DOI:10.16136/j.joel.2022.04.0521

基于渐进式特征推理的裂缝图像修复方法研究

李良福*,黎光耀,王 楠,张 晰

(陕西师范大学 计算机科学学院,陕西 西安 710000)

摘要:图像修复是计算机图形学与计算机视觉中的研究热点之一。针对传统裂缝图像修复采用一次性补全修复方法并没有语义理解能力,当语义场景较为复杂且图像缺陷较大时修复效果不佳的问题,提出了一种基于渐进式特征推理的裂缝图像修复方法。该方法从孔洞边缘逐步修复图像,加强对孔洞中心的约束。首先通过计算掩膜占比确定修复比例,使用部分卷积更新掩膜,然后利用VGG-16网络进行特征提取,再使用语义注意力机制,生成高质量的图像特征,并采用跳跃连接方法,增强遥远距离的梯度相关性,为后续图像修复提供多尺度多层次的特征信息。最后,将递归生成的特征图进行融合解码生成修复图像。实验结果表明,与传统图像修复方法相比,本方法修复的裂缝图像峰值信噪比提升了0.5 dB—1.2 dB,产生语义明确的修复结果。

关键词:图像修复;注意力机制;部分卷积;裂缝图像

中图分类号:TP391.9 **文献标识码**:A **文章编号**:1005-0086(2022)04-0393-10

Research on crack image inpainting method based on progressive feature reasoning

LI Liangfu*, LI Guangyao, WANG Nan, ZHANG Xi

(College of Computer and Science, Shaanxi Normal University, Xi'an, Shaanxi 710000, China)

Abstract: Image inpainting is one of an activate research topic in the domain of computer vision and computer graphics. Aiming at the problem that the traditional crack image restoration method using one-time completion restoration method does not have the ability to understand semantics, and the repair effect is not good when the semantic scene is more complex and the image defect is large, a crack image restoration based on progressive feature reasoning is proposed. This method gradually restores the image from the hole edge and strengthens the constraint on the hole center. At first, partial convolution is used to update the mask, and the update ratio is determined by calculating the mask proportion. Then, use the VGG-16 network for feature extraction, semantic attention mechanism is used to generate high-quality image features, and use the jump connection method to enhance the gradient correlation of remote distances, so as to provide multi-scale and multilevel feature information for subsequent image restoration. Finally, the recursive feature map is fused and decoded to generate a repair image. The experimental results show that the proposed method, compared with traditional image inpainting methods, can improved the peak signal-to-noise ratio of the crack image repaired for 0.5 dB—1.2 dB and produce semantic clear inpainting results.

Key words: image inpainting; attention mechanism; partial convolution; crack image

1 引 言

图像修复的主要目的是在给定掩膜的情况下,基于背景信息修复受损图像中的缺失像素。图像修复算法在老旧照片修复和图像编码等领域有着广泛的应用前景[1-3]。裂缝是最常见的公路桥梁安全问题之一[4],对其检测必不可少,现实生活中采集到的裂缝图像除了裂缝以外还包含泥土、油漆、落叶等障碍物,这些障碍物容易造成裂缝检测漏检和误检等问题,影响裂缝检测的精确

* E-mail: longford@xjtu.edu.cn

收稿日期:2021-07-25 修订日期:2021-09-01

基金项目:国家自然科学基金(61573232,61401263)和陕西省自然科学基金(2022JM335)资助项目

度。因此，消除裂缝图像中的障碍物并预测障碍物遮挡位置是否包含裂缝以及裂缝的分布走向变得十分重要。

传统的图像修复基于扩散法与示例法，例如Barnes的PatchMatch方法[5]，在修复过程中执行图像块匹配，每次构建缺失部分像素点的同时保持与周围像素点的一致性。尽管此方法具有修复简单语义场景的优点，但是它没有语义理解能力，对复杂语义场景的修复效果不佳。近些年随着深度网络在图像修复领域的广泛应用[6-9]，大多数修复方法采用解码器与编码器结构[10-12]，并且假设图像编码后有足够的特征信息对受损图像进行修复，然后一次性修复图像。YU等[14]提出了Contextual attention[13]，允许图像从遥远区域借用特征块。IZUKA等[14-16]把图像修复定义为条件图像生成问题，将生成器与判别器联合训练，能够生成合理的新内容。该方法在裂缝图像遮挡区域较小时是可行的，因为局部区域内的像素点具有很强的相关性，可以从遮挡区域周围环境推断像素。但是当裂缝图像遮挡区域过大，已知像素与缺失像素间的相关性就会减弱，已知区域像素无法给缺失区域像素提供准确的信息，使修复的裂缝图像产生语义不明的结果。另一种图像修复方法是从孔边界到中心的渐进式修复[17]。ZHANG等[18]提出的PGN(progressive generative networks)，是在图像级别执行的，计算成本过大，并且因为反复的从特征图转化成RGB图像会导致信息失真。XIONG等[19]和KAMYAR等[20]采用图像补全逐步填充图像，能够加强结构间的约束，但缺乏对图像特征的语义理解能力。基于上述问题，本文提出了一种渐进式特征推理的裂缝图像修复方法CI-Net，CI-Net的网络结构是U-net类型，由上下文编码器、解码器、裂缝特征推理(crack featurereasoning，CFR)等3个模块组成，能够将输入图像编码为高级语义特征，在图像特征层面使用多次循环和渐进的方式由外向内逐步完善特征图，再将特征解码为图像Pconv(partial convolutions)的使用能够让不同层的掩膜对损失表现出不同贡献，使得训练过程表现出从孔洞外面逐渐缩小孔洞学习的机制，提高了网络的学习效率。经实验证明，相比于传统方法[21,22]，CI-Net对去除障碍物后的裂缝图像修复质量更高。

2　基本原理

2.1　部分卷积

部分卷积[23]将卷积分为输入图像的卷积和输入掩膜的卷积2部分，让卷积只在图像的有效区域进行处理。无效掩膜区域的像素值为0，即相对应的图像特征不参与当前的卷积运算。在CI-Net中部分卷积用于识别输入特征图中需要更新的区域，并在卷积操作后进行标准化特征映射，其表达式为：

$$f_{x,y,z}=\begin{cases}W_z^T\left(f_{x,y}\odot m_{x,y}\dfrac{sum(1)}{sum(m_{x,y})}\right)+b,\text{if } sum(m_{x,y})!=0\\0,\text{else}\end{cases},\quad(1)$$

式中，$f_{x,,y,z}$表示第z层通道(x,y)位置的特征值，W_z表示第z层卷积层滤波器的权重，$f_{x,y}$表示当前滑动窗口的特征值。$m_{x,y}$是相对应的二进制掩膜，$\odot$表示特征图对应位置相乘，$sum(1)/sum(m_{x,y})$是比例因子，当卷积的有效输入数值的数量发生变化时调整结果，b为卷积滤波器的偏置。

每次卷积操作后更新掩膜，其表达式为：

$$m_{x,y}=\begin{cases}1,\text{if } sum(m_{x,y})>0\\0,\text{othwise}\end{cases}。\quad(2)$$

如果卷积操作调整了有效输入值的输出，则移除该位置的掩膜。更新后的掩膜在本次递归过程中会保留，直到下次递归更新。根据上述方法，CI-Net每次都使用更新后的掩膜，输入特征图经过部分卷积后，输出特征图经过Leaky Relu激活函数处理输入至裂缝特征推理模块。随着网络层数的增加，$m_{x,y}$中值为0的像素越来越少，$f_{x,y,z}$中有效区域的面积越来越大，掩膜对整体的loss的作用会越来越小。

2.2　语义注意力机制

图像修复中，注意力机制能够合成质量更高的特征[24]。CI-Net使用语义注意力机制，能够用已知区域的图像纹理特征替换未知区域的纹理特征，合成高质量的图像特征，在实验中证实具有更高的修复能力。语义注意力机制的计算细节如下：

输入的裂缝特征图定义为F，其宽、高和通道数分别是w、h、c。F为在(x,y)位置的值，其大小为$1\times1\times c$。接下来计算f中位置(x,y)与(x',y')处特征的余弦相似性。其表达式为：

$$\hat{sim}^{\,i}_{x,y,x',y'}=\left\langle\frac{f_{x,y}}{\|f_{x,y}\|},\frac{f_{x',y'}}{\|f_{x',y'}\|}\right\rangle,\quad(3)$$

式中，$\hat{sim}^{\,i}_{x,y,x',y'}$表示$f$中位置$(x,y)$与$(x',y')$处特征的余弦相似性，上标$i$为当前循环的次数。通过将$1\times c\times w\times h$的特征图按像素抽取$w\times h$个$1\times1\times c$卷积核，经过归一化后与特征图求卷积得到大小为$1\times wh\times w\times h$的张量，张量的第$(0,a,b,c)$处元素的值代表索引为$a$的特征向量与索引为$(b,c)$的特征向量之间的相似度。然后利用相邻区域中目标像素的相似度在长宽维度上做$k\times k$滤波来平滑注意力分数，其表达式为：

$$score'^{i}_{x,y,x',y'} = \frac{\sum_{p,q\in[-k,\cdots,k]} \hat{sim}_{x+p,y+q,x',y'}}{k\times k}。\quad (4)$$

本实验将卷积核尺寸 k 设为 3,使用步长为 1 的平均池化函数实现式(4)的计算。接下来使用 soft max 函数在通道维度上处理 sim',生成的分数表示 $score'^{i}$。再利用注意力分数重建特征图,这一过程并不会改变特征图的尺寸。重建特征图 $\hat{F}$ 在(x,y)处的值为 $\hat{f}_{x,y}$,其计算式为:

$$\hat{f}^{i}_{x,y} = \sum_{x'\in1,\cdots,W\ y'\in1,\cdots,H} score^{i}_{x,y,x',y'} f^{i}_{x',y'}。\quad (5)$$

最后,将重建的特征图 $\hat{F}$ 与原特征图 F 在通道维度上拼接,其尺寸为 $1\times2c\times w\times h$,接下来使用 c 个尺寸为 $2c\times w\times h$ 的卷积核对拼接后的特征图进行一次卷积,得到当前循环最终更新的尺寸为 $1\times c\times w\times h$ 的特征图 F'。

3　裂缝图像修复模型

3.1　裂缝图像修复模型结构

传统的图像修复无法解决语义模糊问题。本文针对裂缝图像修复提出了 CI-Net。CI-Net 的输入为原始图像与掩膜,原始图像与掩膜首先经过 2 次部分卷积,为了防止模型过拟合,2 次部分卷积后分别对输出特征图进行归一化处理。其网络结构图如图 1 所示。

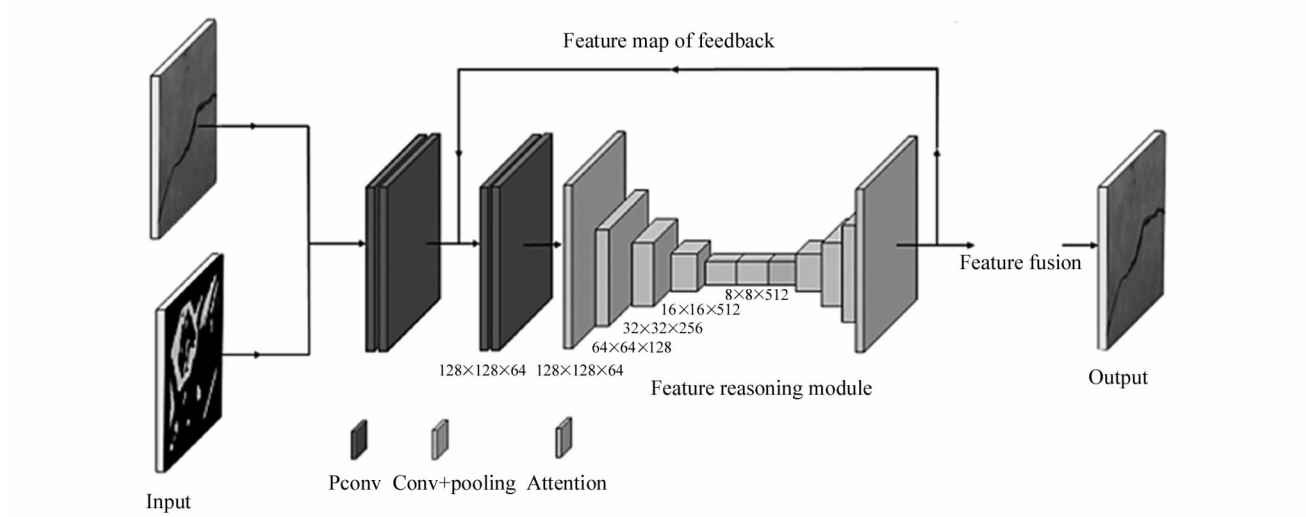

图 1　CI-Net 网络结构图

Fig. 1　The network structure of CI-Net

传统图像修复方法基于临近像素[25],只能修复遮挡区域较小的图像,并且要求遮挡区域的像素和邻近像素差不多。CI-Net 使用 Pconv,能够利用掩膜对卷积操作进行正则化,使其不对损坏区域进行激活响应,让卷积只在图像的有效区域进行,保证卷积操作的输出值与缺失像素的值无关。

在修复过程中,CI-Net 首先计算掩膜占比确定更新范围的大小,其表达式为:

$$r = \frac{E_{1,n}\Psi_{n,n}E_{n,1}}{n\times n},\quad (6)$$

式中,r 为掩膜占比,n 为掩膜的大小,$E_{1,n}$ 与 $E_{n,1}$ 为值全为 1 的向量,$\Psi_{n,n}$为图像相对应的掩膜。

根据 r 确定部分卷积的卷积核大小,将两个卷积核大小分为 7 与 5 的部分卷积层联在一起,用来更新掩膜和特征图。在循环部分 CI-Net 使用一个部分卷积更新掩膜,再通过归一化与 Leaky_Relu 处理后发送到裂缝特征推理模块。

在裂缝特征推理模块,使用语义注意力机制提高 CI-Net 恢复高质量特征图的能力,将原始图像的特征作为卷积核处理掩膜遮挡区域产生 Attention score,接下来让 Attention score 作用于特征图,达到用原始图像的纹理替换遮挡区域的纹理的目的,能够产生语义明确的修复结果,并使用 skip-connection 方法能够将较浅的卷积层特征引入到相对应的反卷积过程,为修复图像遮挡区域提供更多的特征信息,提高生成图像的质量。

3.2　裂缝特征推理模块

裂缝特征推理模块作为修复图像的单元,对遮挡区域的特征值进行估计,使用高质量特征修复遮挡区域,不仅有利于产生更好的修复结果,还有利于后续的特征推理,提高了网络的修复性能,并采用跳跃连接[27],能够增强遥远距离的梯度相关性,给后期图像修复提供多尺度和多层次的信息,提高了图像的修复质量。裂缝特征推理模块网络结构如表1所示。

表 1 裂缝特征推理模块网络结构

Tab. 1 Network structure of fracture feature reasoning module

M_name	Input	In_Size	K_Size	S_size	C_Num	O_Size	BN	Act_Func
PConv2	PConv1	Ori/2	7	1	64	Ori/2	T	Relu
PConv3	FPConv2	Ori/2	5	1	64	Ori/2	T	Relu
Conv1	FPConv3	Ori/2	3	2	128	Ori/4	T	Relu
Conv2	FConv1	Ori/4	3	2	256	Ori/8	T	Relu
Conv3	FConv2	Ori/8	3	2	512	Ori/16	T	Relu
Conv4	FConv3	Ori/16	3	2	512	Ori/16	T	Relu
Conv5	FConv4	Ori/32	3	1	512	Ori/32	T	Relu
Conv6	FConv5	Ori/32	3	1	512	Ori/32	T	Relu
Conv7	FConv6	Ori/32	3	1	512	Ori/32	T	Relu
Conv8	Cat(FConv7,FConv6)	Ori/32	3	1	512	Ori/32	T	Leaky_Relu
Conv9	Cat(FConv8,FConv5)	Ori/32	3	1	512	Ori/32	T	Leaky_Relu
Attention	FConv9	Ori/32				Ori/32		None
DeConv1	Cat(Attention,FConv4)	Ori/32	4	2	512	Ori/32	T	Leaky Relu
DeConv2	Cat(FDeconv1,FConv3)	Ori/16	4	2	256	Ori/16	T	Leaky_Relu
DeConv3	Cat(FDeconv2,FConv2)	Ori/8	4	2	128	Ori/8	T	Leaky_Relu
DeConv4	Cat(FDeconv3,FConv1)	Ori/4	4	2	64	Ori/4	T	Leaky_Relu

CI-Net 的裂缝特征推理模块使用 VGG-16 模型[28]提取图像特征。输入的特征图第 1 次经过 64 个 K_size 为 3 的卷积核的 2 次卷积后，采用 1 次 pooling。第 2 次经过 128 个 K_size 为 3 的卷积核的 2 次卷积后，采用 1 次 pooling。第 3 次经过 256 个 K_size 为 3 的卷积核的 3 次卷积后，采用 1 次 pooling。最后特征图经过 512 个 K_size 为 3 的卷积核的 3 次卷积后采用 1 次 pooling，此时特征图的尺寸为 8×8×512。再进行 3 次不改变输出尺寸的卷积操作后将特征传入语义注意力机制。

由于在特征提取的时候通过最大池化降低特征图的分辨率，不利于生成高质量的裂缝图像。在特征解码的过程中通过 skip-connection 将 Conv1-4 的输出拼接在转置卷积层 DeConv1-4 相对应级别的输入特征图上，使得解码器在上采样过程中可以获取到更多图像特征信息。通过底层特征与高层特征的融合，CI-Net 能够保留更多高分辨率信息，进而更完善的恢复图像中的细节信息，提高修复图像的质量。

3.3 特征融合

当图像的孔洞已经被完全填充或者循环次数达到上限，进入到图像特征融合部分. 如果只使用最后 1 次循环输出的裂缝图像特征图，梯度传播过远会导致梯度消失的问题. 为了解决上述问题，图像特征融合部分加入了中间特征图，通过计算 5 次循环在 DeConv4 输出的有效结果的平均值，有效结果指特征图被填充过的区域，之后得到融合的特征图，其计算式为：

$$\overline{f}_{x,y,z}=\frac{\sum_{i=1}^{N}f_{x,y,z}^{i}}{\sum_{i=1}^{N}m_{x,y,z}^{i}},\tag{7}$$

式中，F^i 表示第 i 次裂缝特征推理模块生成的特征图，$f_{x,y,z}$ 为特征图 F 在 x，y，z 位置的值，M^i 表示 F^i 相对应的二进制掩膜，$m_{x,y,z}$ 表示掩膜在 x，y，z 位置的值，N 为特征图的数量。

3.4 损失函数

对于裂缝图像生成学习，为了提高像素修复的精度，CI-Net 的损失函数由 4 个部分组成，分别是 $L1$ 损失、感知损失、风格损失和 TV 损失。$L1$ 损失确保像素级重建的准确性，其表达式为：

$$\begin{aligned}L_{\text{hole}}&=\frac{1}{N_{I_{\text{gt}}}}\|(1-M)\odot(I_{\text{out}}-I_{\text{gt}})\|_1,\\L_{\text{valid}}&=\frac{1}{N_{I_{\text{gt}}}}\|M\odot(I_{\text{out}}-I_{\text{gt}})\|_1,\end{aligned}\tag{8}$$

式中，I_{gt} 与 I_{out} 分别表示网络的真实值与输出值。M 是二进制掩膜，0 表示孔洞，1 表示有效像素。$N_{I_{\text{gt}}}$ 表示图像中像素值的总数，等于 $C\times W\times H$，C 表示通道数，W 和 H 分别表示 I_{gt} 的高和宽。

为了使填充图像和背景真实图像具有类似的特征表示，引入由 Gatys 等[29] 提出的感知损失。将裂缝真实图像与填充图像输入到经过预先训练的 VGG-16 中提取特征，然后计算所有或几层特征值之间的 $L1$ 距离，其表达式为：

$$L_{\text{perceptual}}=\sum_{n=0}^{N-1}\frac{\|\Psi_n(I_{\text{out}})-\Psi_n(I_{\text{gt}})\|_1}{N_{\Psi_n^{I_{\text{gt}}}}}+\sum_{n=0}^{N-1}\frac{\|\Psi_n(I_{\text{comp}})-\Psi_n(I_{\text{gt}})\|_1}{N_{\Psi_n^{I_{\text{gt}}}}},\quad(9)$$

式中,Ψ_n 表示第 n 个特征图层,I_{comp} 是原始输入图像的输出值,$N_{\Psi_n^I}$ 是 Ψ_n^I 中元素的个数,CI-Net使用 VGG-16 pool1、pool2 和 pool3 层计算感知损失。

风格损失类似于感知损失,在使用 $L1$ 距离之前计算每个特征图的自相关性,也被称为格雷姆矩阵,格雷姆矩阵包含裂缝图像的风格信息,如图像颜色与纹理。之后计算完成图像的 Gram 矩阵和真实裂缝图像之间的 $L1$ 距离,其计算式为:

$$I_{\text{style}_{\text{out}}}=\sum_{n=0}^{N-1}\frac{1}{C_nC_n}\|K_n((\Psi_n(I_{\text{out}}))^{\text{T}}(\Psi_n(I_{\text{out}}))-(\Psi_n(I_{\text{gt}}))^{\text{T}}(\Psi_n(I_{\text{gt}})))\|_1,\quad(10)$$

$$I_{\text{style}_{\text{comp}}}=\sum_{n=0}^{N-1}\frac{1}{C_nC_n}\|K_n((\Psi_n(I_{\text{comp}}))^{\text{T}}(\Psi_n(I_{\text{comp}}))-(\Psi_n(I_{\text{gt}}))^{\text{T}}(\Psi_n(I_{\text{gt}})))\|_1,\quad(11)$$

式中,K_n 为归一化因子,大小为 $1/C_nH_nW_n$,由第 n 层特征图的空间大小决定。Ψ_n 的大小为 $(H_nW_n)\times C_n$,Ψ_n 的转置与 Ψ_n 的矩阵乘积输出的矩阵大小为 $C_n\times C_n$。TV 损失表示为内一个像素和该像素的右侧像素和下面像素的 $L1$ loss,衡量了原始孔洞区域与水平方向右移一个像素的区域或者垂直方向下移一个像素的区域在水平方向和垂直方向上的差异。其计算式为:

$$I_{\text{tv}}=\sum_{(i,j)\in R,(i,j+1)\in R}\frac{1}{N_{I_{\text{comp}}}}\|I_{\text{comp}}^{i,j+1}-I_{\text{comp}}^{i,j}\|_1+\sum_{(i,j)\in R,(i+1,j)\in R}\frac{1}{N_{I_{\text{comp}}}}\|I_{\text{comp}}^{i+1,j}-I_{\text{comp}}^{i,j}\|_1,\quad(12)$$

式中,$N_{I_{\text{comp}}}$ 表示 I_{comp} 中像素的数量。

总损失 L_{total} 为上述损失函数的加权和,其计算式为:

$$L_{\text{total}}=L_{\text{valid}}+6L_{\text{hole}}+0.05L_{\text{perceptual}}+120(L_{\text{style}_{\text{out}}}+L_{\text{style}_{\text{comp}}})+0.01L_{\text{tv}}。\quad(13)$$

CI-Net 模型中损失函数的系数在已有的实验中证明是有效的。L_{total}需要更新的参数较少,所以 CI-Net 具有较高的训练效率。

4　分析与讨论

4.1　训练模型

本文实验环境为 Ubuntu16.04 操作系统,使用 i7-6700K 4.00 GHz CPU 和 11G NVIDIA RTX2080Ti GPU。使用 Adam 训练模型,batch_size 的值设为 8。初始学习率设为 10^{-4},然后将学习率设为 10^{-5}对模型进行微调。超参数设置部分,pconv 的卷积核大小分别为 5 与 7,将 1 作为 L_{valid}的系数、6 作为 L_{hole} 的系数、0.05 作为 $L_{\text{perceptual}}$ 的系数、0.1 作为 L_{tv}的系数、120 作为 L_{styleout}与 $L_{\text{stylecomp}}$的系数。

4.2　数据集

本文使用的数据集为实拍的 8 272 张 256×256 大小的 RGB 彩色图像作为原始裂缝图像数据集。通过对裂缝图像进行旋转,左右翻转扩充数据集,最终得到 16 544 张裂缝图像。其中 14 944 张裂缝图像作为 CI-Net 的训练数据集,1 600 张裂缝图像作为测试数据集。

4.3　CI-Net 修复效果对比

在这一部分,本文通过在裂缝图像数据集上进行试验,与其它图像修复方法进行了比较。分别使用结构相似性指数(structural similarity,*SSMI*)、$L1$ loss、模糊系数(KBlur)和峰值信噪比(peak signal to noise ratio,*PSNR*)对裂缝图像修复效果进行评估。

本组实验将 CI-Net 与 Pconv 对裂缝图像修复效果进行比较。首先随机获取裂缝图像与掩膜,使用掩膜遮挡裂缝图像相对应区域的图像信息。遮挡比例 10%—30%修复结果示例图,如图 2 所示。图 2 从左至右分别为原始图像、遮挡图像、Pconv 修复图像和 CI-Net 修复图像。由图 2 可以分析出,当裂缝图像的遮挡区域较小时,CI-Net 对裂缝图像的修复效果与图像质量指标均优于 Pconv 算法。

为了验证 CI-Net 更适用于大孔洞裂缝图像的修复,本实验将 CI-Net 与 Pconv 对遮挡比例为 30%—50%的裂缝图像的修复结果进行了对比,结果如图 3 所示。由图 3 可知,当遮挡比例为 30%—50%时,Pconv 与 CI-Net 的 *SSIM* 差的绝对值由 0.009 提升至 0.015,*PSNR* 差的绝对值由 0.272 提升至 0.629、$L1$ loss 差的绝对值由 0.771 提升至 0.947。实验结果表明当图像遮挡比例为 30%—50%时,CI-Net 对裂缝图像的修复效果与图像质量指标更优于 Pconv 算法。

接下来将 CI-Net 与 Pconv、Edge-Connect[21] 和 Patch-Match 等图像修复方法生成的结果进行对比实验,修复结果如图 4 所示,从左至右依次为 Patch-Match 生成的图像、Pconv 生成的图像、Edge-Connect 生成的图像和 CI-Net 生成的图像。

不同算法修复效果评价,如表 2 所示。表 2 表明

与传统图像修复方法 Patch-Match 相比，CI-Net 的修复结果语义明确，生成了质量较高的裂缝图像。与渐进式修复方法 Edge-Connect 相比，当遮挡比例为 10%—30%，CI-Net 的 *PSNR* 与 *SSIM* 分别提高了 0.223 和 0.005。当遮挡比例为 30%—50%，CI-Net 的 *PSNR* 与 *SSIM* 分别提高了 0.617 和 0.007，修复性能有了进一步的提升。

为了测试 CI-Net 对含有真实障碍物图像的修复效果。本文制作掩膜遮挡裂缝图像中的障碍物信息，接下来使用 CI-Net 对使用掩膜遮挡障碍物信息后的裂缝图像进行修复，修复效果如图5所示，从左至右依次为原始图像、含有障碍物图像、遮挡图像和

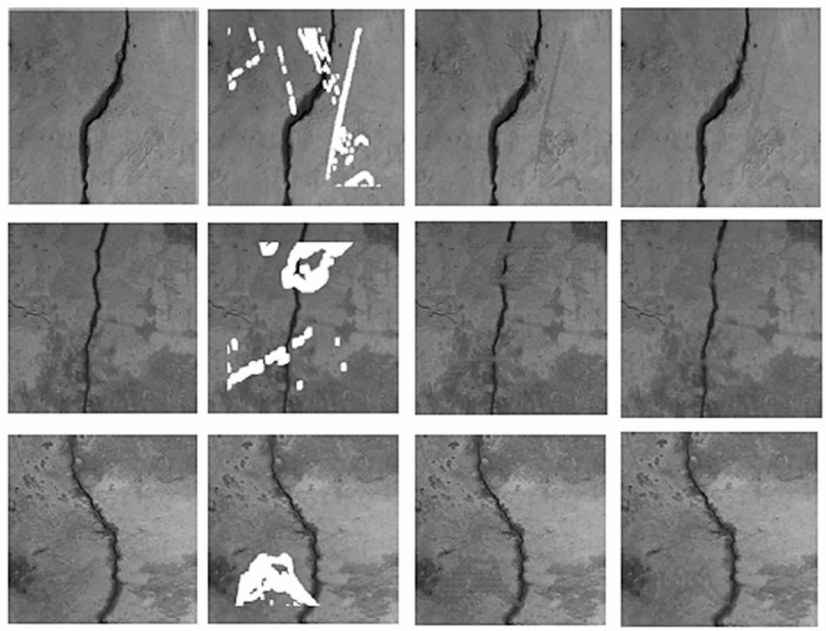

图 2　遮挡比例 10%—30%修复结果示例图

Fig. 2　Example image of inpainting result with mask ratio of 10%—30%

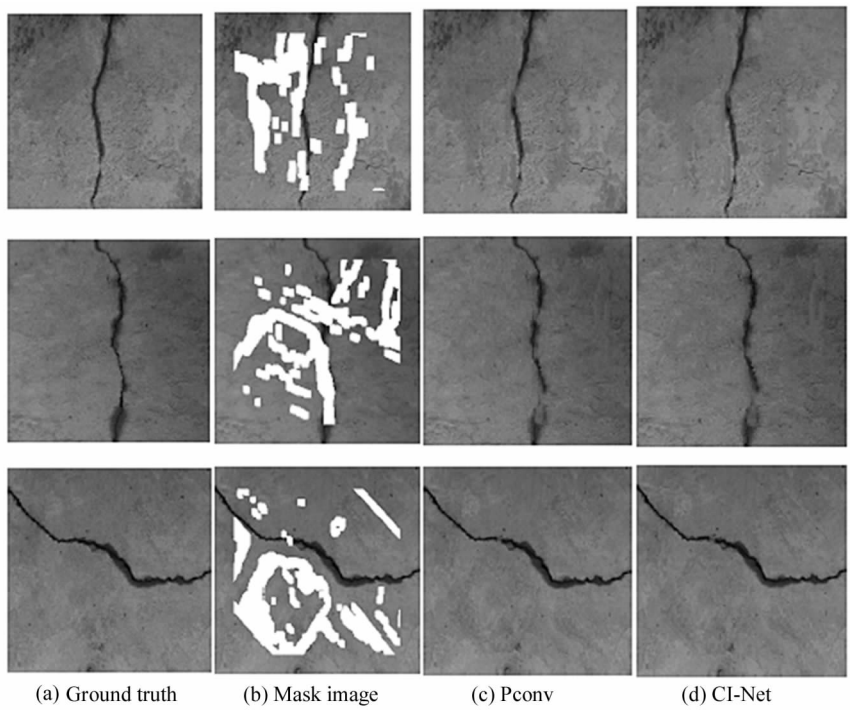

图 3　遮挡比例 30%—50%修复结果示例图

Fig. 3　Example image of inpainting result with mask ratio of 30%—50%

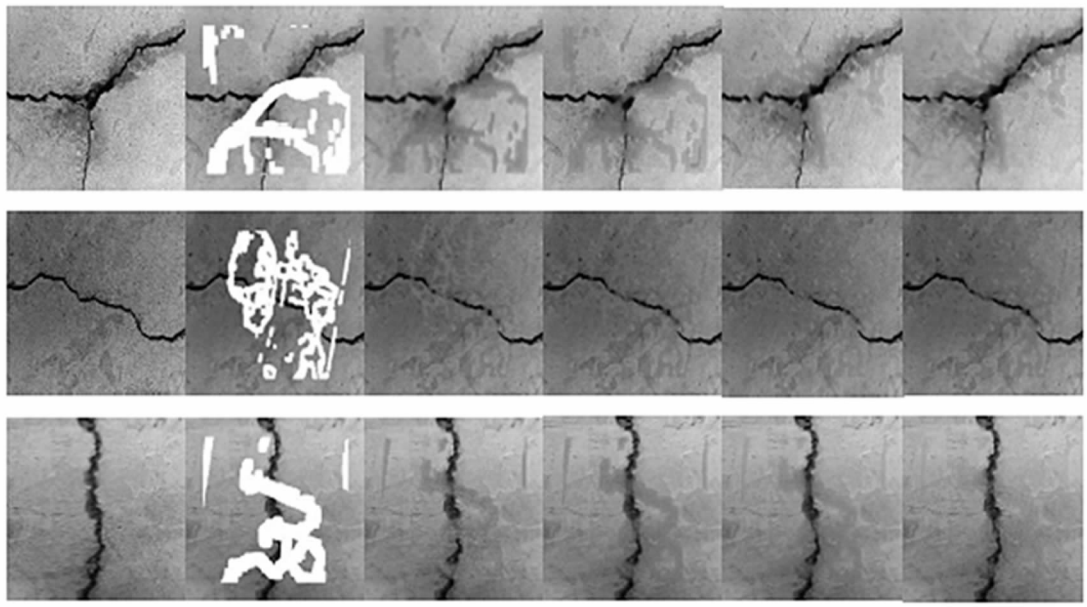

图 4　不同算法修复效果示例图

Fig. 4　Example image of inpainting effects with different algorithms

表 2　不同算法修复效果评价

Tab. 2　Evaluation of inpaint effects of different algorithms

Mask ratio	Algorithm	*L*1 loss	*PSNR*	*SSIM*
10%—30%	Patch-Match	11.356	26.997	0.908
	Pconv	9.257	31.864	0.932
	Edge-Connect	8.658	31.913	0.936
	CI-Net	8.468	32.536	0.949
30%—50%	Patch-Match	13.588	23.291	0.786
	Pconv	10.783	27.824	0.912
	Edge-Connect	10.621	27.836	0.921
	CI-Net	9.842	29.453	0.932

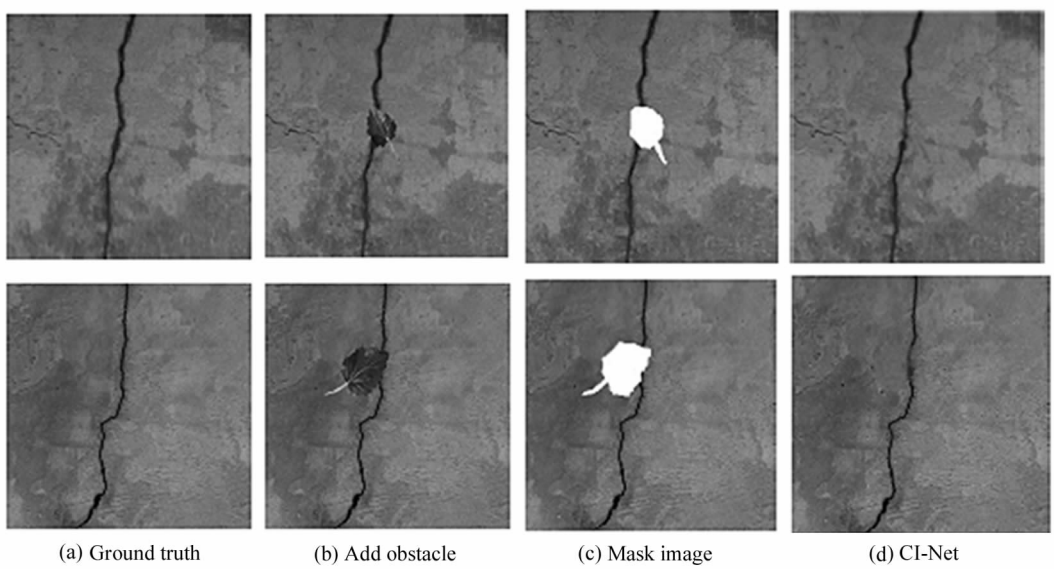

图 5　遮挡障碍物修复效果示例图

Fig. 5　Example image of the inpainting effect with obstructed obstacles

CI-Net生成图像。

遮挡障碍物修复效果评价，如表3所示。由表3可知，图像1的*PSNR*为34.543，*L*1 loss为0.974，图像2的*PSNR*为34.198，*L*1 loss为0.963，CI-Net生成了质量较高的裂缝图像。

表3 遮挡障碍物修复效果评价

Tab. 3 Evaluation of the repair effect of obstructed obstacles

Image	*L*1 loss	*PSNR*	*SSIM*
Image 1	7.294	34.543	0.974
Image 2	7.963	34.198	0.963

4.4 消融实验

为了验证裂缝特征推理模块的有效性，本节通过移动裂缝特征推理模块证明其能安装到其他网络的任意部分。在遮挡比例低于20%的情况下对裂缝推理模块前分别增加1次下采样、增加2次下采样、增加3次下采样后保持特征推理模块不变进行实验。不同下采样次数修复效果对比，如表4所示。由表4可知将裂缝特征推理模块移动较深时性能受影响较小，表明裂缝特征推理模块具有较强的稳定性。

表4 不同下采样次数修复效果对比

Tab. 4 Comparison of inpaint effects of different downsampling times

Down sampling quantity	*L*1 loss	*PSNR*	*SSIM*
1	8.425	32.214	0.951
2	8.468	32.168	0.948
3	8.471	32.128	0.939

接下来使用裂缝推理模块替换Edge-Connect的第2个网络中的残差块，其余网络结构与原文保持一致，对遮挡比例30%—50%的裂缝图像进行修复，并对结果进行比较。添加CFR模块修复结果示例图如图6所示，从左至右依次为输入图像、遮挡图像、Edge Connect生成图像、Edge Connect＋CFR生成图像。

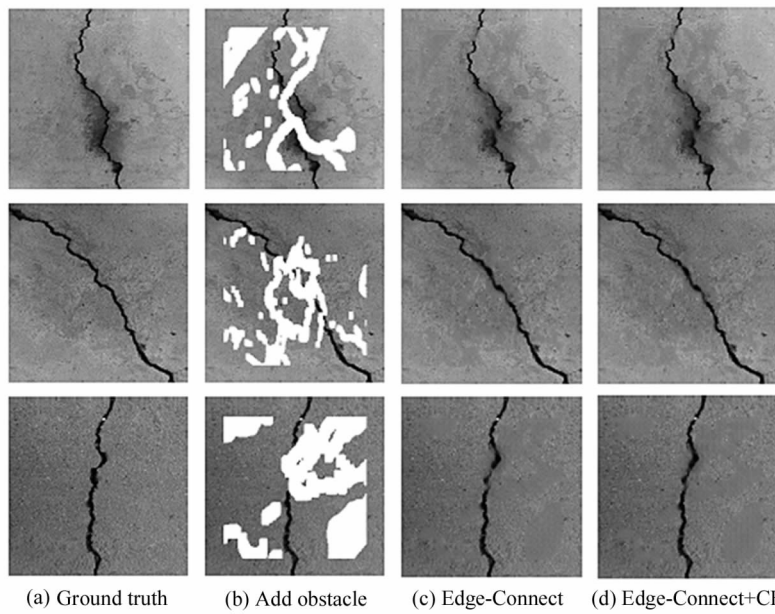

图6 添加CFR模块修复结果示例图

Fig. 6 Example image of adding CFR module repair results

添加CFR模块修复效果对比如表5所示。由表5可以得出添加CFR模块后*PSNR*与*L*1 loss分别提高了0.025和0.002，修复性能得到了提升，验证了CFR模块的高效性。

表5 添加CFR模块修复效果对比

Tab. 5 Add CFR module repair effect comparison

Module	*L*1 loss	*PSNR*	*SSIM*
Edge-Connect	10.621	27.836	0.921
Edge-Connect＋CFR	10.583	27.861	0.923

5 结　论

本文详细的描述了一种基于语义注意力机制的裂缝图像修复方法，并提出了渐进式修复网络CI-

Net,其使用部分卷积与注意力机制,逐步对裂缝图像上的孔洞进行修复,生成了语义明确的修复结果。本文设计的裂缝特征推理模块可以与现有的图像修复方法相结合,提高了图像的修复质量。通过大量的比较实验和消融实验证明了 CI-Net 网络的性能有效性以及 CFR 模块的高效性。目前本文使用部分卷积更新掩膜,上一次卷积操作覆盖的像素在下一层的掩膜都将设置为 1,掩膜更新的灵活性有待提高。在未来的工作中将研究如何合理化更新掩膜,从而达到提高图像修复质量的目的。

参考文献:

[1] DAI T,CAI J,ZHANG Y,et al. Second-order attention network for single image super-resolution[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE,2019:11065-11074.

[2] SONG L, JIE C, SONG L, et al. Geometry-aware face completion and editing[EB/OL]. (2019-02-13) [2021-07-25]. https://arxiv.org/abs/1809.029672018.

[3] SAJJADI M S M, SCHOLKOPF B, HIRSCH M. Enhancenet: single image super-resolution through automated texture synthesis[C]//IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York:IEEE,2017:4491-4500.

[4] LI L F, MA W F,LI L,et al. Research on detection algorithm for bridge cracks based on deep learning[J]. Acta Automatica Sinica,2019,45(9):1727-1742.
李良福, 马卫飞, 李丽, 等. 基于深度学习的桥梁裂缝检测算法研究[J]. 自动化学报, 2019,45(9):1727-1742.

[5] BARNES C, SHECHTMAN E, FINKELSTEIN A, et al. PatchMatch: a randomized correspondence algorithm for structural image editing[J]. ACM Transactions on Graphics,2009,28(3):24:1-24:11.

[6] ZHANG Y, TIAN Y, KONG Y, et al. Residual dense network for image super-resolution[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23,2018,Salt Lake City,UT,USA. New York: IEEE, 2018:2472-2481.

[7] KÖHLER R, SCHULER C, SCHÖLKOPF B, et al. Mask-specific inpainting with deep neural networks[C]//German Conference on Pattern Recognition, September 2-5, 2014, Munster, Germany. Berlin: Springer, 2014: 523-534.

[8] XIE C,LIU S,LI C,et al. Image inpainting with learnable bidirectional attention maps[C]//IEEE/CVF International Conference on Computer Vision, October 27-28, 2019, Seoul, Korea (South). New York: IEEE, 2019: 8858-8867.

[9] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Advances in Neural Information Processing Systems,2014,3:2672-2680.

[10] WANG X, YU K, WU S, et al. ESRGAN:enhanced super-resolution generative adversarial networks[C]//European Conference on Computer Vision, September 8-14, 2018, Munich, Germany. Berlin: Springer, 2018.

[11] YANG C, LU X, LIN Z, et al. High-resolution image inpainting using multi-scale neural patch synthesis[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, Hawaii, USA. New York:IEEE,2017:6721-6729.

[12] TANG X, DU Y, LIU Y, et al. Image recognition with conditional deep convolutional generative adversarial networks [J]. Acta Automatica Sinica,2018,44(5):855-864.

[13] YU J, LIN Z, YANG J, et al. Generative image inpainting with contextual attention[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE, 2018: 5505-5514.

[14] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics (TOG),2017,36(4):107.

[15] LI Y, LIU S, YANG J, et al. Generative face completion [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, Hawaii, USA. New York:IEEE,2017:3911-3919.

[16] YEH R, CHEN C, LIM T Y, et al. Semantic image inpainting with perceptual and contextual losses [EB/OL]. (2016-07-26) [2021-07-25]. https://arxiv. org/abs/1607.07539.

[17] GUO Z, CHEN Z, YU T, et al. Progressive image inpainting with full-resolution residual network[C]//27th ACM International Conference on Multimedia, October 11-25, 2019, New York, NY, USA. New York: ACM, 2019:2496-2504.

[18] ZHANG H, HU Z, LUO C, et al. Semantic image inpainting with progressive generative networks[C]//26th ACM in-

ternational conference on Multimedia, October 22-26, 2018, Seoul, Korea (South). New York: ACM, 2018: 1939-1947.

[19] XIONG W, YU J, LIN Z, et al. Foreground-aware image inpainting[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 5840-5848.

[20] NAZERI K, NG E, JOSEPH T, et al. Edgeconnect: structure guided image inpainting using edge prediction[C]//IEEE/CVF International Conference on Computer Vision, October 27-28, 2019, Seoul, Korea (South). New York: IEEE, 2019.

[21] CRIMINISI A, PEREZ P, TOYAMA K. Object removal by exemplar-based inpainting[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 16—22, 2003, Madison, Wisconsin, USA. New York: IEEE, 2003.

[22] YANG C, LU X, LIN Z, et al. High-resolution image inpainting using multi-scale neural patch synthesis[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, Hawaii, USA. New York: IEEE, 2017: 6721-6729.

[23] LIU G, REDA F A, SHIH K J, et al. Image inpainting for irregular holes using partial convolutions[C]//European Conference on Computer Vision, September 8-14, 2018, Munich, Germany. Berlin: Springer, 2018.

[24] YU J, LIN Z, YANG J, et al. Free-form image inpainting with gated convolution[C]//IEEE/CVF International Conference on Computer Vision, October 27-28, 2019, Seoul, Korea (South). New York: IEEE, 2019: 4471-4480.

[25] LI K, WEI Y, YANG Z, et al. Image inpainting algorithm based on TV model and evolutionary algorithm[J]. Soft Computing, 2016, 20(3): 885-893.

[26] LI J, WANG N, ZHANG L, et al. Recurrent feature reasoning for image inpainting[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 14-19, 2020, Seattle, WA, USA. New York: IEEE, 2020: 7760-7768.

[27] DROZDZAL M, VORONTSOV E, CHARTRAND G, et al. The importance of skip connections in biomedical image segmentation[M]//Deep learning and data labeling for medical applications. Berlin: Springer, Cham, 2016: 179-187.

[28] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2014-09-04)[2021-07-25]. https://arxiv.org/abs/1409.1556.

[29] GATYS L A, ECKER A S, BETHGE M. A neural algorithm of artistic style[EB/OL]. (2015-08-26)[2021-07-25]. https://arxiv.org/abs/1508.06576.

作者简介：

李良福 (1977—)，男，江西人，博士，教授，硕士生导师，主要从事计算机视觉方面、图像处理、人工智能等方面的研究.