

# 基于高效注意力的密集残差卷积修复算法

李 然, 钟杏苑, 陈 泉, 李海燕\*

云南大学信息学院, 云南 昆明

收稿日期: 2023年4月27日; 录用日期: 2023年5月24日; 发布日期: 2023年5月31日

## 摘 要

为解决修复人为破损图像大面积不规则缺失情况时, 修复结果存在的结构不合理、伪影、失真等问题, 提出了一种基于高效注意力的密集残差图像修复算法。首先, 将破损图像进行下采样, 捕获图像细节特征, 将其传递至包含空洞卷积的密集残差网络, 增大信息感受野, 提升修复细节特征的细粒度。同时增加高效注意力机制, 使用特征图加权聚合生成全局上下文向量, 对输入特征进行全局描述。此外, 引入超图卷积模块, 用于捕捉全局信息并增强修复细节的完整性。然后进行上采样, 还原图像的结构纹理信息。最后将生成的图像输入SN-Patch GAN鉴别器进行判别优化, 进一步提升修复效果。在公开数据集上进行训练和测试, 实验结果表明: 本文算法可以实现大面积随机缺失图像的修复, 有效捕捉全局上下文信息, 修复结果具有平滑过渡边界和清晰细节, 满足视觉连贯性及真实性, 在修复的视觉效果、峰值信噪比、结构相似度和平均误差方面, 均优于对比算法。

## 关键词

图像修复, 高效注意力, 超图卷积, 密集残差网络

# Dense Residual Convolution Inpainting Algorithm Based on Efficient Attention

Ran Li, Xingyuan Zhong, Quan Chen, Haiyan Li\*

School of Information Science & Engineering, Yunnan University, Kunming Yunnan

Received: Apr. 27<sup>th</sup>, 2023; accepted: May 24<sup>th</sup>, 2023; published: May 31<sup>st</sup>, 2023

## Abstract

In order to solve the problems of unreasonable structure, artifact and distortion of inpainting results in the case of irregular images and large-area missing, a dense residual image inpainting al-

\*通讯作者。

gorithm based on efficient attention is proposed. Firstly, the damaged image is downsampled to gradually reveal the detailed information of the image features, which is then passed to the dense residual network containing the dilated convolution to increase the information receptive field and captures more detailed texture information. At the same time, an efficient attention mechanism is added, and the weighted aggregation of feature maps is used to generate a global context vector, which is used to reflect the global description of the input features. Furthermore, a hypergraph convolution module is introduced to capture global information and enhance the inpainting fine-grainedness. Thereafter, upsampling is performed to restore the structural texture information of the image. Finally, the generated image is fed into the SN-Patch GAN discriminator for discriminant optimization to further improve the repaired effect. Training and testing are carried out on public datasets. The experimental results show that the proposed algorithm can achieve large-area missing image inpainting and effectively capture the global context information and enhance the fine-grained inpainting, which the inpainting results have smooth boundaries and clear details, satisfying visual coherence and authenticity, and are superior over the state of the art in terms of visual effects, peak signal-to-noise ratio, structural similarity and average error.

## Keywords

Image Inpainting, Efficient Attention, Hypergraph Convolution, Dense Residual Network

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

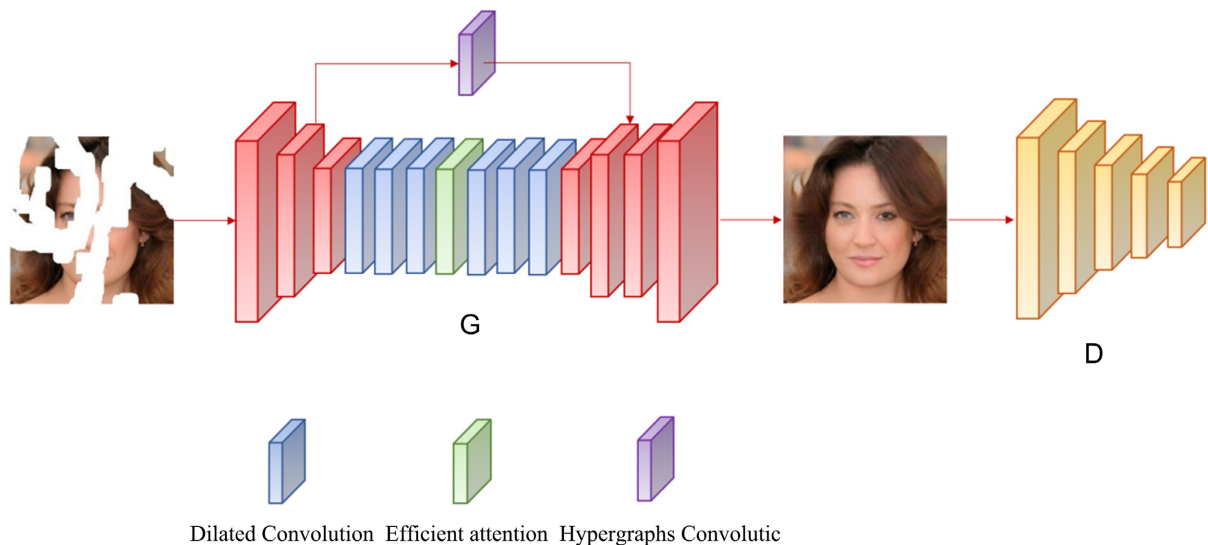
图像修复是根据残缺图像的完好部分对图像缺失信息进行重建,使修复图像与原始图像具有相似结构的一项技术。现有的图像修复方法主要有两类:传统方法和基于深度学习的方法。传统方法中,基于偏微分方程的方法利用物理学中的热扩散方程[1] [2] [3],将缺失区域相邻信息传播到缺失区域中,该方法对细小损伤有较好的修复效果,但当受损面积大时修复效果差。当图像大量信息丢失时,多采用基于纹理合成的修复方法[4] [5] [6],通过模仿图像其他区域的纹理,生成纹理片段填充缺失区域。此种方法不能捕获图像的语义信息,易导致图像不连贯。

为弥补传统方法的不足,研究者们进一步探索基于深度学习的算法。Pathak 等[7]提出 Context Encoder 网络,在类似“沙漏”的网络结构中增加对抗损失,使得生成的图像具有合理结构,但该算法的修复结果存在伪影和边缘模糊的问题。为解决上述不足, Li 等[8]在此基础上增加了一项新的对抗损失,将修复区域补全到原图后再判断整个修复图像的真伪,保证生成的内容更加合理的同时,维持边界像素值的连续性,但此种方法仅能处理背景较简单的规则缺失。此后, Yu 等[9]提出基于内容感知层的前馈生成网络,将修复过程分成粗修复及细修复两阶段,可有效修复大面积规则缺失,可以生成语义合理的结果,但在修复随机孔洞时,缺乏整体与局部信息的语义一致性。为有效修复随机不规则缺失区域, Liu 等[10]提出用部分卷积代替普通卷积的修复算法,利用有效像素自动掩码更新机制推断并修复缺失区域,能处理背景简单、随机不规则的小面积缺失区域,但修复边缘数据稀疏的区域中,修复质量降低。其后, Nazeri 等[11]提出了一个包括边缘生成器和图像补全网络的二阶段生成对抗网络,用边缘生成器预测生成边缘,再使用预测边缘作为先验填充缺失区域,该网络在背景简单时可得到逼真的细节纹理,但无法重建大面积不规则缺失图像。Zheng 等[12]提出一种多元图像修复算法(PIC),采用重建路径与生成路径并行的结构,

可修复较大面积不规则缺失, 但该算法在复杂背景下会出现结构变形且色彩不协调的问题。为修补背景复杂的大面积缺失区域, Peng [13]等提出基于分层的多解修复算法(VQ-VAE), 其运用条件自回归网络和结构注意力模块生成具有高度多样性的合理结构, 但缺失区域面积增大至 30%及以上时, 该算法的修复结果包含纹理细节失真和局部模糊。Jam [14]等提出反向掩码网络(R-MNet), 将 Wasserstein GAN 与一种新型反向掩码算子相结合, 可捕捉数据分布并生成兼具逼真与连贯性的图像, 但在修复背景复杂且缺失区域达 30%及以上缺失时, 生成图像边缘过渡模糊, 视觉质量下降。此后, Suvorov 等[15]利用快速傅立叶卷积对具有全局感受域的频域特征进行编码, 捕捉图像高感受野以获取图像整体结构, 该算法可修复大面积随机破损的图像, 但主要提取纹理特征, 损失大量的结构信息, 因而生成图像结构细节模糊、清晰度低且掩码处含有色彩伪影。

综上所述, 现有图像修复算法在处理人为破损图像中含复杂背景的大面积不规则缺失区域时, 难以把握整体结构的完整度和纹理细节的清晰度, 生成的图像存在结构不合理, 边缘模糊, 分辨率较低等问题。为克服此缺陷, 本文仅针对人为破损图像, 提出一种基于高效注意力的密集残差图像修复算法, 主要创新点在于: 1) 通过下采样提取破损区域的细节特征, 将其传递至包含空洞卷积的密集残差网络, 增大信息感受野的同时, 捕获更多的纹理细节信息, 提升了在含有大面积区域缺失的残缺图像上的修复效果。2) 引入了高效注意力机制, 通过(Key, Query, Value)的三元组有效捕捉全局上下文信息, 将通道按特征图加权聚合生成全局上下文向量, 描述输入的全局特征, 并增强修复纹理的细粒度, 增强了在纹理细节复杂的残缺图像上的修复效果。3) 引入超图卷积模块, 用于捕捉全局结构信息, 复原图像的结构纹理信息。

## 2. 基于高效注意力的密集残差卷积图像修复



**Figure 1.** The proposed algorithm model

**图 1.** 提出算法模型

基于高效注意力的密集残差卷积图像修复算法如图 1 所示: 首先, 对人为破损图像进行下采样操作, 即在生成对抗网络中设置卷积核大小对破损图像卷积得到 FeatureMap, 再采用基本的区域归一化并调节池化层步长对其进行特征的粗提取, 至此完成下采样过程。该部分主要负责降低图像的特征维度并保留精简的特征信息, 然后, 将提取出的特征信息输入包含空洞卷积的密集残差生成网络, 此密集残差生成

网络中，外层是包含三层密集残差卷积的多尺度融合空洞卷积块组合的卷积网络，内层是包含着高效注意力机制的卷积网络，超图卷积模块与之平行，连接第二层与倒数第二层的卷积块，其中，密集残差部分主要负责细化提取特征，进一步扩大感受野，捕获更加精细的特征纹理信息。同时，通过增加高效注意力机制，使用特征图的加权聚合有效捕获上下文信息。然后将其输入到超图卷积模块，修复多元关联对象的关系以更准确地描述全局语义信息并减小语义信息丢失。最后，将生成图像输入 SN-Patch GAN 鉴别器进一步判别优化，加强修复图像的结构完整性和纹理真实性。

## 2.1. 高效注意力机制

注意力机制对于学习图像修复中跨空间位置的模式具有重要意义[9] [16] [17]，有利于主干特征提取网络对颜色、线条等浅层语义特征的提取。该模块的主要作用是通过学习通道间的相关性，为每个通道分配权重，从而使网络更关注关键语义特征。然而，注意力模块的计算成本很高，且在图像修复中不易并行化[10]。为有效捕获远程特征并保证修复模型的效率，提出算法引入高效注意力(Efficient Attention)模块。首先将输入特征  $X \in R^{h \times w \times d}$  和掩码  $M \in R^{h \times w \times 1}$  重构为  $R^{h \times w \times d}$  和  $R^{h \times w \times 1}$ ，分别作为矢量数据。从而可将高效注意力的过程记为：

$$\begin{cases} Q = \text{softmax}_{\text{row}} [W_q X + (M \bullet -\infty)] \\ K = \text{softmax}_{\text{col}} [W_k X + (M \bullet -\infty)] \\ V = W_v X \\ E = Q(K^T V) \end{cases} \quad (1)$$

上式中， $W_{q,k,v}$  表示由输入像素矩阵  $X$  生成的学习参数矩阵，分别对应 query、key、value 三个特征，其中： $+(M \bullet -\infty)$  表示在 softmax 操作之前对输入图像屏蔽缺损部分的像素。 $\text{softmax}_{\text{row}}$  和  $\text{softmax}_{\text{col}}$  分别表示对输入的行和列应用柔性最大值传输操作，得到非线性归一化的处理结果。然后将输出  $E \in R^{h \times w \times d}$ ，重新整形为  $R^{h \times w \times d}$ 。提出算法使用的是普通注意力操作的近似值[18]，输出  $E$  是由  $K^T V \in R^{d \times d}$  而不是由标准的  $QK^T \in R^{h \times w \times h \times w}$  计算所得，即  $(QK^T)V$  的计算顺序变换为  $Q(K^T V)$ ，将计算复杂度从  $O(n^2)$  转变为  $O(d_k * d_v)$ 。由于  $d_k * d_v$  实际上永远小于  $n^2$ ，复杂度呈线性，该方法可以显著降低计算成本。在实际计算机应用中，维度  $d$  远小于  $hw$ ，因此提出算法引入多头注意力[19]进一步将维度从  $d$  减小到  $d' = d/n_{\text{head}}$ 。因此， $K^T V$  可在特征级聚合查询  $Q$  并获得全局上下文  $E$ 。Efficient Attention 模块是受激励的，以掩盖缺损区域的注意力分数而聚合未缺损区域的特征。

## 2.2. 超图卷积

简单图的每个边仅连接两个节点，能基本表示数据之间的成对关系，但很难描述图像空间特征之间的复杂关系，因此提出算法使用超图卷积捕捉缺失区域的全局信息并增强修复细节的完整性。

令  $G(V, E, \omega)$  为由有限顶点集  $V$ 、超边集  $E$  和超边的权重  $\omega$  组成的超图。每个超边都有一个权重  $\omega(e)$ 。超边  $G$  表示为  $|V| \times |E|$  关联矩阵  $H$  方程式，其中每条边定义为：

$$h(v_i, e_j) = \begin{cases} 1, & \text{当 } v_i \in e_j \\ 0, & \text{其他} \end{cases} \quad (2)$$

基于关联矩阵  $H$ ，顶点的度数可以定义为  $d(v) = \sum_{e \in E} \omega(e) h(v, e)$ ，即对每一个边度进行加权求和；超边的度数定义为  $\delta(e) = \sum_{v \in V} h(v, e)$ ，即对每一个节点度加权求和。其中， $D_v$  和  $D_e$  分别表示顶点度和超边度的对角矩阵， $W$  表示超边权重的对角矩阵。

节点特征通过求和操作构建出所连接的超边特征，相应地，超边特征经过聚合后重新应用于节点特征的更新。用于超图学习的归一化拉普拉斯方法可表示为：

$$\arg \min \Omega(f) + \lambda R_{emp}(f) \quad (3)$$

其中， $f$  表示要学习的相关性分数， $\Omega(f)$  表示归一化成本函数， $R_{emp}(f)$  是经验损失， $\lambda$  是用于平衡正则化器和经验损失的参数。上式中通过对  $\Omega(f)$  最小化操作，共享多数偶发超边的顶点可以计算出相似性质的相关分数。整理得到超图的正则化器定义为：

$$\Omega(f) = \frac{1}{2} \sum_{e \in E} \sum_{u, v \in e} \frac{\omega(e) h(u, e) h(v, e)}{\delta(e)} \times \left( \frac{f(u)}{\sqrt{d(u)}} - \frac{f(v)}{\sqrt{d(v)}} \right)^2 \quad (4)$$

在此对正则器作出简化：令  $\Theta = D_v^{-\frac{1}{2}} H W D_e^{-1} H^T D_v^{-\frac{1}{2}}$ ，代表超边特征聚合预测出节点在下一回合特征表示的操作，再令  $\Delta = I - \Theta$ ，则正则化器可以重写为  $\Omega(f) = f^T \Delta f$ ，其中超图拉普拉斯算子  $\Delta$  是一个半正定矩阵。

### 2.3. 损失函数

#### 1) 生成器损失函数

定义生成器损失函数  $L_G$  评估预测图像的两个方面：缺失像素区域的质量和整个图像的感知质量。围绕这两个指标构建  $L_G$  确保生成器生成准确的缺失像素。使用特征空间鼓励对抗性训练生成具有相似特征的图像，从而获得更真实的结果。提出算法的损失函数将逐像素对比改进为用 VGG 提取高级特征，计算损失模型输入和输出的平方差作为感知损失[20] ( $L_p$ )，如式(5)：

$$L_p = \frac{1}{k} \sum_{i \in \phi} \left( \varphi_i [I] - \varphi_i [I_{pred}] \right)^2 \quad (5)$$

其中  $k$  是  $\phi$  的大小， $\varphi_i [I]$  是使用  $I$  作为输入运行 VGG19 的前向传递获得的特征， $\varphi_i [I_p]$  是对生成器  $G_\theta [I_M]$  的输出运行前向传递获得的特征。

定义反向掩码损失 ( $L_{rm}$ )，仅针对掩码区域创建的有效特征进行重建。反向掩码损失对应图像上掩码区域相应像素和掩码图像重建像素的平方差，使用反向掩码 ( $M_r$ ) 和原始图像 ( $X_{input}$ ) 来获得  $L_{rm}$ ，其中，

$$L_{rm} = \frac{1}{k} \sum_{i \in \phi} \left( \varphi_i [I_{Mpred}] - \varphi_i [\Theta M_r] \right)^2 \quad (6)$$

最后通过线性组合感知损失  $L_p$  和反向掩码损失  $L_{rm}$ ，得到生成器损失函数：

$$L_G = (1 - \lambda) L_p + \lambda L_{rm} \quad (7)$$

其中  $\lambda \in [0, 1]$ ，通过最小化缺失区域的误差来匹配与真实情况相当的预测，从而对特征进行最佳评估。

#### 2) 判别器损失函数

使用 Wasserstein 距离损失函数 ( $L_w$ ) 训练网络，对其定义如式(8)：

$$L_w = E_{I \sim P_x} [D_\theta(I)] - E_{I_{pred} \sim P_z} [D_\theta(I_{pred})] \quad (8)$$

其中， $E_{I \sim P_x} [D_\theta(I)]$  是真实数据分布的概率， $E_{I_{pred} \sim P_z} [D_\theta(I_{pred})]$  是生成的数据分布。

### 3. 实验结果与分析

在 CelebA-HQ 和 Paris-street 数据集上，将提出算法与现有三种算法进行实验对比。CelebA-HQ 和 Paris-street 数据集分别包含高清人脸和街景图像。设备参数为 CPU: Intel i7-9700K, GPU: RTX2080Ti。



代码在 Pytorch 深度学习框架下运行, 将提出算法与 PIC 算法[12]、VQ-VAE 算法[13]和 RM-Net 算法[14]进行对比, 掩码为人工设置的不规则掩码, 掩码大小为 40%~60%。生成网络的学习率为  $1 \times 10^{-4}$ , 鉴别器的学习率为  $1 \times 10^{-1}$ 。

### 3.1. 主观实验结果

为了保证比较的公平性, 所有模型的迭代次数相同。

图 2 是不规则掩码大小为 40%~60%时, 各算法分别在 Celeba-HQ 数据集及 Paris-street 数据集上的修复效果, 直观比较细节已在图中标出。

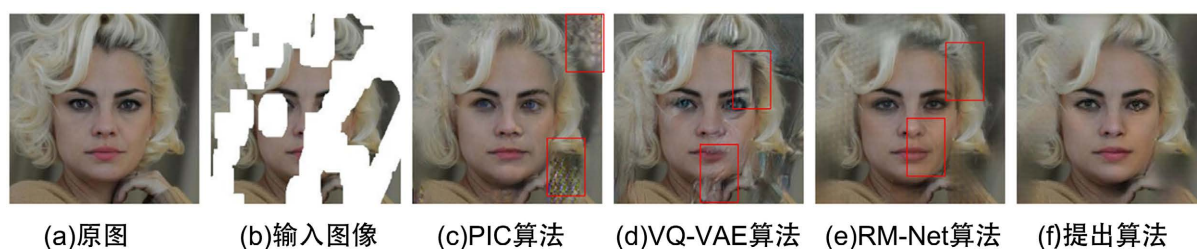


Figure 2. The comparison experiment 1 of CelebA-HQ

图 2. CelebA-HQ 对比实验 1

图 2(a)为原图, 图 2(b)是掩码图, 缺失区域达到前景的 45%左右, 且为随机不规则缺失。图 2(c)~(f)分别为 PIC 算法、VQ-VAE 算法、RM-Net 算法以及本文算法的修复结果。由图可见 PIC 算法能生成较为合理的面部结构, 但图像边缘出现明显的修复斑点。VQ-VAE 算法生成的人像存在大面积修复痕迹, 且右侧嘴角结构扭曲。RM-Net 算法的修复效果整体较好, 但嘴角及头发边缘处存在较大面积伪影, 修复细节不够清晰。相对而言, 提出算法修复结果更接近原始图像, 且结构完整, 面部细节纹理清晰。

图 3(b)中随机缺失区域面积超过 50%, 且主要缺失集中, 覆盖右侧主体区域, 存在局部细节的修复困难。PIC 算法的修复结果结构合理, 但右侧头发细节丢失, 图像四周浮现出大块修复斑点。VQ-VAE 算法的修复结果信息缺失, 右眼的结构不合理, 眼睛和嘴角存在不同程度的形态扭曲。RM-Net 算法生成的图像相对完整, 但头发出现大面积可见斑块, 鼻翼和嘴角的纹理细节较为扭曲, 同时眼睛的修复效果也比较模糊。而本文提出算法不仅能生成合理的结构, 且能兼顾整体与细节的语义一致性, 眼睛和头发细节的修复结果较其他算法更为准确细致。

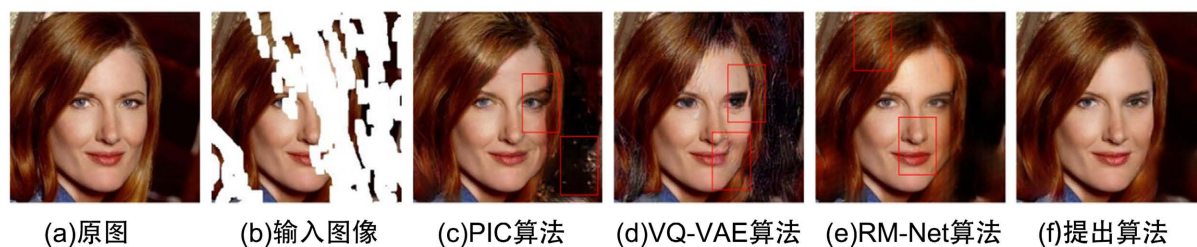


Figure 3. The comparison experiment 2 of CelebA-HQ

图 3. CelebA-HQ 对比实验 2

图 4(b)的掩码区域遮挡了人脸的 60%以上, 且掩码分布较为分散。从修复结果看: VQ-VAE 算法修复效果差, 整体存在大面积的修复斑点和不均匀色块, 眼部结构较为不合理。RM-Net 算法的修复结果能保持较好的整体一致性, 但同样是眼睛细节不够清晰, 且额头处存在明显色差斑块。PIC 算法能正确复

原图像纹理，面部细节也较为精细，但发梢边缘存在大面积模糊光影，轮廓语义丢失。相对而言，使用提出算法修复得到的整体结构更为真实精致，颜色一致性保持较好。

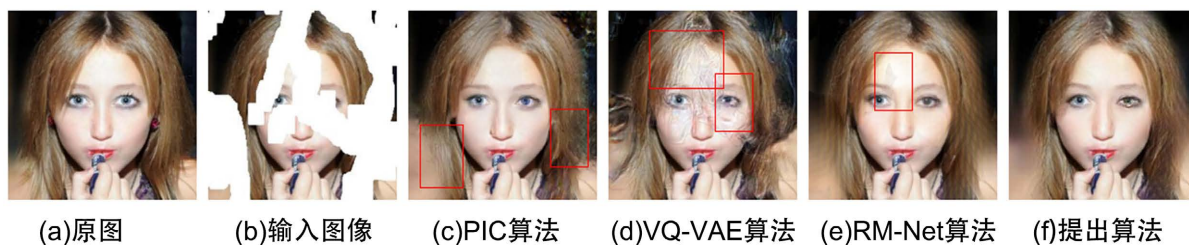


Figure 4. The comparison experiment 3 of CelebA-HQ

图 4. CelebA-HQ 对比实验 3

图 5 是对 Paris-street 的修复结果对比。图 5(b)中部分缺失面积达到 85%左右，且背景复杂，包含大量纹理细节。从实验结果可看出：VQ-VAE 算法的修复效果较差，大理石柱的轮廓及建筑的结构扭曲，可见原缺失痕迹及较大面积的明显掩码痕迹。PIC 算法修复的整体结构相对完整，轮廓较为笔直锋利，但边缘区域存在像素丢失及内容不连续的问题，重影明显。RM-Net 算法能较好地捕获整体语义信息，但不能呈现出精细的纹理，表现为边缘模糊，存在掩码痕迹。提出算法修复的目标边界清晰，结构明确，纹理有序，掩码处不存在可见伪影。

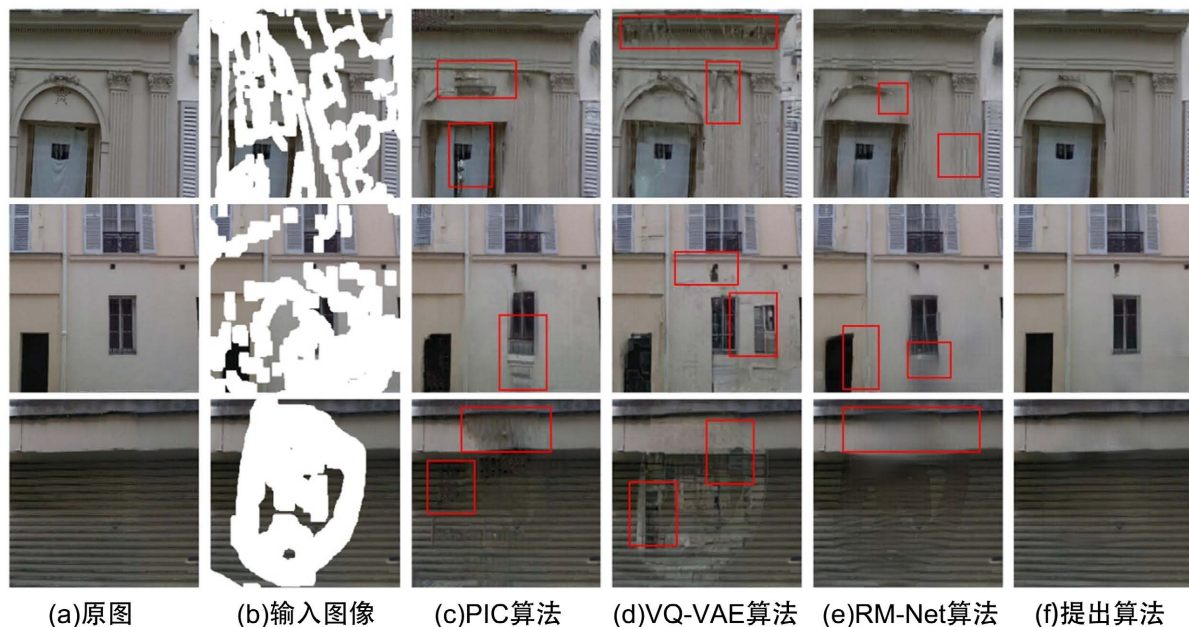


Figure 5. The comparison experiment of Paris-street

图 5. Paris-street 对比实验

### 3.2. 定量评估

为客观评估本文提出算法的效果，选择三种定量评估指标进行效果对比。使用峰值信噪比(PSNR)比较图像在纹理及像素层面的差异，凭借结构相似度(SSIM)区分图像在结构、亮度以及对比度上的差异，通过  $l_1$ (%)损失度量预测值和真实值之间的差距。对比算法均导入统一的训练集和测试集，采取相同的训

练机制，以减少训练和测试过程中可能产生的误差。

表 1 为各算法在 CelebA-HQ 数据集上叠加不规则掩码的修复性能指标，由实验结果可知：提出算法的 PSNR、SSIM、 $l_1(\%)$  3 个指标都高于比较算法。相比整体修复效果较好的 RM-Net 算法，提出算法的 PSNR 提高了 2.56，SSIM 提高了 0.088， $l_1(\%)$  损失降低了 0.86。客观实验指标也验证了提出算法的有效性。

**Table 1.** Quantitative comparison experiments

**表 1.** 定量实验对比

	PIC	VQ-VAE	RM-Net	Ours
PSNR	16.84	18.89	20.16	22.72
SSIM	0.75	0.73	0.806	0.894
$l_1(\%)$	4.82	5.36	4.14	3.28

以上评价结果皆表明提出算法在把握全局结构，纹理细节及语义信息方面均表现出较好性能，均优于比较算法。

#### 4. 结论

本文提出了一种基于高效注意力的密集残差修复算法。首先，对人为破损图像下采样，将提取的细节特征输入包含空洞卷积的密集残差网络，扩大信息感受野从而对细节纹理信息进行更加深入的提取。然后引入补充高效注意力机制，通过生成的全局上下文向量，表达输入特征的全局结构。同时叠加超图卷积模块，捕获全局信息，增强修复纹理的细粒度。最后，将生成的图像输入 SN-Patch GAN 鉴别器进行判别优化，进一步提升修复结构的整体一致性和细节纹理的精细度。将提出算法与现有的三种算法进行比较，实验结果表明：当待修复图像存在大面积不规则缺失时，提出算法有效提高了修复结果的结构合理性及纹理细节的精度，主客观指标优于对比算法。

#### 项目基金

国家级大学生创新创业训练计划支持项目(202110673087)。

#### 参考文献

- [1] Bertalmio, M., Sapiro, G., Caselles, V., *et al.* (2000) Image Inpainting. *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, New Orleans, 23-28 July 2000, 417-424. <https://doi.org/10.1145/344779.344972>
- [2] Ballester, C., Bertalmio, M., Caselles, V., Sapiro, G. and Verdera, J. (2001) Filling-In by Joint Interpolation of Vector Fields and Gray Levels. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, **10**, 1200-1211. <https://doi.org/10.1109/83.935036>
- [3] Esedoglu, S. and Shen, J.H. (2002) Digital Inpainting Based on the Mumford-Shah-Euler Image Model. *European Journal of Applied Mathematics*, **13**, 353-370. <https://doi.org/10.1017/S0956792502004904>
- [4] Criminisi, A., Pérez, P. and Toyama, K. (2004) Region Filling and Object Removal by Exemplar-Based Image Inpainting. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, **13**, 1200-1212. <https://doi.org/10.1109/TIP.2004.833105>
- [5] Chen, Q., Zhang, Y.X. and Liu, Y.C. (2007) Image Inpainting with Improved Exemplar-Based Approach. *International Workshop on Multimedia Content Analysis and Mining*, Weihai, 30 June-1 July 2007, 242-251.
- [6] Barnes, C., Shechtman, E., Finkelstein, A. and Goldman, D.B. (2009) PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing. *ACM Transactions on Graphics*, **28**, Article No. 2. <https://doi.org/10.1145/1531326.1531330>



- 
- [7] Pathak, D., Krähenbühl, P., Donahue, J., Darrell, T. and Efros, A.A. (2016) Context Encoders: Feature Learning by Inpainting. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 2536-2544. <https://doi.org/10.1109/CVPR.2016.278>
- [8] Li, Y., Liu, S., Yang, J. and Yang, M.-H. (2017) Generative Face Completion. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 5892-5900. <https://doi.org/10.1109/CVPR.2017.624>
- [9] Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X. and Huang, T.S. (2018) Generative Image Inpainting with Contextual Attention. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 5505-5514. <https://doi.org/10.1109/CVPR.2018.00577>
- [10] Liu, G., Reda, F., Shih, K., *et al.* (2018) Image Inpainting for Irregular Holes Using Partial Convolutions. *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 85-100. [https://doi.org/10.1007/978-3-030-01252-6\\_6](https://doi.org/10.1007/978-3-030-01252-6_6)
- [11] Nazeri, K., Ng, E., Joseph, T., Qureshi, F.Z. and Ebrahimi, M. (2019) EdgeConnect: Generative Image Inpainting with Adversarial Edge Learning.
- [12] Zheng, C.X., Cham, T.-J. and Cai, J.F. (2019) Pluralistic Image Completion. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 1438-1447. <https://doi.org/10.1109/CVPR.2019.00153>
- [13] Peng, J., Liu, D., Xu, S., *et al.* (2021) Generating Diverse Structure for Image Inpainting with Hierarchical VQ-VAE. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, 20-25 June 2021, 10775-10784. <https://doi.org/10.1109/CVPR46437.2021.01063>
- [14] Jam, J., Kendrick, C., Drouard, V., *et al.* (2021) R-MNet: A Perceptual Adversarial Network for Image Inpainting. 2021 *IEEE Winter Conference on Applications of Computer Vision*, Waikoloa, 5-9 January 2021, 2713-2722. <https://doi.org/10.1109/WACV48630.2021.00276>
- [15] Suvorov, R., *et al.* (2022) Resolution-Robust Large Mask Inpainting with Fourier Convolutions. 2022 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, 3-8 January 2022, 3172-3182. <https://doi.org/10.1109/WACV51458.2022.00323>
- [16] Liu, H.Y., Jiang, B., Xiao, Y. and Yang, C. (2019) Coherent Semantic Attention for Image Inpainting. <https://arxiv.org/abs/1905.12384>
- [17] Li, J., Wang, N., Zhang, L., Du, B. and Tao, D. (2020) Recurrent Feature Reasoning for Image Inpainting. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 7757-7765. <https://doi.org/10.1109/CVPR42600.2020.00778>
- [18] Shen, Z.R., Zhang, M.Y., Zhao, H.Y., *et al.* (2021) Efficient Attention: Attention with Linear Complexities. 2021 *IEEE Winter Conference on Applications of Computer Vision*, Waikoloa, 5-9 January 2021, 3530-3538. <https://doi.org/10.1109/WACV48630.2021.00357>
- [19] Vaswani, A., Shazeer, N., Parmar, N., *et al.* (2017) Attention Is All You Need. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Curran Associates Inc., Red Hook, 6000-6010.
- [20] Johnson, J., Alahi, A. and Li, F.-F. (2016) Perceptual Losses for Real-Time Style Transfer and Super-Resolution. 14th *European Conference Computer Vision—ECCV* 2016, Amsterdam, 11-14 October 2016, 694-711. [https://doi.org/10.1007/978-3-319-46475-6\\_43](https://doi.org/10.1007/978-3-319-46475-6_43)