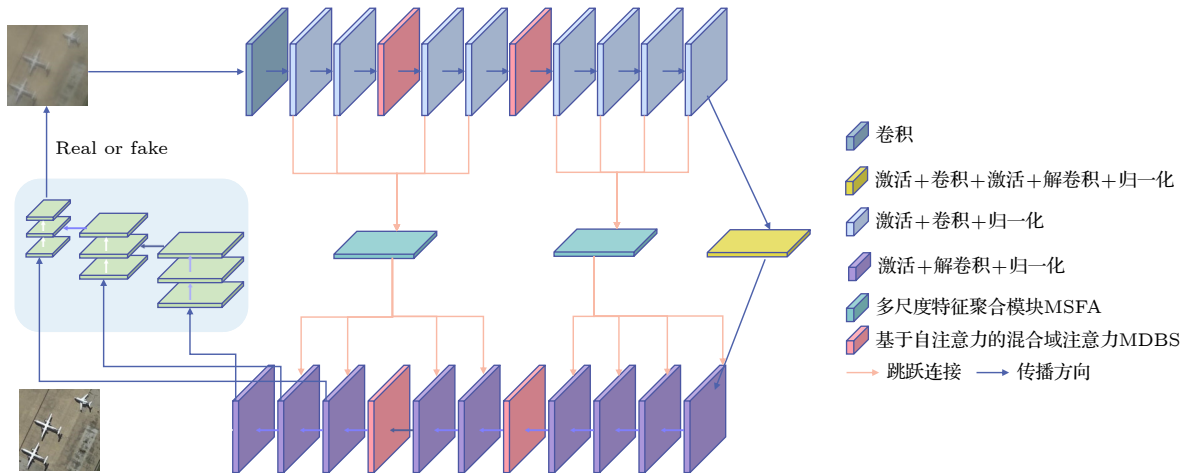


图 2 MFE-GAN 总体框架图

Fig. 2. MFE-GAN General frame diagram.

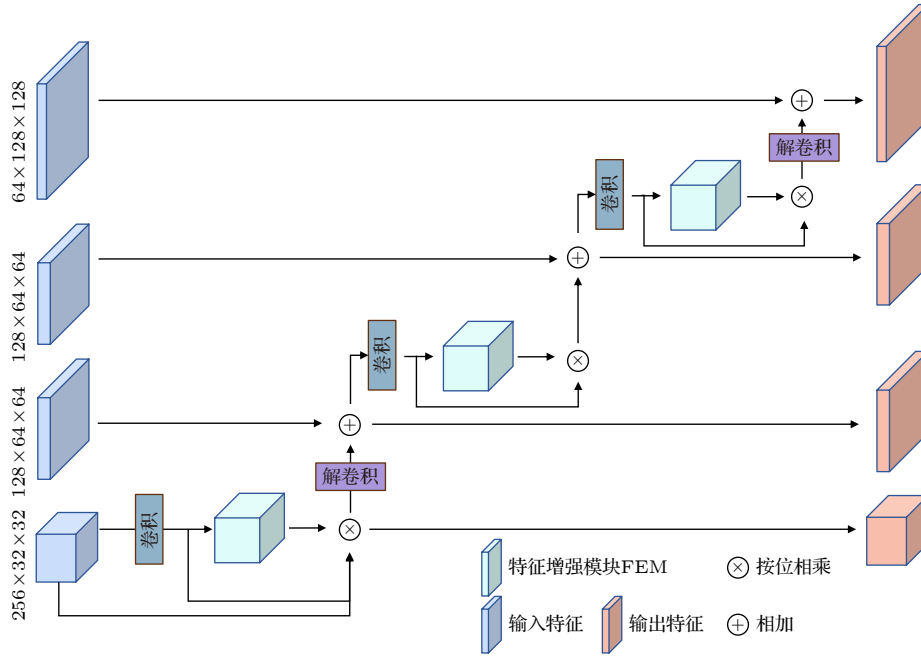


图 3 多尺度特征聚合模块

Fig. 3. Multi-scale feature aggregation module.

再跳跃连接到解码区域,使得跳跃连接的特征图不仅具有浅层特征的信息,还融合了深层特征的信息,更好地利用了原始图像的多尺度特征,进一步挖掘图像深层信息.为了防止尺度差异过大带来的特征信息差异问题,在本文生成器中采用两个多尺度特征聚合模块,对图像的浅层特征和深层特征分别进行特征聚合操作.

图 3 所示为网络浅层中的多尺度特征聚合模块,共有 4 个输入,其中最上层特征大小为  $64 \times 128 \times 128$ ,其中 64 为通道数、128 为宽和高,中间两层特征大小为  $128 \times 64 \times 64$ ,最下层特征大小为  $256 \times 32 \times 32$ .首先利用  $1 \times 1$  的卷积层将特征的通道数调整为与上一层特征的通道数相同,再经过特征增强模块,得到对应特征的权重图后,让输入特征的每个通道特征图和对应的像素权重进行按位相乘.得到全新特征后,一方面输出到解码层与对应生成特征进行跳跃连接,另一方面利用反卷积层、归一化层和激活层将其恢复成与上级特征图相同尺度大小,再与上级特征进行相加.为了使级间特征之间融合更加自然,保证本级特征在多尺度融合的基础上不会被其他级的特征信息过度影响,本文采用门控卷积的思想,利用门控机制让网络自适应的去决定每种尺度特征融入下一尺度特征的多少.

由于合成孔径光学退化图像与传统模糊图像

不同,其模糊类型呈对称状且会产生振铃现象,为使网络更加关注图像的真实结构信息,本文提出一种特征增强模块 FEM,进一步加强每级的尺度特征表达,使生成图像的结构信息更加锐化. MSFA 中,在本级尺度特征与下级尺度特征融合后,通过特征增强模块 FEM 对融合后的特征进行特征增强,获取更加突出的特征表达后一路继续向上级特征融合,一路跳跃连接到解码区对应尺度特征.具体特征加强模块如图 4 所示.

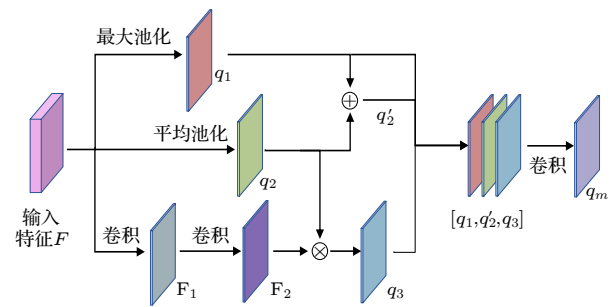


图 4 特征加强模块 FEM

Fig. 4. Feature enhancement module FEM.

FEM 模块在获得输入特征后,根据特征图的不同尺度对其进行通道上的最大池化和平均池化操作,分别获得特征权重图  $q_1$  和  $q_2$ .将  $q_1$  和  $q_2$  进行融合后输出  $q'_2$ ,再对输入特征进行两次卷积后与权重图  $q_2$  进行相乘操作获得特征权重图  $q_3$ ,再与权重图  $q_1$  和  $q_3$  融合后,共同构成特征空间权重

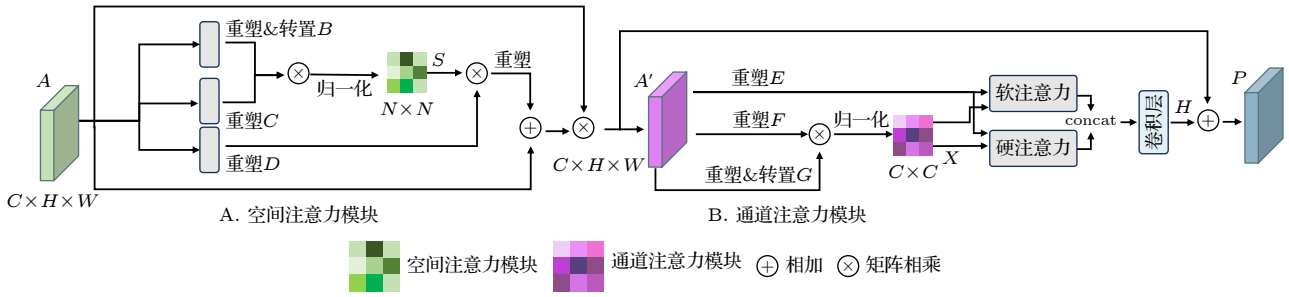


图 5 基于自注意力的混合域注意力 MDBS

Fig. 5. Attention in mixed soft and hard domains MDBS.

图  $[q_1, q_2, q_3]$ . 最终通过卷积操作获得特征空间的像素位权重图  $q_m$ . 在多尺度融合模块中, 融合的不同尺度特征根据每层的特征空间权重图  $q_m$  对特征信息进行重新学习, 提高了网络对关键信息的关注能力及表达能力.

### 2.5 基于自注意力的混合域注意力模块

在提取原始图像尺度特征的过程中, 为了更好地关注图像信息重点, 提出一种基于自注意力的混合域注意力模块 MDBS, 具体结构如图 5 所示. 在 CBAM(convolutional block attention module) 注意力<sup>[16]</sup>和文献<sup>[17]</sup>的研究基础上, 引入自注意力的计算方式, 先对输入特征在空间域上进行增强, 更好关注空间信息. 将通道数为  $C$ 、高为  $H$ 、宽为  $W$  的特征  $A$  经过卷积层后, 通过重塑和转置操作得到大小为  $N \times C$  的特征  $B$  (其中  $N = H \times W$ ), 再将特征  $A$  经过卷积层后仅通过重塑操作得到特征  $C$  和特征  $D$ , 随后将特征  $B$  与特征  $C$  进行矩阵相乘和归一化操作得空间特征图  $S$ . 设置两个自学习的权重因子  $\vartheta$  和  $\beta$  (初始均为 0), 将提取到的特征  $D$  与空间特征图  $S$  进行相乘, 重塑为原始大小尺寸的特征后, 在权重因子  $\vartheta$  的影响下与原始输入特征  $A$  进行相加融合, 再将原始输入特征  $A$  在权重因子  $\beta$  的影响下与融合后的特征进行矩阵相乘, 最终得到重点关注空间信息的特征  $A'$ . 因为在通道域上不需要关注空间信息, 因此直接对特征  $A'$  进行与空间域类似重塑和转置操作, 得到特征  $E$ ,  $F$  和特征  $G$ , 特征  $F$  与特征  $G$  在相乘后进行归一化操作得到通道特征图  $X$ , 随后用特征  $E$  与通道特征图  $X$  分别计算软注意力和硬注意力, 软注意力即计算二者对应位置的注意力权重, 硬注意力即只计算二者对应位置的特征最大值, 最终将得到的两个特征结果在通道域上拼接后利用卷积层进行融合得到特征  $H$ , 再添加一个自学习权重因子  $\delta$  (初

始为 0), 融合后的输出特征在  $\delta$  的影响下与之前提取到的空间特征  $A'$  再次进行相加, 进一步增强模块对空间上的关注能力, 得到最终的输出特征  $P$ .

### 2.6 判别器模型

在生成对抗网络训练中, 若判别器判别能力与生成器的生成能力不匹配, 极易发生判别器无法准确区分原始图像和生成图像的现象, 从而导致整体训练失败. 为使判别器具有更准确、高效的识别能力, 采用三输入的判别器在不同尺度上进行综合判别. 判别器接受生成器生成图像尺度  $P$ ,  $(P - 1)$  和尺度  $(P - 2)$ , 尺度  $P$  经过 3 次卷积后和经过 1 次卷积后的特征图尺度  $(P - 1)$  进行融合, 得到的特征信息再经过一次卷积后与卷积后的尺度  $(P - 2)$  的特征信息进行融合, 最终输出对生成图像的综合评分. 判别器结构如图 6 所示.

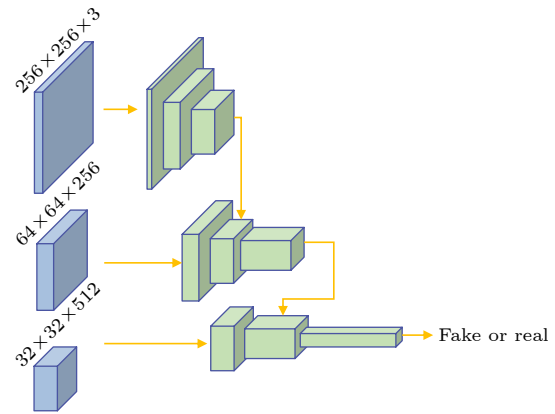


图 6 判别器结构图

Fig. 6. Discriminator structure diagram.

### 2.7 损失函数

为了进一步提高生成对抗网络的训练效率, 缓解生成对抗网络训练中梯度消失和模式崩溃等现象的发生, 本文采用 Wasserstein 距离代替传统

Jensen-Shannon (JS) 散度作为代价函数, 更准确地衡量真实数据分布与生成数据分布之间的差异. 具体定义式为

$$W(p_r, p_g) = \frac{1}{K} \sup_{\|f\|_L \leq K} (E_{x \sim p_r(x)}(f(x)) - E_{x \sim p_g(x)}(f(x))), \quad (4)$$

其中  $p_r$  和  $p_g$  分别表示真实分布和生成分布,  $\sup(\cdot)$  代表函数值的上确界;  $K$  为 Lipschitz 常数,  $\|f\|_L \leq K$  表示函数  $f$  满足  $K$ -Lipschitz 连续,  $E$  为不同分布下的数学期望.

生成器的损失函数定义为生成样本与真实样本之间的 Wasserstein 距离的相反数,

$$L_{1(G)} = -E(D(G(z))). \quad (5)$$

其中,  $E$  为数学期望,  $D$  表示为判别器输出,  $G$  表示为生成器输出,  $z$  表示输入的退化图像. 此外, 在生成器网络中引入了峰值信噪比 (peak signal to noise ratio, PSNR) 损失和结构相似度 (structural similarity, SSIM) 损失, 分别用于减少生成图像与原始图像之间的像素误差和结构误差. PSNR 损失基于均方差 (mean squared error, MSE) 函数设定, SSIM 损失基于两张图像窗口化后的均值、方差和协方差设定. 具体见 (6) 式和 (7) 式:

$$L_{\text{PSNR}} = 10 \lg \left[ \frac{1}{(K-1)^2} \frac{1}{h \times w} \sum_0^{h-1} \sum_0^{w-1} [M(i, j) - N(i, j)]^2 \right], \quad (6)$$

其中,  $K$  为常量, 表示图像的亮度范围, 一般取 255;  $\lg$  函数用于减少欧几里得波动;  $h, w$  分别表示图像宽和高;  $M, N$  分别表示生成图像和原始图像.

$$L_{\text{SSIM}} = -\frac{2\mu_a\mu_b + C_1}{\mu_a^2 + \mu_b^2 + C_1} \times \frac{2\sigma_{ab} + C_2}{\sigma_a^2 + \sigma_b^2 + C_2}, \quad (7)$$

其中,  $a$  和  $b$  分布表示生成区域与原图区域,  $\mu_a$  和  $\mu_b$  表示均值,  $\sigma_a$  和  $\sigma_b$  表示标准差,  $\sigma_{ab}$  表示  $a$  和  $b$  的协方差, 常数  $C_1 = 0.0001$  和  $C_2 = 0.0009$  用于保证分子分母均不为 0.

综上所述, 整体生成器网络的最终损失函数定义为

$$L_G = \lambda L_{1(G)} + \lambda_{\text{psnr}} L_{\text{psnr}} + \lambda_{\text{ssim}} L_{\text{ssim}}, \quad (8)$$

其中  $\lambda$ ,  $\lambda_{\text{psnr}}$  和  $\lambda_{\text{ssim}}$  分别表示用于控制生成器损失、信噪损失和结构损失权重的超参数.

## 3 实验及结果分析

### 3.1 实验环境

具体实验环境配置如表 1 所列.

表 1 实验环境  
Table 1. Experimental environment.

Hardware and software	Configuration
Operating system	Windows10
programming language	
Programming framework	Pytorch2.0.0+ python3.9.16
CPU	12th Gen Intel(R) Core(TM) i9-12900 KF
GPU	Nvidia GeForce RTX3090
Memory	32G
Video memory	24G

### 3.2 合成孔径光学系统图像退化实验

本文实验采用 NWPU-RESISC45 遥感数据集, 其中共包含大小为  $256 \times 256$  的 31500 张图像, 涵盖 45 个类别, 每个类别包含 700 张图像, 为了证明本方法的普遍适用性, 从中随机挑选 20 类, 每类随机选出 400 张图像, 共计 8000 张图像. 为了模拟合成孔径光学系统成像异常的情况, 本文采用八孔径环形阵列结构的合成孔径光学系统实际测得的系统点扩散函数, 利用 (1) 式和 (2) 式计算生成退化图像, 并对退化图像加入 55—65 dB 加性高斯噪声来模拟真实成像过程中的系统噪声, 作为本实验数据集. 其中随机选出 7000 张作为训练集, 1000 张作为测试集. 图 7 为原始图像与退化图像对比图, 其中图 7(a) 为原始遥感图像, 图 7(b) 为退化模型生成的带有振铃现象和系统噪声的退化图像.

### 3.3 对比实验

针对合成孔径光学系统成像带来的降质模糊和振铃现象等问题, 目前基本采用传统算法来解决. 因此本文选取传统的复原算法维纳滤波 (Wiener filtering, WF) 算法 [5], 基于深度学习的去模糊算法 Deblur GAN 算法 [10]、SRN 算法 [6]、MPR-Net 算法 [7] 和 Stripformer 算法 [20] 与 MFE-GAN 算法进行对比. 具体可视化结果如图 8 所示, 其中图 8(a) 为原始清晰图像, 图 8(b) 为经过合成孔径退化模型后的退化图像, 图 8(c)—(g) 分别为对比算法可视化结果, 图 8(h) 为 MFE-GAN 算法可视化结果.

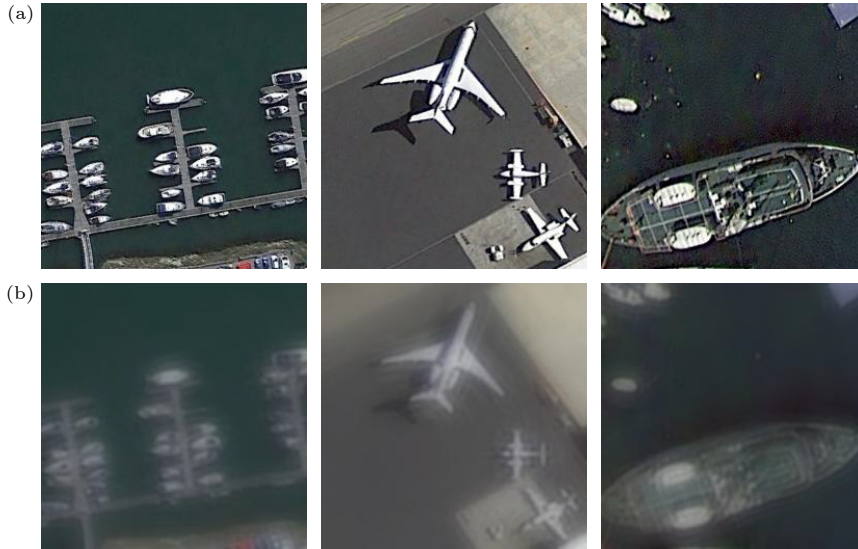


图 7 (a) 原始图像与 (b) 退化图像对比

Fig. 7. Comparison between original image (a) and degraded image(b).

为了进一步评估本方法的复原效果,引入峰值信噪比、结构相似度、可学习感知图像块相似度 (learned perceptual image patch similarity, LPIPS) 和弗雷歇距离 (Fréchet inception distance, FID) 4 种图像质量评价标准. 其中, 峰值信噪比指标在像素层面上对生成图像和原始图像进行对比; 结构相似度指标综合考虑了生成图像和原始图像在亮度、对比度和结构方面的相似性; 可学习感知图像块相似度通过学习感知特征来评估两张图像之间的差异和相似度, 能较好地反映人类视觉感知中的差异; 弗雷歇距离则是将生成图像和原始图像输入至预训练的 Inception 网络中进行特征比较来判断两幅图像的差异. 数值上, PSNR, SSIM 值越大, 而 LPIPS, FID 值越小表明生成的图像更加合理、真实. 不同方法定量对比结果如表 2 所列.

表 2 不同算法定量实验结果  
Table 2. Quantitative experiment results of different algorithms.

对比算法	PSNR/dB	SSIM	LPIPS	FID
Wiener filtering	15.227	0.5452	0.0896	3.9776
Deblur GAN	22.379	0.7011	0.0383	1.2534
SRN	17.639	0.6588	0.0697	2.1246
MPR-Net	24.725	0.7364	0.0515	<b>0.5728</b>
Stripformer	26.013	0.7556	0.0384	0.7138
Ours	<b>26.408</b>	<b>0.7890</b>	<b>0.0363</b>	0.8979

由各算法对比可视化结果图 8 可见, 传统算法维纳滤波方法生成结果较差, 图像具有明显伪影和噪点, 如表 2 所示定量实验结果的峰值信噪比和

结构相似性都非常低, 仅为 15.227 和 0.5452, 同时 LPIPS 和 FID 指标分别为 0.0896 和 3.9776, 表示生成图像素质差, 与原始图像差距较大; Deblur GAN 算法的生成图像在整体色彩结构上较为明亮, 其在细节部分仍存在噪点, 整体结构恢复较为完整, 峰值信噪比和结构相似度为 22.379 和 0.7011, LPIPS 和 FID 分别为 0.0383 和 1.2534; SRN 算法的生成图像效果略差, 在关键结构附近会生成伪影, 且具有明显噪点, 但与传统维纳滤波算法相比较好, 其峰值信噪比为 17.639, 结构相似度为 0.6588, LPIPS 和 FID 分别为 0.0697 和 2.1246; MPR-Net 算法的生成结果较优秀, 整体图像噪点较少, 但因为生成图像过于平滑, 部分图像的锐化细节丢失, 其峰值信噪比为 24.725, 结构相似度为 0.7364; Stripformer 算法虽在峰值信噪比上大大提高, 但其生成图像会丢失部分细小的边线信息, 如图 8(g) 中飞机图像的右上部裂缝未能很好复原; 相比于上述算法, 本文所提方法可以有效地复原退化图像, 不仅能很好恢复图像的整体结构, 对于细节信息也不会过度丢失, 其峰值信噪比和结构相似度提高了 0.395 和 0.0334, LPIPS 和 FID 分别为 0.0363 和 0.8979.

针对合成孔径光学系统成像中出现的振铃现象, 本方法提出的特征加强模块可以重点关注退化图像的结构信息, 为进一步验证该方法的独特性, 选用受 3 种不同振铃增益因子  $n$  影响下的退化图像, 与正常退化图像进行复原对比, 如图 9 所示.

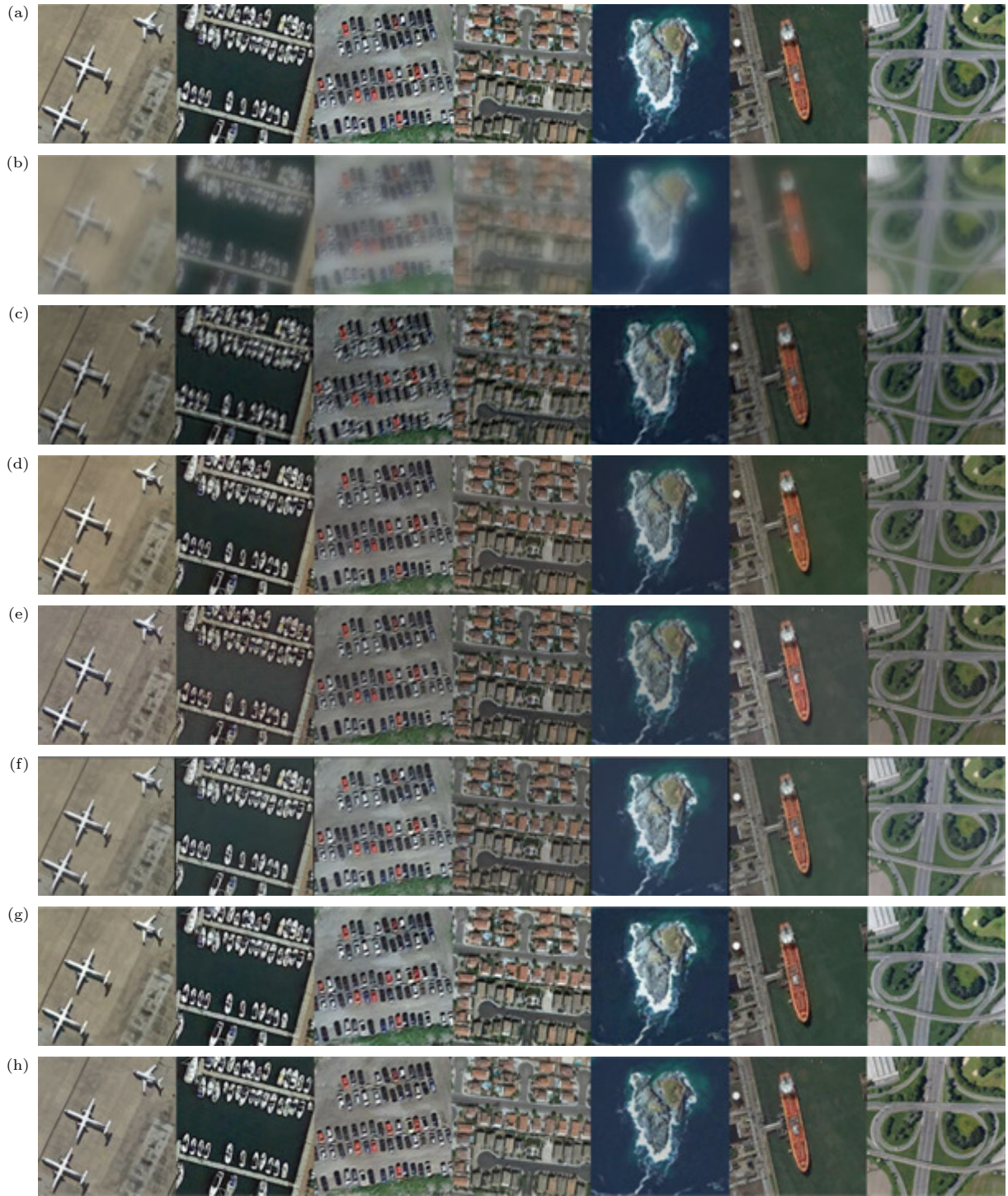


图 8 对比试验可视化结果 (a) 原始清晰图像; (b) 退化图像; (c) 维纳滤波算法; (d) Deblur GAN 算法; (e) SRN 算法; (f) MPR-Net 算法; (g) Stripformer 算法; (h) 本文算法

Fig. 8. Visualization results of comparison experiment: (a) Clear image; (b) degraded image; (c) Wiener filtering; (d) Deblur GAN; (e) SRN; (f) MPR-Net; (g) Stripformer; (h) our proposed method.

图 9(a) 为  $n = 0$  时真实光学仪器系统成像结果, 图 9(b)—(d) 是提高振铃增益因子后的结果, 随着增益因子  $n$  的逐渐提高, 原始图像结构附近出现更多伪影, 会对图像复原的结果造成较大干扰. 本文提出的特征加强模块 FEM 可以锐化图像特

征, 复原出真实结构信息, 具体复原结果如图 10 所示.

由图 10 可见, 当振铃增益因子  $n = 0$  和  $n = 1$  时修复结果几乎无区别, 当  $n$  增大到 10 时, 才会在局部细节区域出现伪影. 为了进一步证明本方法的

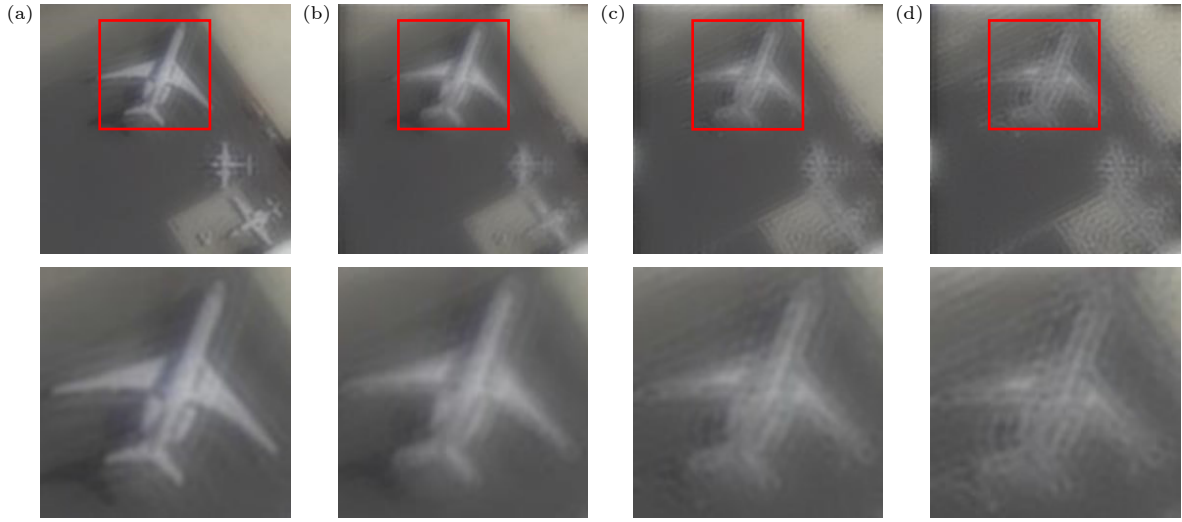


图 9 不同振铃增益影响下的退化图像 (a)  $n = 0$ ; (b)  $n = 1$ ; (c)  $n = 2$ ; (d)  $n = 10$

Fig. 9. Degraded images under the influence of different ringing gains: (a)  $n = 0$ ; (b)  $n = 1$ ; (c)  $n = 2$ ; (d)  $n = 10$ .

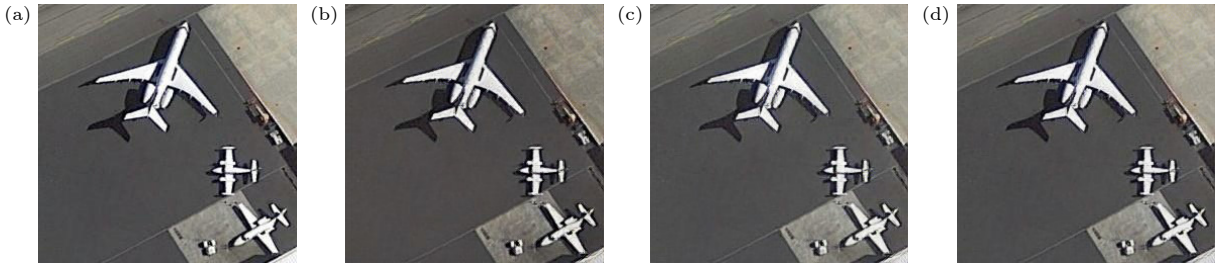


图 10 不同振铃增益影响下本文方法对图像的复原结果 (a)  $n = 0$ ; (b)  $n = 1$ ; (c)  $n = 2$ ; (d)  $n = 10$

Fig. 10. Results of image restoration obtained by this method under the influence of different ringing gains: (a)  $n = 0$ ; (b)  $n = 1$ ; (c)  $n = 2$ ; (d)  $n = 10$ .

优越性, 选取振铃增益因子  $n = 10$  的受损图像进行复原, 与 3 种深度学习复原方法进行定量对比. 由图 11 可见, Deblur GAN 的复原结果在飞机周围出现明显伪影, 且整体色彩有偏差; SRN 针对较大目标的复原结果较好, 但在小目标复原时仍在细节处出现错位伪影; MPR-Net 方法对结构特征处理较好, 基本复原了特征信息, 但整体图像噪点较高; Stripformer 方法生成的图像会丢失部分边线细节信息; 本文方法相对更好地生成了原始图像的结构信息, 几乎未出现伪影, 一定程度上也降低了部分噪点. 具体定量结果如表 3 所列.

此外, 本文方法对于部分颜色异常的图像也有很好的复原效果, 如过度曝光和整体光线较暗等异常情况, 具体结果如图 12 所示, 其中, 图 12(a) 为光线异常图像, 图 12(b) 为退化后的光线异常图像, 图 12(c) 为本文复原后的生成图像. 针对于过度曝光情况, 本文方法可以将其复原为正常颜色, 能较好突出原始图像的结构信息和细节信息; 对于

整体光线较暗的情况, 本文方法可以提高其整体明亮程度, 并突出相应细节信息.

表 3 在  $n = 10$  振铃增益影响下不同算法的定量实验结果

Table 3. Quantitative experimental results of different algorithms under the influence of ringing gain ( $n = 10$ ).

对比算法	PSNR/dB	SSIM	LPIPS
Deblur GAN	16.867	0.6278	0.1006
SRN	19.369	0.7454	0.0545
MPR-Net	20.166	0.7624	0.0437
Stripformer	22.023	0.7472	0.0483
Ours	<b>22.162</b>	<b>0.7702</b>	<b>0.0426</b>

### 3.4 消融实验

为验证本文方法中各模块对整体网络复原性能的影响, 进行消融实验, 在此过程中除消融项外, 其他结构和参数保持不变. 具体消融实验结果如表 4 所列, 标记“√”表示模型在训练过程中包含了

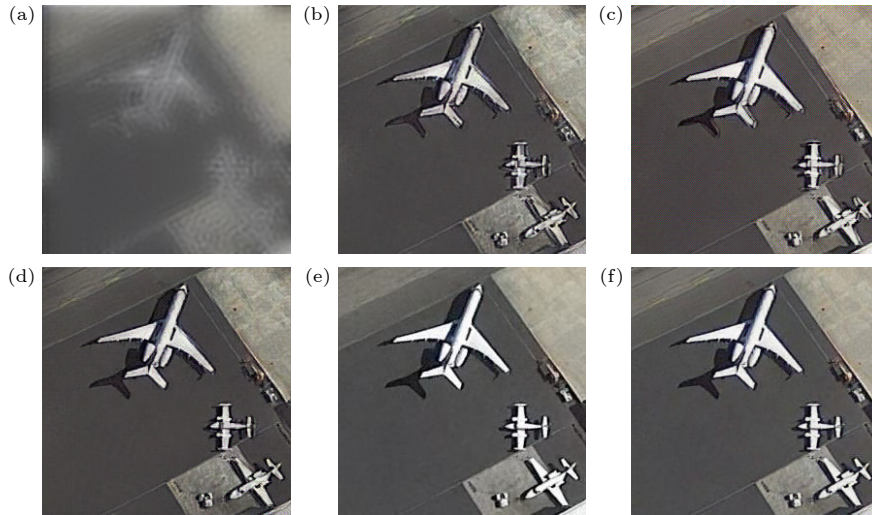


图 11 振铃增益  $n = 10$  时, 本文方法与其他方法对图像的复原结果对比 (a) 受损图像; (b) Deblur GAN 算法; (c) SRN 算法; (d) MPR-Net 算法; (e) Stripformer 算法; (f) 本文方法

Fig. 11. Comparison of image restoration results between the proposed method and other methods under the influence of ringing gain ( $n = 10$ ): (a) Damaged image; (b) Deblur GAN; (c) SRN; (d) MPR-Net; (e) Stripformer; (f) our proposed method.



图 12 对光线异常退化图像的修复结果 (a) 光线异常图像; (b) 退化后的光线异常图像; (c) 本文复原后的生成图像

Fig. 12. Repair results of abnormal light degradation images: (a) Abnormal light image; (b) degraded images of light anomalies; (c) generated image after restoration of ours.

对应项. 其中策略 1—策略 4 采用传统判别器, 策略 1 表示整体网络仅采用 U-Net 结构作为生成器; 策略 2 表示在 U-Net 基础上引入多尺度特征融合模块; 策略 3 表示在多尺度特征融合的基础上引入

特征加强模块; 策略 4 表示在策略 3 的基础上引入混合域注意力; 策略 5—策略 9 采用多尺度判别器进行判别, 策略 5 表示在生成器中引入 U-Net 结构、多尺度特征融合模块和特征加强模块; 策略 6

表 4 消融实验数据对比

Table 4. Comparison of G1 ablation data of coarse repair network.

消融策略	U-Net	多尺度特征聚合模块	特征增强模块	混合域注意力	多尺度判别器	PSNR/dB	SSIM	LPIPS	FID
策略1	✓					19.306	0.5865	0.0841	1.2976
策略2	✓	✓				21.904	0.6714	0.0833	1.0742
策略3	✓	✓	✓			23.872	0.6927	0.0579	0.9103
策略4	✓	✓	✓	✓		25.425	0.7563	0.0515	0.9157
策略5	✓	✓	✓		✓	23.261	0.7401	0.0604	0.9247
策略6	✓	✓		✓	✓	22.558	0.7342	0.0739	0.9508
策略7	✓		✓	✓	✓	23.074	0.7095	0.0637	0.9460
策略8	✓			✓	✓	20.839	0.6584	0.0799	0.9882
策略9	✓		✓		✓	22.944	0.6627	0.0682	0.9763
本文策略	✓	✓	✓	✓	✓	<b>26.408</b>	<b>0.7890</b>	<b>0.0363</b>	<b>0.8979</b>

表示在策略 5 的基础上使用混合域注意力, 而不采用特征加强模块; 策略 7 表示不采用多尺度特征融合模块, 对 U-Net 跳连接分别采用特征加强模块, 同时引入注意力模块和多尺度判别器; 策略 8 表示不采用多尺度特征融合模块和特征加强模块, 仅测试注意力模块和多尺度判别器组合对整体网络带来的性能影响; 策略 9 表示仅采用 U-Net 结构、特征加强模块和多尺度判别器. 实验结果对比如表 4 所列.

策略 1 仅采用 U-Net 结构的生成对抗网络时, 峰值信噪比和结构相似度都较低, 分别为 19.306 和 0.5865, 这代表整体生成结果较差; 策略 2 引入多尺度特征融合模块, 更好地利用了图像的多尺度特征, 峰值信噪比和结构相似度大幅上升; 策略 3 进一步在多尺度特征融合模块中加入特征加强模块, 使得整体网络更加关注图像的多尺度特征, 复原结果进一步提高, 峰值信噪比和结构相似性分别提高到 23.872 和 0.6927, LPIPS 和 FID 分别为 0.0579 和 0.9103; 策略 4 在编码器中加入混合域注意力, 进一步关注图像的高级特征, 优化了复原结果, 峰值信噪比和结构相似性分别提高到 25.425 和 0.7563; 策略 5 取消了注意力模块, 将判别器改为多尺度判别器, 使得整体网络更好收敛, 但因缺少注意力因素, 结果相比于策略 4 较差; 策略 6 和策略 7 在与本文策略对比下, 证明了将多尺度融合模块和特征加强模块相结合的效果最佳; 策略 8 直接将二者取消, 结果峰值信噪比和结构相似性大幅度下降; 策略 9 在策略 7 的基础上取消了注意力模块, 使峰值信噪比和结构相似性分别下降到 22.944 和 0.6627, LPIPS 和 FID 分别为 0.0682 和 0.9763, 证明了注意力模块的必要性.

## 4 结 论

本文提出一种基于多尺度特征增强的合成孔径光学图像复原方法, 针对合成孔径光学系统成像出现的振铃现象专门提出特征加强模块, 关注原始图像的真实结构信息, 同时利用基于自注意力的混合域注意力增加整体网络对关键信息的关注能力. 在 NWPU-RESISC45 遥感数据集上采用真实光学仪器测得模糊核进行模拟退化实验并进行复原实验, 相比于传统光学系统图像复原算法复原效果明显提高, 且具有更好的适用性. 相比于深度学习主流复原算法, 能更好复原带有振铃现象的退化图像, 最终结果在 PSNR 和 SSIM 评估指标上分别提高了 1.51% 和 4.42%, LPIPS 和 FID 指标上分别达到 0.0363 和 0.8979. 此外原始图像数据过度曝光或光线不足等异常情况, 本文方法也能将其复原, 证明了本方法在实际场景中的实用性. 本文方法注重采集原始图像的多尺度信息进行特征融合, 造成整体网络参数量略大, 部署到移动端后对退化图像复原时间略长, 无法做到实时复原, 未来工作将针对各模块进行轻量化设计, 包括注意力机制轻量化和减轻特征融合中的冗余计算过程, 使网络参数量更小, 从而做到实时复原的需求.

感谢中国科学院光电技术研究所钟烁、范斌老师为本文退化实验提供的真实光学系统模型参数.

## 参考文献

- [1] Li J J, Zhou N, Sun J S, Zhou S, Bai Z D, Lu L P, Chen Q, Zuo C 2022 *Light Sci. Appl.* **11** 154
- [2] Li D J, Gao J H, Cui H J, Zhou K, Wu J 2022 *Chin. J. Lasers* **49** 0310001 (in Chinese) [李道京, 高敬涵, 崔岸婧, 周

- 凯, 吴疆 2022 中国激光 **49** 0310001]
- [3] Shinwook K, Youngchun Y 2023 *Opt. Express* **31** 4942
- [4] Liu Z, Wang S Q, Rao C H 2012 *Acta Phys. Sin.* **61** 039501 (in Chinese) [刘政, 王胜千, 饶长辉 2012 物理学报 **61** 039501]
- [5] Sun J, Cao W F, Xu Z B, Ponce J 2015 *IEEE Conference on Computer Vision and Pattern Recognition* Boston, MA, USA, June 7–12, 2015 p769
- [6] Tao X, Gao H Y, Shen X Y, Wang J, Jia J Y 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* Salt Lake City, UT, USA, June 18–23, 2018 p8174
- [7] Zamir S W, Arora A, Khan S, Hayat M, Khan F S, Yang M H, Shao L 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* Nashville, TN, USA, June 20–25, 2021 p14816
- [8] Chong M, Wang Q, Zhang J 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* New Orleans, LA, USA, June 18–24, 2022 p17378
- [9] Li, D S, Zhang Y, Cheung K, Wang X G, Qin H W, Li H S 2022 *European Conference on Computer Vision (ECCV)* Tel Aviv, Israel October 23–27, 2022 p736
- [10] Kupyn O, Budzan V, Mykhailych M, Mishkin D, Matas J 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* Salt Lake City, UT, USA, June 18–23, 2018 p8183
- [11] Kupyn O, Martyniuk T, Wu J R, Wang Z Y 2019 *IEEE/CVF International Conference on Computer Vision* Seoul, Korea (South), Oct. 27–Nov. 2, 2019 p8877
- [12] Jiang Z T, Qin L L 2020 *J. Electron.* **48** 258 (in Chinese) [江泽涛, 覃露露 2020 电子学报 **48** 258]
- [13] Chen B Q, Zhu X, Wang Z Y, Liang Y C 2022 *J. Hunan Univ. (Nat. Sci.)* **49** 124 (in Chinese) [陈炳权, 朱熙, 汪政阳, 梁寅聪 2022 湖南大学学报 (自然科学版) **49** 124]
- [14] Wang S B, Liang D, Shen 2023 *J. Comput. Aided Design Comput. Graphics* **35** 1109 (in Chinese) [王山豹, 梁栋, 沈玲 2023 计算机辅助设计与图形学学报 **35** 1109]
- [15] Liu J, Qi R, Han K 2023 *Opt. Precis. Eng.* **31** 2080 (in Chinese) [刘杰, 祁睿, 韩轲 2023 光学精密工程 **31** 2080]
- [16] Woo S H, Park J, Lee J Y, Kweon I S 2018 *Proceedings of the European Conference on Computer Vision* Munich, Germany, September 8–14, 2018 p3
- [17] Wang X J, Ouyang W S 2022 *Infrared Laser Eng.* **51** 460 (in Chinese) [王向军, 欧阳文森 2022 红外与激光工程 **51** 460]
- [18] Zamir S W, Arora A, Khan S, Hayat M, Khan F S, Yang M H 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* New Orleans, LA, USA, June 18–24, 2022 p5718
- [19] Li Y W, Fan Y C, Xiang X Y, Demandolx D, Ranjan R, Timofte R, Gool L V 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* Vancouver, BC, Canada, June 17–24, 2023 p18278
- [20] Tsai F J, Peng Y T, Lin Y Y, Tsai C C, Lin C W 2022 *European Conference on Computer Vision (ECCV)* Tel Aviv, Israel, October 23–27, 2022 p146
- [21] Chen L Y, Chu X J, Zhang X, Sun J 2022 *European Conference on Computer Vision (ECCV)* Tel Aviv, Israel, October 23–27, 2022 p17
- [22] Tang J, Wang K Q, Ren Z R, Zhang W, Wu X Y, Di J L, Liu G D, Zhao J L 2020 *Opt. Lasers Eng.* **139** 106463
- [23] Tang J, Wu J, Wang K Q, Ren Z B, Wu X Y, Hu L S, Di J L, Liu G D, Zhao J L 2021 *Opt. Lasers Eng.* **146** 106707

## Synthetic aperture optical image restoration based on multi-scale feature enhancement\*

Zhang Yin-Sheng<sup>1)2)</sup> Tong Jun-Yi<sup>2)</sup> Chen Ge<sup>2)</sup> Shan Meng-Jiao<sup>2)</sup>  
Wang Shuo-Yang<sup>2)</sup> Shan Hui-Lin<sup>1)2)†</sup>

1) (*Jiangsu Province Engineering Research Center of Integrated Circuit Reliability Technology and Testing System, Wuxi University, Wuxi 214105, China*)

2) (*School of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China*)

( Received 6 November 2023; revised manuscript received 17 January 2024 )

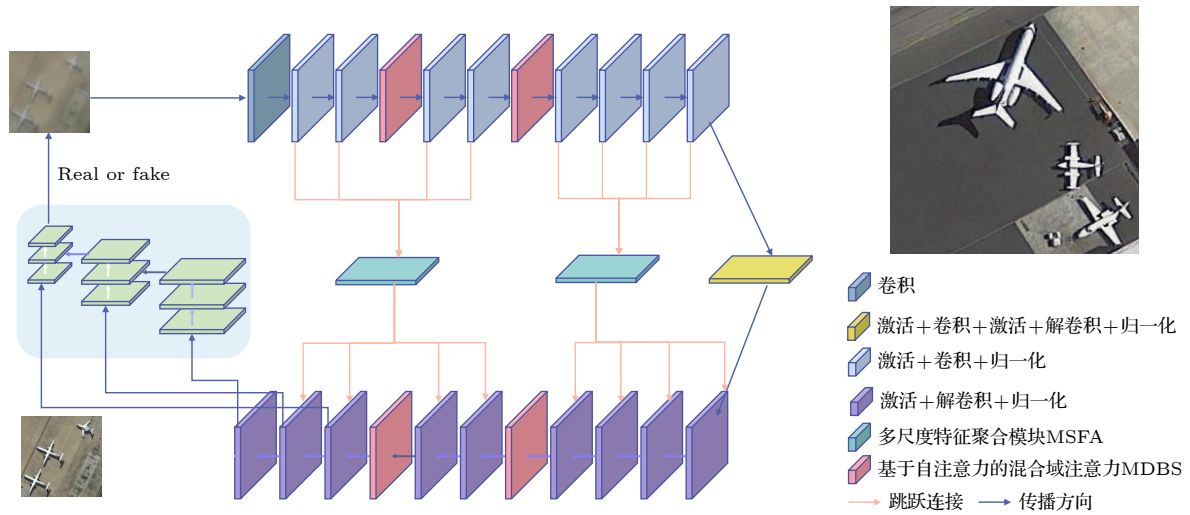
### Abstract

With the wide applications of high-resolution imaging technology in topographic mapping, astronomical observation, and military reconnaissance and other fields, the requirements for imaging resolution of optical system are becoming higher and higher. According to the diffraction limit and Rayleigh criterion, the imaging resolution of the optical system is proportional to the size of the aperture of the system, but affected by the material and the processing of the optical component: the single aperture of the optical system cannot be

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 62071240, 62106111).

† Corresponding author. E-mail: shanhuilin@nuist.edu.cn

infinitely enlarged. Therefore the synthetic aperture technology is proposed to replace the single large aperture optical system. Owing to the effect of sub-aperture arrangement and light scattering, the imaging of synthetic aperture optical system will be degraded because of insufficient light area and phase distortion. The traditional imaging restoration algorithm of synthetic aperture optical system is sensitive to noise, overly relies on degraded model, requires a lot of manually designed models, and has poor adaptability. To solve this problem, a multi-scale feature enhancement method of restoring the synthetic aperture optical image is proposed in this work. U-Net is used to obtain multi-scale feature, and self-attention in mixed domain is used to improve the ability of of the network to extract the features in space and channel. Multi-scale feature fusion module and feature enhancement module are constructed to fuse the information between features on different scales. The information interaction mode of the codec layer is optimized, the attention of the whole network to the real structure of the original image is enhanced, and the artifact interference caused by ringing is avoided in the process of restoration. The final experimental results are 1.51%, 4.42% and 5.22% higher than those from the advanced deep learning algorithms in the evaluation indexes of peak signal-to-noise ratio, structural similarity and perceived similarity, respectively. In addition, the method presented in this work has a good restoration effect on the degraded images to different degrees of synthetic aperture, and can effectively restore the degraded images and the images with abnormal light, so as to solve the problem of imaging degradation of synthetic aperture optical system. The feasibility of deep learning method in synthetic aperture optical image restoration is proved.



**Keywords:** image processing, synthetic aperture, multi-scale, feature fusion

**PACS:** 42.30.Tz, 07.05.Pj, 95.75.Mn

**DOI:** [10.7498/aps.73.20231761](https://doi.org/10.7498/aps.73.20231761)



## 基于多尺度特征增强的合成孔径光学图像复原

张银胜 童俊毅 陈戈 单梦姣 王硕洋 单慧琳

### Synthetic aperture optical image restoration based on multi-scale feature enhancement

Zhang Yin-Sheng Tong Jun-Yi Chen Ge Shan Meng-Jiao Wang Shuo-Yang Shan Hui-Lin

引用信息 Citation: *Acta Physica Sinica*, 73, 064203 (2024) DOI: 10.7498/aps.73.20231761

在线阅读 View online: <https://doi.org/10.7498/aps.73.20231761>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 气泡在超声场中绕圈运动的高速摄影及其图像分析

High-speed photography and image analysis of orbital motion of gas bubbles in ultrasonic field

物理学报. 2021, 70(5): 054701 <https://doi.org/10.7498/aps.70.20201381>

#### 基于多层声速模型的合成孔径超声皮质骨成像

Multi-layer velocity model based synthetic aperture ultrasound imaging of cortical bone

物理学报. 2019, 68(18): 184302 <https://doi.org/10.7498/aps.68.20190763>

#### 基于多引导滤波的图像增强算法

Image enhancement based on multi-guided filtering

物理学报. 2018, 67(23): 238701 <https://doi.org/10.7498/aps.67.20181425>

#### 多尺度水下偏振成像方法

Multi-scale analysis method of underwater polarization imaging

物理学报. 2018, 67(5): 054202 <https://doi.org/10.7498/aps.67.20172009>

#### 基于变分模态分解与多尺度排列熵的生物组织变性识别

Recognition of denatured biological tissue based on variational mode decomposition and multi-scale permutation entropy

物理学报. 2019, 68(2): 028702 <https://doi.org/10.7498/aps.68.20181772>

#### 基于多模光纤散斑的压缩感知在光学图像加密中的应用

Application of compressive sensing based on multimode fiber specklegram in optical image encryption

物理学报. 2020, 69(3): 034203 <https://doi.org/10.7498/aps.69.20191143>