基于离散小波变换的细节增强的图像修复优化方法

郑博伟[1] 李宗辉[1] 陈锐彬[1] 黄梅佳[1]
ZHENG Bowei LI Zonghui CHEN Ruibin HUANG Meijia

摘 要 针对现有基于小波变换的深度图像修复方法存在的不足进行优化，提出了一种基于离散小波变换的细节增强的深度图像修复优化方法。首先，采用一种求和方法将现有方法中的冗余高频信息加以利用，进一步增强修复图像的高频细节；其次，提出一种基于稠密神经网络（DenseNet）的判别网络结构，改进原有方法中的判别网络结构，以提高修复图像的质量；最后，在多个公共数据集上进行大量实验，实验结果表明，所提出的优化方法具有有效性。

关键词 图像修复；高频信息；稠密神经网络；判别网络

doi：10.3969/j.issn.1672-9528.2024.06.010

0 引言

图像修复[1]作为一种图像编辑技术，旨在用合理的内容填充图像中缺失或损坏的部分，并且能在视觉上不留痕迹，使得修复后的整张图片看起来语义完整。近年来，深度学习的方法在图像修复领域取得了重大突破，它可以提取缺失图像的高级语义特征，从而合理地填补缺失区域。

Pathak 等人[2]首次利用生成对抗网络[3]（generative adversarial network，GAN）来进行图像修复工作，模型基于自编码器[4]（auto-encoder，AE）的生成网络和基于 GAN 的判别网络，但其算法缺乏对图像语义信息的有效利用。Iizuka 等人[5]在他们的图像修复工作基础上引入了全局和局部判别器，通过使用全局和局部判别器来更好地捕捉图像的全局和局部特征，但对于一些复杂和特殊的缺失区域，可能需要进行更精细的处理和调整。近些年来，深度学习发展迅速，图像修复算法也不断完善。Zeng 等人[6]提出了一种基于金字塔式编码网络的图像修复算法。Li 等人[7]提出基于小波变换的细节增强（detail-enhanced）的图像修复算法，首先对图像进行小波变换，得到不同频率的子图像，然后通过并行的两个网络分别对图像的高频和低频进行修复，最后通过小波逆变换，将处理后的子图像重新组合成修复后的图像。该方法是一种有效的图像修复算法，它能够通过增强图像的细节信息来提高图像的质量，适用于不同类型的图像修复任务。Deng 等人[8]的方法是通过对图像的每个部分赋予不同的权重，从而突出重要的区域并抑制不重要的区域，网络可以自适应地学习图像的几何变换和先验知识，从而更好地修复图像。

其中，Li 等人[7]提出基于离散小波变换的细节增强（detail-enhanced image inpainting，DEII）的图像修复方法，利用小波变换将缺失内容的高频部分和低频部分分别在两个并行的双分支结构（内容分支和纹理分支）中进行填补，最后通过小波逆变换得到重建后的修复图像。DEII 算法虽然实现了修复图像细节信息的增强，但其修复网络结构中有比较多的冗余高频信息没有被使用，该部分信息应加以利用，以进一步增强修复图像细节；并且该算法对于一些场景的修复，依旧存在语义结构扭曲、纹理模糊等问题，应改进生成对抗网络结构，以提高修复图像的质量。

针对上述 DEII 图像修复方法的问题，本文进行优化，提出了一种基于离散小波变换的细节增强的图像修复优化方法，分别对原有方法的修复网络和判别网络进行改进。首先，针对原有修复网络中没有被使用的高频信息，即内容分支填补完成的三个高频小波子带，本文使用求和方法（add）将其加到对应纹理分支中生成的三个高频小波子带，以进一步增强生成的三个高频小波子带的纹理细节，再与内容分支填补完成的低频子带一起通过小波逆变换得到重建后的修复图像。然后，本文设计一个基于稠密神经网络结构（DenseNet）的判别网络，改进原有判别网络，以稳定训练，提高修复图像的质量。最后，在几个公共数据集上进行的大量实验表明，本文所提出的方法能有效改进 DEII 图像修复方法，提高其修复图像的质量。

1 DEII 方法简介

1.1 网络整体框架

如图 1 所示，DEII 网络是由两个模块构成，即修复网络

1. 揭阳职业技术学院信息工程系 广东揭阳 522000
[基金项目] 2022 年揭阳职业技术学院科学研究项目“基于深度学习的图像修复技术的研究”（2022JYCKY05）

（G）和判别网络（D），每个网络又由两个分支构成。G 由内容网络（G_{con}）和纹理网络（G_{txt}）构成，D 由全局判别网络（D_{glb}）和局部判别网络（D_{loc}）构成。在训练阶段，修复网络和判别网络一起训练，实际测试阶段，只需使用修复网络完成修复。

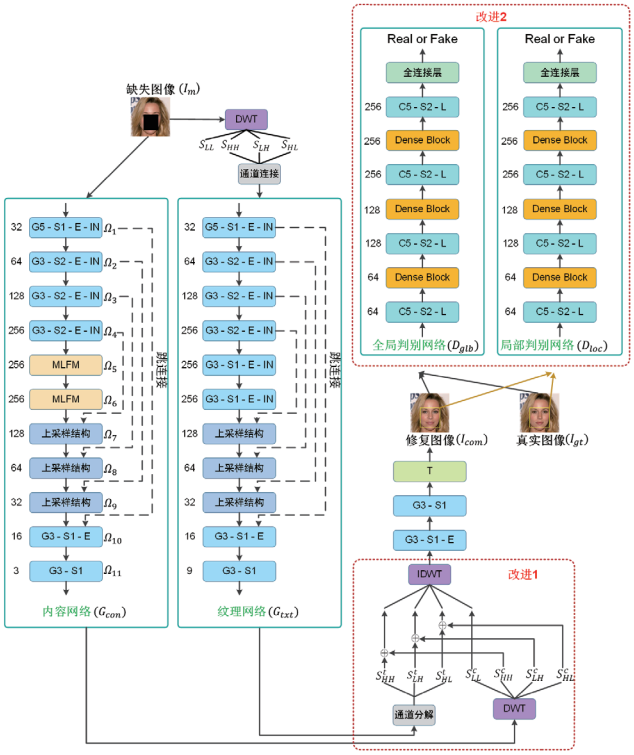


图 1 网络整体框架

注：“Ga-Sb-E-IN”表示 a*a 门控卷积（GConv）—步长 b—ELU 激活函数—实例规范。“Ca-Sb-L”表示 a*a 普通卷积（VConv）—步长 b—LeakyRelu 激活函数。“T”表示 Tanh 激活函数。除了 MLFM，该网络所有卷积的空洞率为 1。$\Omega_i(i \in \{1,2,\cdots,11\})$ 表示第 i 个操作的输出特征。“DWT”和“IDWT”分别为离散小波变换和逆变换，其小波基为 Daubechies2(db2)。

给定一张缺失图像（I_m），它可以被认为是一张真实图像（I_{gt}）的退化，由一张掩膜 M（1 代表缺失区域，0 代表真实区域）来指示缺失区域，即 $I_m = I_{gt} \odot (1 - M)$，$\odot$ 代表逐元素乘法，则一张修复图像（I_{com}）可以被定义为：

$$I_{com} = G(I_m) \odot M + I_m \tag{1}$$

1.2 修复网络结构设计

在 DEII 的修复网络 G 中，两个并行的分支网络（内容网络 G_{con} 和纹理网络 G_{txt}）分别完成图像的语义内容和纹理细节的填补。第一分支在空域中填补缺失图像的语义内容，第二分支在小波域中填补缺失图像的高频子带细节。为了同步这两个分支的输出，第一分支的输出通过离散小波变换（discrete wavelet transform，DWT）分解，其低频子带与第二分支的输出一起通过离散小波逆变换（inverse discrete wavelet transform，IDWT）来重建出修复图像。两个分支网络的结构都是基于 U-net[9] 的类型结构和通用门控卷积（generalized gated convolution[10]，GConv）。U-net 结构是将编码器下采样的特征通过跳连接到解码器，再和解码器特征一起做上采样。通用门控卷积公式为：

$$Y_j = \phi(\sum_i W_i^f * X_i + b) \odot \sigma(\sum_i W_i^g * X_i + c) \tag{2}$$

式中：在一个通用门控卷积中，X_i 和 Y_j 是第 i 个输入特征图和第 j 个输出特征输图，W_i^f 和 W_i^g 是两个不同的卷积核，b 和 c 是两个不同的偏置，ϕ 为任意激活函数（本文为 ELU[11] 激活函数），σ 是 Sigmoid 激活函数。

1.2.1 内容网络分支

内容网络（G_{con}）的目标是使最终的修复图像（I_{com}）有合理的图像结构和语义内容，网络结构（G_{con}）是基于编码器-解码器的 U-net 结构，如图 1 所示。

编码器端总共有 4 组卷积，分别是 1 个步长为 1 的 5*5 门控卷积作为第一组和 3 个步长为 2 的 3*3 门控卷积作为其余三组。多级融合模块（MLFM）用来扩大感受野，在该模块中总共有 5 个卷积组，在前 4 个卷积组中，采用逐步增加空洞门控卷积数量和核大小的方式来扩大感受野；在第 2 组到第 4 组的末尾进行特征拼接，可以使具有不同大小感受野的特征融合在一起，使内容网络能够学习到更多不同尺度的内容；最后一个卷积组中的 3*3 门控卷积用于恢复输出的特征维度和输入的特征维度一致。

解码器端使用了 3 个上采样模块来将编码的特征恢复到输入的维度，该上采样模块中有两个输入，一个来自该模块的前一个输出特征，另一个来自跳连接的特征，它们首先拼接在一起，然后分别经过两种操作得到两个上采样的特征，分别是：一个步长为 2 的 4*4 反卷积和一个步长为 1 的 3*3 卷积组合，一个 2 倍双线性扩大插值和一个步长为 1 的 3*3 卷积组合。两个上采样特征在通道拼接后，最终经过一个步长为 1 的 3*3 门控卷积进行融合。

1.2.2 纹理网络分支

纹理网络（G_{txt}）的目标是对缺失区域的小波高频子带进行填充，因此网络的输入是 3 个带缺失区域小波高频子带的拼接。除了 MLFM 没有应用于纹理网络，取而代之的是两个步长为 1 的 3*3 门控卷积，其他网络结构与内容网络相似。

1.3 判别网络结构设计

DEII 采用了全局（D_{glb}）与局部（D_{loc}）结合的判别形式，并使用了 GAN 的变体 WGAN-GP[12] 来提高训练能力。D_{glb} 的输入是整张 256×256 图像，D_{loc} 的输入是包含修复区域的 128×128 局部区域，它们都是使用 4 个步长为 2 的 5*5 普通卷积进行下采样和 1 个全连接得到输出图像的判别值，每个卷积之后会进行非线性激活函数（LeakyRelu），结构如图 1 所示。

1.4 空间衰减掩膜

DEII 还使用了一种空间衰减掩膜（SD-mask），可以应

用于损失函数。计算方法为：

$$m'_{i,j}=\begin{cases}0, & m_{i,j}=0\\ \gamma^{t_{i,j}-1}, & m_{i,j}=1\end{cases} \tag{3}$$

式中：$m_{i,j}\in M$，$m'_{i,j}\in M_{\mathrm{sd}}$，$\gamma$ 设为 0.99，$t_{i,j}$ 是腐蚀次数，使用一个 3*3 的全 1 矩阵对 M 进行腐蚀，其中，M 的 1 像素将变为 0。这样，不需要计算每个缺失像素的距离，大大提高了实现效率。对于一个尺寸缩小的特征图，对应的 SD-mask 可以通过对全尺寸的 SD-mask 相应缩小得到，而无需再次执行公式（3）的操作。使用$M_{\mathrm{sd}}^{(1/2^i)}$来代表 SD-mask 的维度变为原来的 $1/2^{2i}$。

1.5 损失函数

修复网络（G）的总损失函数定义为：

$$L_G=\omega_1\times L_{\mathrm{con}}+\omega_2\times L_{\mathrm{dwt}}+\omega_3\times L_{\mathrm{per}}+\omega_4\times L_{\mathrm{sty}}+\omega_5\times L_{\mathrm{adv}}^{G} \tag{4}$$

式中：L_{con} 是内容损失，L_{dwt} 是小波子带损失，L_{per} 是感知损失，L_{sty} 是风格损失，L_{adv}^{G}是修复网络对抗损失。其中，ω_1=10，ω_2=10，ω_3=0.5，ω_4=1000，ω_5=0.001。

判别网络（D）的损失函数定义为$L_D=L_{\mathrm{adv}}^{D}$。其中，L_{adv}^{D}是判别网络的对抗损失。

1.5.1 内容损失

采用 L1 范数来衡量 SD-mask 指定的缺失区域中修复图像（I_{com}）和真实图像（I_{gt}）之间的像素内容损失（Content loss）：

$$L_{\mathrm{con}}=\left\|\left(I_{\mathrm{com}}-I_{\mathrm{gt}}\right)\odot M_{\mathrm{sd}}\right\|_1 \tag{5}$$

1.5.2 小波子带损失

采用小波子带损失（DWT loss）来保证纹理网络生成的小波高频子带和内容网络生成的小波低频子带与它们对应的真实小波子带相近。

$$\begin{aligned}L_{dwt}=&\left\|\left(S_{HL}^{t}-S_{HL}^{gt}\right)\odot M_{sd}^{(1/2)}\right\|_1+\left\|\left(S_{LH}^{t}-S_{LH}^{gt}\right)\odot M_{sd}^{(1/2)}\right\|_1+\\&\left\|\left(S_{HH}^{t}-S_{HH}^{gt}\right)\odot M_{sd}^{(1/2)}\right\|_1+\left\|\left(S_{LL}^{c}-S_{LL}^{gt}\right)\odot M_{sd}^{(1/2)}\right\|_1\end{aligned} \tag{6}$$

1.5.3 感知损失

采用感知损失（Perceptual loss）来保证修复图像（I_{com}）和真实图像（I_{gt}）之间在高级语义层次相近。使用在 ImageNet 数据集[13]上预训练好的 VGG19 模型来计算该损失：

$$L_{\mathrm{per}}=\sum_{i=1}^{4}\left\|\left(\varphi_{I_{\mathrm{gt}}}^{S_i}-\varphi_{I_{\mathrm{com}}}^{S_i}\right)\odot M_{\mathrm{sd}}^{(1/2^i)}\right\|_1 \tag{7}$$

式中：φ^{S_0}、φ^{S_1}、φ^{S_2}、φ^{S_3}和φ^{S_4}分别代表 VGG19 网络[14]中的 relu1_1 层、relu2_1 层、relu3_1 层、relu4_1 层、relu5_1 层。

1.5.4 风格损失

采用风格损失（Style loss）来保证修复图像（I_{com}）和真实图像（I_{gt}）之间在风格上相近。它是计算在特征图的格拉姆矩阵（Gram Matrix，$\mathcal{G}[\cdot]$）上：

$$L_{\mathrm{sty}}=\sum_{i=1}^{3}\left\|\mathcal{G}[\varphi_{I_{\mathrm{gt}}}^{Q_i}\odot M_{\mathrm{sd}}^{(1/2^i)}]-\mathcal{G}[\varphi_{I_{\mathrm{com}}}^{Q_i}\odot M_{\mathrm{sd}}^{(1/2^i)}]\right\|_1 \tag{8}$$

$$\mathcal{G}[x]=\frac{f(x)*f^{\mathrm{T}}(x)}{H_xW_xC_x} \tag{9}$$

式中：对于特征 x，H_x 代表高度，W_x 代表宽度，C_x 代表通道数。$f(x)$ 表示将特征 x 的形状 $(H\times W\times C)$ 重组为 $(H\times W)\times C$，$f^{\mathrm{T}}(x)$ 表示 $f(x)$ 的转置。φ^{Q_0}、φ^{Q_1}、φ^{Q_2} 和 φ^{Q_3} 分别代表 VGG19 网络中的 relu2_2 层、relu3_4 层、relu4_4 层、relu5_2 层。

1.5.5 对抗损失

D_{glb} 和 D_{loc} 采用 WGAN-GP 形式的对抗损失。假设 (x_r, x_g) 为一对数据对，表示真实数据（real）和生成数据（generated），它们分别从真实图像（I_{gt}）和修复图像（I_{com}）采样得到。判别网络的对抗损失（WGAN dloss）表示为：

$$\begin{aligned}L_{\mathrm{adv}}^{D}=&-E_{x_r}\left[D_{\mathrm{loc}}(\mathcal{C}[x_r])\right]+E_{x_g}\left[D_{\mathrm{loc}}(\mathcal{C}[x_g])\right]+\lambda E_{\hat{x}}\left[(\|\nabla_{\hat{x}}D_{\mathrm{loc}}(\mathcal{C}[\hat{x}])\|_2-1)^2\right]\\&-E_{x_r}\left[D_{\mathrm{glo}}(x_r)\right]+E_{x_g}\left[D_{\mathrm{glo}}(x_g)\right]+\lambda E_{\hat{x}}\left[(\|\nabla_{\hat{x}}D_{\mathrm{glb}}([\hat{x}])\|_2-1)^2\right]\end{aligned} \tag{10}$$

式中：E 表示求期望值，$\mathcal{C}[\cdot]$ 表示从完整尺寸图像中截取（如 128×128）局部区域。$\hat{x}=\epsilon x_r+(1-\epsilon)x_g$，$\epsilon\sim U[0,1]$。$\lambda$=10。

修复网络的对抗损失（WGAN gloss）为：

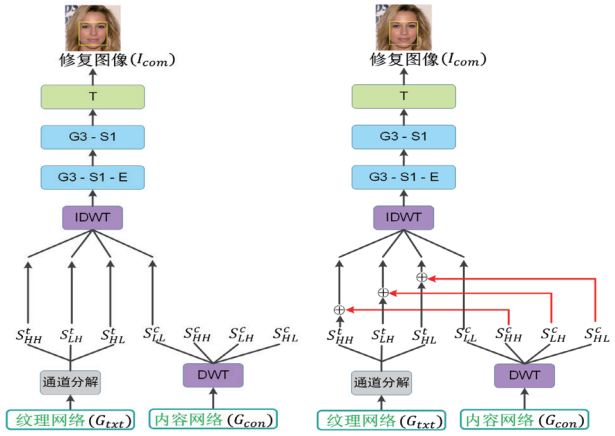
$$L_{\mathrm{adv}}^{g}=-E_{x_g}\left[D_{\mathrm{loc}}\left(\mathcal{C}[x_g]\right)\right]-E_{x_g}\left[D_{\mathrm{glo}}\left(x_g\right)\right] \tag{11}$$

2 基于离散小波变换的细节增强的图像修复优化方法

DEII 方法虽然实现了修复图像细节信息的增强，但其修复网络结构中有比较多的冗余高频信息没有被使用，该部分信息应加以利用，以进一步增强修复图像细节；并且该算法对于一些场景的修复，依旧存在语义结构扭曲、纹理模糊等问题，应改进判别网络结构，以提高修复图像的质量。

2.1 修复网络的优化结构

修复网络的优化结构如图 2 所示。



（a）原先方法　　（b）本文优化方法

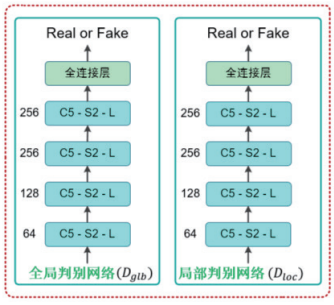
图 2 改进的修复网络结构

注：“Ga-Sb-E”表示 a*a 门控卷积（GConv）—步长 b—ELU 激活函数。“T”表示 Tanh 激活函数。“DWT”和“IDWT”分别为离散小波变换和逆变换，其小波基为 Daubechies2(db2)，⊕表示对应位置元素相加 (add)

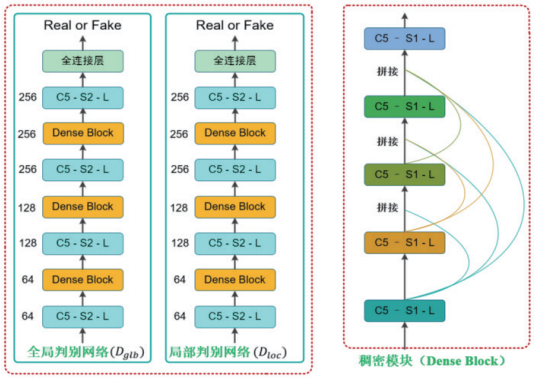
左图（a）为原先 DEII 的图像修复结构，其中内容网络（G_{con}）生成的三个高频小波子带（S^c_{HL}，S^c_{LH}，S^c_{HH}）没有被加以利用，而这些小波子带也是经由自编器结构的内容网络生成，里面的缺失区域也已被填补完成，该部分信息应加以利用，以进一步增强修复图像细节。右图（b）为本文所改进后的结构，采用求和方法（add）将其加到对应纹理分支中生成的三个高频小波子带（S^t_{HL}，S^t_{LH}，S^t_{HH}），以进一步增强生成的三个高频小波子带的纹理细节。最后，三个增强的高频小波子带再与内容分支填补完成的低频子带（S^c_{LL}）一起通过小波逆变换得到重建后的修复图像。

2.2 判别网络的优化结构

如图 3 所示，左图（a）为原先 DEII 算法的图像修复结构中的判别网络结构（D），采用了全局（D_{glb}）与局部（D_{loc}）结合的判别形式，该结构与修复网络（G）一起进行对抗训练，最终训练内容网络修复出高质量的图像，但对于一些场景的修复，依旧存在语义结构扭曲、纹理模糊等问题，因此应改进生成判别网络结构，以提高修复图像的质量。本文设计一个基于稠密神经网络（DenseNet）的判别网络结构， 如图 3 右图（b）所示，D_{glb} 的输入是整张 256×256 图像，D_{loc} 的输入是包含修复区域的 128×128 局部区域，它们的结构相同，使用的是普通卷积进行下采样和稠密模块（Dense Block）的结合，最后再经过 1 个全连接得到输出图像的判别值。



（a）原先判别网络结构



（b）本文优化的判别网络结构

图 3 改进的判别网络结构

注：“Ca-Sb-L”表示 a*a 普通卷积（VConv）—步长 b—LeakyRelu 激活函数

其中，下采样卷积为步长为 2 的 5*5 普通卷积，稠密模块（Dense Block）结构如图 3 右图（2）所示，是一个包含 5 层卷积特征图的结构，层与层之间采用密集连接方式，具体来说就是每一层的输入都来自它前面所有层的特征图，每一层的输出均会直接连接到它后面所有层的输入，卷积大小都为 5*5，步长为 1，每个卷积之后会进行非线性激活函数（LeakyRelu）。使用稠密模块能有效缓解梯度消失，使特征传递更加有效，网络更易于稳定训练。

判别网络最小化对抗损失函数以区分真实图像和修复图像，而修复网络最大化对抗损失函数以迷惑判别网络的识别，生成更加逼真的修复图像。它们以对抗性的方法进行训练，相互对抗，以提高生成图像的质量。此外，全局和局部的结合判别是用来确保修复内容在全局图像内容和局部图像内容之间具有语义一致性，增强修复效果。

2.3 其他结构

其他结构、参数、损失函数等和 DEII 图像修复方法保持一致，详情请参考第一部分 DEII 方法简介。

3 实验结构及分析

在这一节，本文通过一些实验来验证所提出的优化方法能有效改进 DEII 的图像修复方法，提高其修复图像的质量。总共进行三个实验，第一部分实验是实验设置，包括数据集、对比方法、度量指标和实现细节的说明；第二部分实验是本文方法与对比方法在每个数据集上的性能对比结果，包括定性指标和定量指标，即修复视觉效果对比和度量指标对比情况；第三部分实验是对本文提出的优化方法的消融实验，包括改进修复网络结构的消融实验和改进判别网络结构的消融实验。

3.1 实验设置

3.1.1 数据集

本文选择了图像修复任务中常用的 3 个公开数据集，分别是人脸图像数据集（CelebA-HQ[15]）、复杂纹理图像数据集（describable textures dataset，DTD[16]）和建筑图像数据集（Facade[17]）。前三个数据集按总图像数量的 70%、20% 和 10%，将它们划分成训练集、测试集和验证集。详细数量情况如表 1 所示。

表 1 3 个数据集的划分情况

数据集	图像内容类型	训练集 / 张	测试集 / 张	验证集 / 张
CelebA-HQ	人脸图像	21 000	6000	3000
DTD	复杂纹理图像	3947	1132	564
Facade	建筑图像	424	122	60

3.1.2 对比方法

本研究选择了 3 种深度图像修复方法进行对比，即 ContextEncoder（CE）、Contextual Attention（CA）[18]、Detail-Enhanced Image Inpainting（DEII）。将这些方法在 3 个数据集上进行训练，并使用和本文相同的验证标准在验证集上选出最佳的模型。

3.1.3 度量指标

本文采用了 4 个常用的度量指标来评价本文方法与其他方法的优缺点：平均绝对值损失（mean absolute loss，MAE）、峰值信噪比（peak signal to noise ratio，PSNR）、结构相似性（structural SIMilarity，SSIM）、Frechet Inception Distance（FID）[19]。

3.1.4 实现细节

在训练阶段，本文算法的目标是在一幅 256×256 图像内填补一个具有任意形状和随机位置的 128×128 缺失区域。在测试阶段，本文方法还可以同时修复多个缺失区域，而不局限于 128×128 区域。在图像预处理阶段，对于 CelebA-HQ 数据集，直接将它们缩放为 256×256 尺寸；对于其他数据集，将每张图像的最长边随机裁剪成和最短边一样长，再将这张图像缩放为 256×256 尺寸。训练时，本文采用不规则缺失区域进行训练，它的生成方法如文献 [10] 所述，本文方法有全局和局部判别网络的结合，因此是在图像的 128×128 区域内生成不规则缺失区域。对于一批量（batchsize）训练图像，它们的 128×128 缺失区域位置是一样的，而对于不同的一批批训练图像，位置是不同的。对于其他超常数，本文使用 Adam 优化器的梯度优化算法，它的参数 β_1=0.5，β_2=0.9。学习率为 0.000 1，批量大小为 2。网络总训练的迭代次数（iteration）为 100 万次，每 5 万次模型会使用验证集图像计算 DISTS 数值，从而选出最佳的训练模型，验证图像的缺失区域都是中间 128×128 正方形区域。本文使用 Daubechies2（db2）小波基，整个网络是使用 PyTorch 实现，然后在 NVIDIA RTX 4060（3 GB 显存）硬件环境训练。

3.2 实验结果

3.2.1 定性实验

本文首先进行了定性修复效果的对比实验，将修复中心 128×128 平方区域的视觉效果与对比方法进行比较，如图 4。可以看出，CE 算法和 CA 算法修复效果较差，而对比 DEII 算法，本文改进后的方法可以生成更精细的细节，如人脸（即 CelebA-HQ 数据集）中更完善的眼睛。在复杂图像（即 DTD 数据集）和建筑图像（即 Facade 数据集）中，本文改进后的方法可以生成结构合理的修复图像。总的来说，本文方法在整体内容语义和结构方面表现出更好的修复质量。

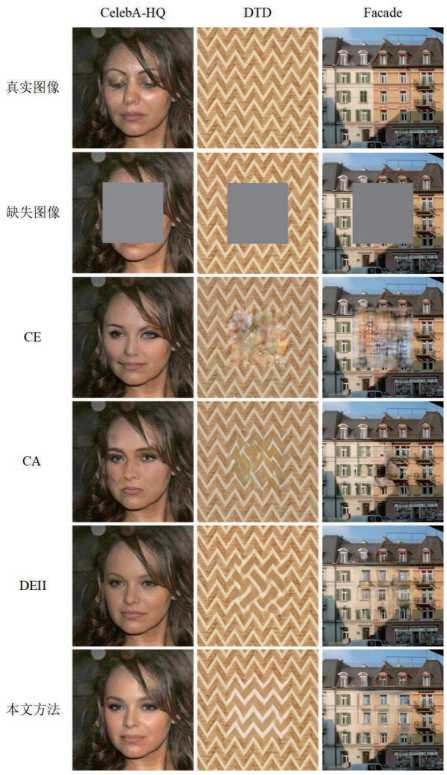


图 4 本文方法与对比方法的定性修复效果

3.2.1 定量实验

本文进行定量分析，具体来说，使用测试数据集计算 4 个评估指标，其中修复区域设置为中心 128×128 平方区域。从表 2 中可以看出，本文方法在 SSIM 和 FIM 这两个最先进的图像质量评估指标方面对所有数据集都取得了最好的结果，这意味着本文方法修复的图像具有更好的合理结构，这也与图 4 进行的定性修复效果性能一致，本文方法在结构性方面表现出优秀的修复效果。另外，本文方法在 CelebA-HQ、DTD 和 Facde 数据集的大多数其余指标上也表现良好。

表 2 本文方法与其他对比方法在 4 个数据集的定量指标表现情况

数据集	方法	PSNR↑	SSIM↑	MAE↓	FID↓
CelebA-HQ	CE	26.03	0.828 6	4.840	5.334
	CA	23.99	0.797 2	6.014	6.220
	DEII	26.17	0.839 0	**4.371**	5.168
	本文	**26.46**	**0.888 0**	5.122	**5.156**
DTD	CE	20.40	0.727 8	15.21	77.01
	CA	22.80	0.767 6	7.739	33.89
	DEII	**23.60**	0.773 2	6.981	32.97
	本文	22.37	**0.812 6**	**6.716**	**32.44**
Facade	CE	20.04	0.731 6	13.91	83.53
	CA	22.87	0.808 2	6.876	32.47
	DEII	**23.15**	0.797 3	**6.788**	28.21
	本文	22.89	**0.840 1**	7.252	**27.30**

注：↑ /↓ 表示数值越高 / 越低越好

4 总结

本文主要是对现有基于小波变换的深度图像修复方法存在的不足进行优化，一方面是改进原来修复网络，用求和方法将原先内容网络生成的 3 个高频子带加到对应纹理分支中生成的三个高频小波子带，以进一步增强生成的 3 个高频小波子带的纹理细节；另一方面，在此基础上，设计一个基于稠密神经网络结构（DenseNet）的判别网络，改进原有判别网络，以稳定训练，提高修复图像的质量。通过与其他修复算法进行定性和定量的对比实验，结果表明，本文改进后的算法可以生成细节更加完善、结构更加合理的修复图像。

参考文献：

[1]BERTALMIO M, SAPIRO G, CASELLES V, et al.Image inpainting[EB/OL].(2020-07-01)[2024-02-20].https://doi.org/10.1145/344779.344972.

[2]PATHAK D, KRAHENBUHL P, DONAHUE J, et al.Context encoders:feature learning by inpainting[C]//29th IEEE Conference on Computer Vision and Pattern Recognition. Piscataway:IEEE,2016:2536-2544.

[3]GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al.Generative adversarial networks[C]//2023 14th International Conference on Computing Communication and Networking Technologies(ICCCNT).Piscataway: IEEE, 2014: 139-144.

[4]WU Z, HOU B, JIAO L.Multiscale CNN with autoencoder regularization joint contextual attention network for sar image classification[J].IEEE transactions on geoscience and remote sensing, 2021,59(2):1200-1213.

[5]IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion[J].ACM transactions on graphics, 2017,36(4):1-14.

[6]ZENG Y, FU J, CHAO H, et al.Learning pyramid-context encoder network for high-quality image inpainting[C]//2019 IEEE/CVP Conference on Computer Vision and Pattern Recognition, [v.3].Piscataway:IEEE,2019: 1486-1494.

[7]LI B, ZHENG B, LI H, et al.Detail-enhanced image inpainting based on discrete wavelet transforms[J].Signal processing,2021,189:108278.

[8]DENG Y, HUI S, MENG R, et al.Hourglass attention network for image inpainting[C]//Computer Vision-ECCV 2022,p.18. Cham:Springer,2022:483-501.

[9]RONNEBERGER O, FISCHER P, BROX T.U-net:convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention--MICCAI 2015,Part III.Cham: Springer, 2015: 234-241.

[10]YU J, LIN Z, YANG J, et al.Free-form image inpainting with gated convolution[C]//2019 IEEE/CVF International Conference on Computer Vision.Piscataway: IEEE, 2019: 4470-4479.

[11]DJORK-ARNÉ C, UNTERTHINER T, HOCHREITER S.Fast and accurate deep network learning by exponential linear units(ELUs)[EB/OL].(2015-11-23)[2024-02-21].https://doi.org/10.48550/arXiv.1511.07289.

[12]GULRAJANI I, AHMED F, ARJOVSKY M, et al.Improved training of wasserstein gans[EB/OL].(2017-03-31)[2024-02-21]. https://doi.org/10.48550/arXiv.1704.00028.

[13]DENG J, DONG W, SOCHER R, et al. Imagenet:a large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition.Piscataway: IEEE, 2009:248-255.

[14]SIMONYAN K, ZISSERMAN A.Very deep convolutional networks for large-scale image recognition[C]//2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR). Piscataway: IEEE,2015:730-734.

[15]KARRAS T, AILA T, LAINE S, et al.Progressive growing of gans for improved quality,stability,and variation[EB/OL].(2017-10-27)[2024-02-15].https://doi.org/10.48550/arXiv.1710.10196.

[16]CIMPOI M, MAJI S, KOKKINOS I, et al. Describing textures in the wild[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition.Piscataway:IEEE,2014:3606-3613.

[17]TYLECEK R, SARA R.Spatial pattern templates for recognition of objects with regular structure[C]//Pattern Recognition. Berlin:Springer,2013:364-374.

[18]YU J, LIN Z, YANG J, et al.Generative image inpainting with contextual attention[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition,[Volume 8 of 13]. Piscataway: IEEE, 2018:5505-5514.

[19]HEUSEL M, RAMSAUER H, UNTERTHINER T, et al.Gans trained by a two time-scale update rule converge to a local nash equilibrium[EB/OL].(2017-06-26)[2024-02-25].https://doi.org/10.48550/arXiv.1706.08500.

【作者简介】

郑博伟（1995—），通信作者（email: 734249947@qq.com），男，广东揭阳人，硕士，助教，研究方向：物联网技术、人工智能、图像处理等。

李宗辉（1982—），男，山东郯城人，硕士，副教授，研究方向：网络工程、机器学习。

陈锐彬（1982—），男，广东揭阳人，硕士，高级工程师，研究方向：数字图像处理；

黄梅佳（1993—），男，广东揭阳人，硕士，助教，高级工程师，研究方向：软件工程、模式识别。

（收稿日期：2024-03-15）