

基于Transformer多分辨率特征融合的图像压缩感知重构

熊承义^{ab}, 马帅^{ab}, 高志荣^c, 李帆^{ab}, 陈文旗^{ab}

(中南民族大学 a. 电子信息工程学院; b. 智能无线通信湖北省重点实验室; c. 计算机科学学院, 武汉 430074)

摘要 利用图像多分辨率特征的交叉融合, 对于改善压缩感知图像的重构质量具有较好潜能. 研究了一种基于Transformer多分辨率特征融合的图像压缩感知重构方法. 输入图像的测量值首先经过初始重构, 得到一组分辨率降维的低分辨率初始重构图像; 然后, 采用两个通路并行提取不同分辨率图像的特征并进行交叉融合; 最后, 将输出的两路特征分别用于原始图像的重构及其降采样重构. 采用Transformer网络执行多分辨率图像特征的交叉融合, 以更好利用图像的远距离相关性. 大量实验比较结果验证了所提出的方法在平衡网络复杂度和改进重构图像质量方面的有效性.

关键词 多分辨率特征; 压缩感知; 交叉融合; Transformer方法

中图分类号 TP391.4 文献标志码 A 文章编号 1672-4321(2025)03-0400-07

doi: 10.20056/j.cnki.ZNMDZK.20250314

Deep reconstruction for compressed sensing based on multi-resolution feature fusion with Transformer

XIONG Chengyi^{ab}, MA Shuai^{ab}, GAO Zhirong^c, LI Fan^{ab}, CHEN Wenqi^{ab}

(South-Central MinZu University, a.College of Electronic and Information Engineering; b.Hubei Key Lab of Intelligent Wireless Communication; c.College of Computer Science, Wuhan 430074, China)

Abstract The cross-fusion of multi-resolution features of images has good potential to improve the reconstruction quality of compressed sensing images. An image compressive sensing reconstruction method based on multi-resolution feature fusion with Transformer has been studied. The measurement values of the input image are initially reconstructed to obtain a set of low-resolution images with reduced dimension. Then two channels are used to extract the features cross-fused of the different resolution images in parallel. Finally, the features output by the two channels are used to reconstruct the original image and reconstruct the downsampled image. Transformer is used to cross-fuse the features of multi-resolution images to better utilize the long-range correlation of the images. Extensive experimental results are compared to verify the effectiveness of the proposed method in balancing the complication of the network and improving image reconstruction quality.

Keywords multi-resolution features; compressed sensing; cross-fusion; Transformer method

压缩传感理论(Compressive Sensing, CS)证明了如果信号具有稀疏性, 则可以低于奈奎斯特采样率采样信号, 并能够被准确恢复^[1]. 基于压缩感知的信号采样可以极大节省信号获取和传输的成本, 因此

该理论自提出以来已广泛应用于单像素相机^[2]、磁共振成像、快照和视频压缩^[3]等领域.

信号的压缩感知可分为压缩采样过程和重构过程. 信号的压缩采样过程可以数学建模表示为 $y =$

收稿日期 2022-11-21

作者简介 熊承义(1969-), 男, 教授, 博士, 研究方向: 图像处理与模式识别, E-mail: xiongcy@mail.scuec.edu.cn

基金项目 多谱信息处理技术国家重点实验室基金资助项目(6142113210303); 中央高校基本科研业务专项资金资助项目(CZY19016)

$\Phi x \in \mathbb{R}^M$,其中 $x \in \mathbb{R}^N$ 为输入信号($M \ll N$), $y \in \mathbb{R}^M$ 为其采样值, Φ 为 $M \times N$ 的采样矩阵, M/N 为采样率.重构过程旨在根据测量值 y 恢复出原始信号 x ,由于 $M \ll N$,其在数学上为欠定方程求解问题,为典型的病态逆问题.为了得到准确解,传统的压缩感知重建建模为规则化最小化优化问题: $\operatorname{argmin} \frac{1}{2} \|y - \Phi x\|_2^2 + \lambda \mathcal{R}(x)$,其中 $\mathcal{R}(\cdot)$ 为规则化项, λ 为规则化系数.

近年来,随着深度学习的兴起,基于深度学习的压缩感知重构方法已成为当前主流.基于深度学习的压缩感知重构通常可分为两类,一是基于传统优化方法^[4-6]的深度展开方法^[7-10],二是基于直接深度网络的方法^[11-16].基于深度学习的重构网络目前大多基于多层卷积神经网络(Convolutional Neural Network, CNN),但是最近随着视觉Transformer网络的提出及其在图像分类识别中的成功应用^[17],基于视觉Transformer的图像重构方法研究也得到了大量关注^[18-20].Transformer网络与卷积神经网络相比,能够更好地提取图像的长距离相关性,因此能极大地提升图像重构性能^[21].

基于视觉Transformer的良好特性,结合考虑不同尺度图像特征互补增强对压缩感知图像重构的潜在影响,本文提出了一种基于视觉Transformer多分辨率特征融合的压缩感知重构方法.主要思想在于利用Transformer模块提取不同尺度图像的特征相关性,增强深度网络的学习能力,从而改进重构图像质量.具体地,重构网络首先基于输入的测量值得到一组多幅图像的低分辨率初始重构,然后分别送入两路并行的多级特征提取单元,并最终分别得到原始图像的重构以及降维图像的重构.Transformer模块结合跨注意力学习策略以融合两路特征实现特征增强.大量实验比较结果验证了,本文方法在平衡网络复杂度和改进重构图像质量方面的有效性.

1 相关工作

1.1 基于卷积神经网络的CS重构

在过去几年,深度神经网络被广泛应用于图像CS重构中,其中基于深度学习的直接重构方法受到了极大的关注.与深度展开方法相比,深度直接方法未使用某种特定先验,而是使用基于CNN的网络结构,这也使得深度直接方法拥有更快的重构速度.比如KULKARNI等提出的ReconNet^[13]被认为是第一个将CNN用于CS重建的方法,旨在通过CNN从

测量值中恢复原始图像,重构质量和计算复杂度均优于传统的迭代算法.DR2-Net^[12]采用全连接层进行采样,并堆叠多个残差块以提高重建质量.SUN等还提出了一种用于CS重建的双路注意力网络DPA-Net^[16],该网络使用双通路结构,分别学习结构特征和纹理特征,然后通过注意力模块进行融合.SHI等在SCSNet^[15]中实现了采样率可伸缩,该方法采用多个并行网络层的设计,每一个网络层用来实现不同采样率的CS重构,同时低采样率网络层的特征会融合至下一个高采样率网络层,以增强高采样率下的CS重构质量.

1.2 基于Transformer的图像恢复

Transformer最初是为了解决自然语言处理(NLP)中长距离依赖的问题,现已广泛应用于计算机视觉各领域中,并取得不错的效果.CHEN等设计了一个用于图像处理的TransformerIPT(Image Processing Transformer)的预训练模型^[18],使用大规模的数据对IPT进行预训练,使得IPT在超分、去雨、去噪等领域均获得不错的效果.WANG等提出的用于图像恢复的U型Transformer网络Uformer^[19],借鉴了U-Net^[22]的思想,并结合Transformer来构建整体网络结构,在去噪与去雨等底层视觉任务上均获得了提升.

在基于参考图像的超分领域,YANG等提出了一个学习相似纹理的Transformer网络TTSR^[20](Texture Transformer Network for Image Super-Resolution),该网络将低分辨率图像和参考图像作为Transformer的输入,旨在从参考图像中提取相似的纹理信息,以获得更高质量的高分辨率图像.值得一提的是,多数研究更加关注使用Transformer的自注意力特性来提取特征的全局表达,而在TTSR中,Transformer能够从不同的图像中提取相似的纹理.受到该方法的启发,本文使用Transformer在多分辨率特征融合方面同样取得了不错的效果.

2 提出的方法

2.1 网络整体结构

网络整体结构如图1所示,整体可以分为三个子网络:采样网络、初始重构网络和深度重构网络,其中, F_{b1}^i 与 F_{b2}^i 则分别表示原始图像重构分支与2倍降采样重构分支中第 i 个残差块的输出(i 为0时表示卷积层的输出特征), F_{t}^i 为第 i 个Transformer模块的输出.采样网络使用基于学习的分块卷积采样方法,得到指定采样率的测量值;初始重构网络通过卷积和Reshape操作实现,得到一组低分辨率图像;

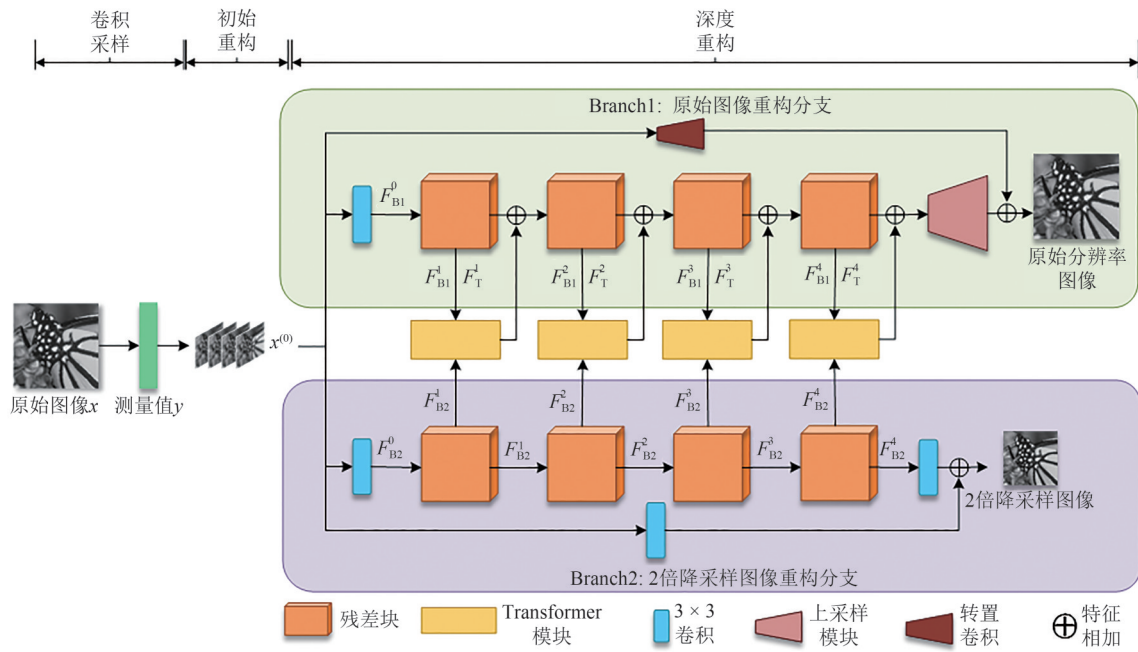


图1 网络整体结构图

Fig. 1 Overall structure of the proposed network

深度重构网络使用双分支结构分别实现原始图像和2倍降采样图像重构,通过Transformer模块提取两个分支的相关特征,并融合至原始图像重构分支,以增强压缩感知重构效果.

2.2 采样网络

为避免采样网络复杂度过高,大部分CS网络会使用BCS(block-based CS)^[23]方法进行采样.具体地,将原始图像 $x \in \mathbb{R}^{H \times W \times C}$ 分为多个大小为 $B \times B$ 的非重叠图像块,图像块的数量 $n = \frac{H}{B} \times \frac{W}{B} \times C$.假定 x_i 为 x 的第 i 个图像块,那么 x_i 的测量值 y_i 可表示为:

$$y_i = \Phi x_i, i = 1, 2, \dots, n, \quad (1)$$

其中, m/B^2 为采样率且测量矩阵 $\Phi \in \mathbb{R}^{m \times B^2}$.那么, x 的采样值 y 可表示为所有图像块采样值的集合:

$$y = [y_1, y_2, \dots, y_n], i = 1, 2, \dots, n, \quad (2)$$

本文使用无偏卷积层 W_s 来实现BCS采样过程,其中, W_s 的卷积核大小设置为 B ,步长设置为 B ,卷积核个数则设置为 m .卷积采样过程可表示为:

$$y = W_s \otimes x, \quad (3)$$

其中: \otimes 为卷积操作, y 为原始图像对应的CS采样值.

2.3 初始重构网络

初始重构网络将采样网络输出的测量值作为输入,并输出为一组2倍降采样初始重构图像.本文使用卷积层 W_{ini} 来实现初始重构过程,其中, W_{ini} 含有 B^2 个大小为 $1 \times 1 \times m$ 的卷积核,且步长设置为1.

将卷积层 W_{ini} 的输出进行Reshape后即可得到一组低分辨率图像 $x^{(0)} \in \mathbb{R}^{(H/2) \times (W/2) \times 4C}$.根据上述过程,本文的初始重构过程可表示为:

$$x^{(0)} = \text{Reshape}(W_{\text{ini}} \otimes y). \quad (4)$$

初始重构图像大小设定为原始图像2倍下采样的大小,一方面,可以在深度重构阶段实现多分辨率特征提取,增强网络的特征表示能力,提升重构质量;另一方面,由于深度重构网络采用双通路结构,若将初始重构图像大小设定为原始分辨率图像大小,则会增加深度重构网络的复杂度.具体地,假设重构网络中的某一卷积层 W_k 的卷积核大小为 $K \times K$,其输入与输出的通道数均为 M .由于原始图像重构分支的特征在上采样之前的大小与初始重构图像的大小一致,若初始重构图像为 $x^{(0)} \in \mathbb{R}^{(H/2) \times (W/2) \times 4C}$,则使用卷积层 W_k 的计算量 Ω 可表示为:

$$\Omega = K^2 H W M^2 / 4, \quad (5)$$

但如果初始重构图像为 $x^{(0)} \in \mathbb{R}^{H \times W \times C}$,即为原始分辨率图像大小,则计算量为:

$$\Omega = K^2 H W M^2, \quad (6)$$

由此可见:当初始重构图像大小为2倍降采样大小时,相同的网络结构和参数量下能够具有更少的计算量.

2.4 深度重构网络

深度重构模块如图1所示,使用双分支结构分别实现原始图像重构和2倍降采样重构.实现原始图像重构的分支作为主干支路,实现降采样重构的分支作

为辅助支路.利用Transformer模块实现主干支路与辅助支路的特征融合.主干支路包含一个转置卷积模块用于获取一幅低质量原始图像,以及并行的由多个堆叠的残差块与上采样模块用于获取残差图像.其中,每个残差块包含两个 3×3 卷积层与一个ReLU激活函数,上采样模块通过亚像素层进行上采样.

辅助支路包含一个 3×3 卷积模块获取低质量的降采样图像,以及并行的与主干支路同样的残差块实现多级特征提取,并通过最后卷积模块生成降维图像的残差图.主干支路与辅助之路对应残差块的输出特征通过互联的Transformer模块进行上下通路的特征融合与增强,具体表示为:

$$F_{B1}^i = \begin{cases} H_{B1}^i(F_{B1}^0), & i = 1, \\ H_{B1}^i(F_T^{i-1} + F_{B1}^{i-1}), & 1 < i \leq N, \end{cases} \quad (7)$$

$$F_{B2}^i = H_{B2}^i(F_{B2}^{i-1}), \quad 1 \leq i \leq N, \quad (8)$$

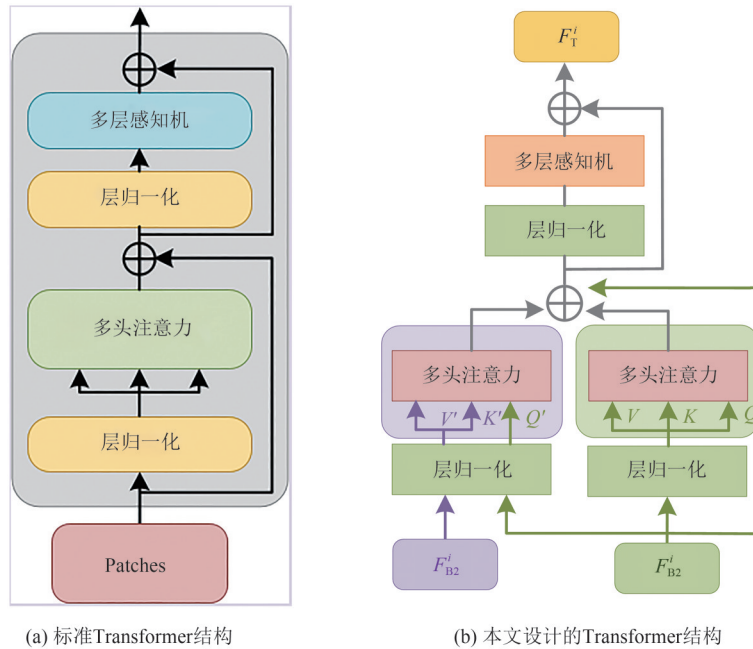


图2 Transformer结构示意图

Fig. 2 Schematic diagram of the structure of Transformer

本文的Transformer模块使用了两路并行的多头注意力,如图2(b)所示.一路使用跨注意力提取多分辨率特征之间的相关性,以获取辅助分支中的相关特征,实现增强原始图像重构质量的目的,其输入 Q' 来自主干分支特征 F_{B1}^i , K' 和 V' 均来自辅助分支特征 F_{B2}^i ;另一路使用自注意力对 F_{B1}^i 进行全局建模,以增强网络的特征表示能力,其输入均来自 F_{B1}^i .

Transformer模块会将两路多头注意力所提取的特征相加融合,然后再经过前馈网络,从而获取多分辨率融合特征.整个过程表示如下:

$$Z_1 = \text{MSA}(\text{LN}(F_{B1}^i), \text{LN}(F_{B2}^i), \text{LN}(F_{B2}^i)), \quad (10)$$

其中: N 表示各分支残差块的数目, H_{B1}^i 与 H_{B2}^i 分别表示主干分支与辅助分支的第 i 个残差块的运算, F_{B1}^i 与 F_{B2}^i 则分别表示主干分支与辅助分支中第 i 个残差块的输出(i 为0时表示卷积层的输出特征), F_T^i 表示第 i 个Transformer模块的输出.

2.5 Transformer特征融合模块

标准的Transformer^[24]采用基于自注意力的结构,如图2(a)所示.当自注意力结构的输入相同时,即 QKV 处于同一特征空间,可以通过内积来估计当前特征空间内不同patch之间的相关性大小;当输入不同时,则能以跨注意力的方式得到不同特征之间的相关性.在标准的Transformer中,多头注意力可以表示为:

$$X = \text{MSA}(Q, K, V), \quad (9)$$

其中:MSA表示多头注意力, X 表示MSA的输出,且输入 QKV 相同.

$$Z_2 = \text{MSA}(\text{LN}(F_{B1}^i), \text{LN}(F_{B1}^i), \text{LN}(F_{B1}^i)), \quad (11)$$

$$Z_3 = Z_1 + Z_2 + F_{B1}^i, \quad (12)$$

$$F_T^i = \text{MLP}(\text{LN}(Z_3)) + Z_3, \quad 1 < i \leq N, \quad (13)$$

其中: N 为Transformer模块数目, LN 表示层归一化, Z_1 、 Z_2 、 Z_3 表示Transformer计算过程中对应的中间特征.

3 实验

3.1 实验设置和训练数据

本文实验数据由BSD500数据集[25]中的训练

集(200张)和测试集(200张)组成,然后从每幅图像上裁剪出14个大小为 96×96 的图像块,再将每个图像块进行翻转、旋转 90° 、旋转 180° 、旋转 270° 操作,最终扩展至89600张。BatchSize设置为32,一共训练100个epoch。初始学习率设置为0.0004,分别在20、50、80、90训练周期时将学习率衰减为原来的0.5倍。使用Python语言在Pytorch框架下使用Adam优化器优化网络参数,在Ubuntu22.04系统上使用一张12 GB的Nvidia TITANX显卡进行训练。

3.2 损失函数

为了能够有效的从降采样重构分支提取相关特征,增强原始图像重构质量,本文方法使用了两个损失函数,分别用来约束生成的原始图像与降采样图像。因此,本文将总损失设置为如下形式:

$$\mathcal{L}_{\text{all}} = \lambda_{\text{H}} \mathcal{L}_{\text{H}} + \lambda_{\text{L}} \mathcal{L}_{\text{L}}, \quad (14)$$

其中: \mathcal{L}_{H} 与 \mathcal{L}_{L} 分别表示原始图像重构损失、降采样图像重构损失, λ_{H} 与 λ_{L} 为对应的损失权重。本文利用了L2损失相比于L1损失在梯度求导上的优势,将原始图像重构损失 \mathcal{L}_{H} 与降采样图像重构损失 \mathcal{L}_{L}

定义为如下形式:

$$\mathcal{L}_{\text{H}} = \frac{1}{N} \sum_{i=1}^N \|\hat{x}_i^{\text{H}} - x_i^{\text{H}}\|_2^2, \quad (15)$$

$$\mathcal{L}_{\text{L}} = \frac{1}{N} \sum_{i=1}^N \|\hat{x}_i^{\text{L}} - x_i^{\text{L}}\|_2^2, \quad (16)$$

这里使用 N 表示训练数据的Batchsize大小, \hat{x}_i^{H} 、 \hat{x}_i^{L} 分别表示重构出的原始分辨率图像与降采样图像, x_i^{H} 、 x_i^{L} 分别表示原始图像、对原始图像使用双三次插值进行2倍下采样后的低分辨率图像。

3.3 量化性能对比

为了验证本文方法在提高重构性能方面的有效性,本文与现有的CS重构方法在定量和定性方面进行了对比,对比的方法包括ISTA-Net+[7]、CSNet+[11]、SCSNet[15]等主流的CS重构方法。表1列出了在5种采样率下,不同CS重构方法在测试数据集Set11上的PSNR和SSIM,其中,各重构方法的指标大多为原文提供的结果,原文中未提供的结果,则根据原文中指定的训练策略复现论文代码得到。可以看出,在Set 11上本文方法在所有采样率均可获得最高的定量指标。

表1 Set11上不同CS采样率的PSNR(dB)/SSIM对比

Tab. 1 PSNR(dB)/SSIM performance comparisons with different CS ratios on Set11

方法\采样率	0.01	0.1	0.3	0.4	0.5
DR ² -Net ^[12]	17.59/0.4198	24.50/0.7326	29.54/0.8722	31.16/0.8999	32.63/0.9181
ReconNet ^[13]	17.56/0.4170	23.29/0.6720	23.30/0.6813	23.22/0.6811	26.08/0.7951
CSNet+ ^[11]	21.03/0.5566	28.37/0.8580	34.30/0.9490	36.34/0.9677	38.52/0.9749
ISTA-Net+ ^[7]	17.45/0.4131	26.49/0.8036	33.70/0.9382	34.32/0.9367	35.54/0.9447
SCSNet ^[15]	21.04/0.5562	28.48/0.8616	34.62/0.9511	36.92/0.9538	39.01/0.9651
本文	21.28/0.5603	29.08/0.8859	35.14/0.9591	37.43/0.9714	39.77/0.9813

为了证明本文方法在视觉效果上的提升,本文与各CS重构方法进行了视觉效果对比。图3与图4分别给出了0.3采样率下各CS方法在barbara上、0.4采样率下各CS方法在butterfly上的局部细节的视觉效果,通过观察可以看出本文方法的重构结果在barbara的头巾纹理处与butterfly的黑斑与白斑分界更加清晰,具有更好的视觉效果。结合定量和定性分析的结果,可以验证本文提出的方法可以获得更好的重构效果。

3.4 消融实验

3.4.1 单路特征增强

在本文中,Transformer模块提取的特征相关性既可用于增强主干分支特征(如图1所示),也可同时用于增强辅助分支特征。考虑到增强辅助分支重

构性能对提升主干分支性能的潜在影响,本文对此进行了相关实验。具体地,在不改变图1网络结构的基础上,增加一个Transformer模块用于从主干分支提取相关特征,并融合至辅助分支。本文对两种方案的训练时间与重构性能进行了比较,结果如表2所示。可以看出:双路特征增强相对于单路特征增强,训练时间增加,但重构效果提升有限。因此,本文方法最终选取单路特征增强策略。

3.4.2 多分辨率特征融合方式

本文方法在深度重构阶段使用Transformer模块来实现多分辨率特征融合。为了验证该融合方式的有效性,本文将Transformer融合方法与相加融合的方法进行了对比,其中,相加融合是指将辅助分支的特征以相加的方式融合至主干分支。通过表3实

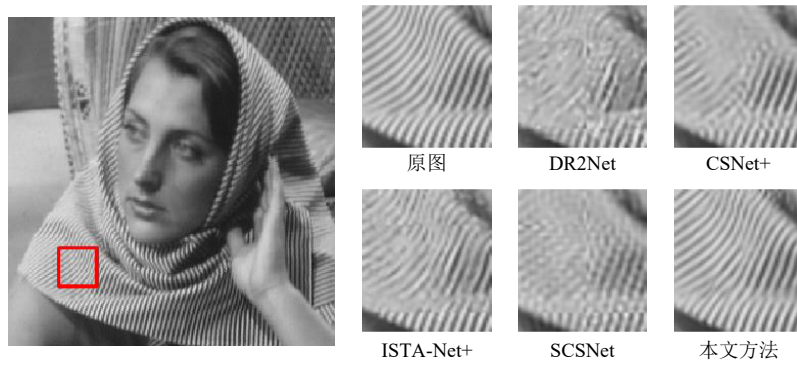


图 3 0.3 采样率下不同 CS 算法对 barbara 重构效果

Fig. 3 Reconstruction results of barbara by different algorithms at 0.3 sampling rate

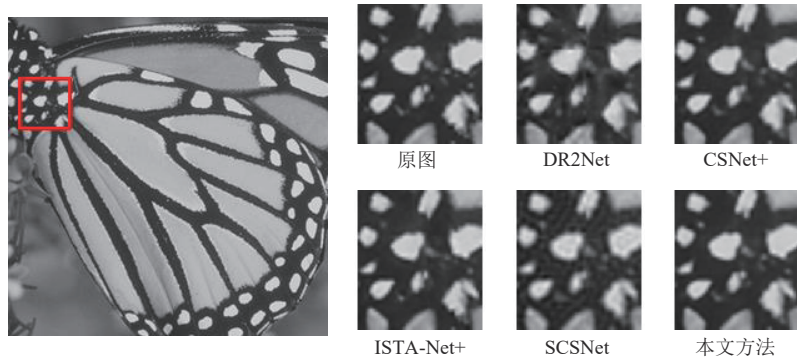


图 4 0.4 采样率下不同 CS 算法对 butterfly 重构效果

Fig. 4 Reconstruction results of butterfly by different algorithms at 0.4 sampling rate

表 2 0.5 采样率下使用不同特征增强策略在 Set11 上的效果对比

Tab. 2 Comparison of results when using different feature enhancement strategy at 0.5 sampling rate on Set11 test datasets

特征增强策略	PSNR	SSIM	训练时间/h
单路	39.77	0.9813	29
双路	39.78	0.9815	50

验结果可以看出,使用 Transformer 进行多分辨率特征融合比相加融合更加有效.

表 3 0.5 采样率下使用不同融合方式在各测试集上的 PSNR (dB) 与 SSIM 对比

Tab. 3 Comparison of the PSNR (dB) and SSIM when using different fusion methods at 0.5 sampling rate on different test datasets

测试集	Set5	Set11	Set14
相加融合	41.97/0.9837	39.46/0.9806	38.28/0.9702
Transformer 融合	42.40/0.9844	39.77/0.9813	38.74/0.9714

3.4.3 自/跨注意力

本文的 Transformer 模块使用了双路并行的多头注意力设计,即在跨注意力通路的基础上,增加了自注意力通路.为验证增加的自注意力通路对网络性能的影响,本文进行了相关实验.通过表 4 实验结果可以看出,Transformer 模块使用自注意力通路可以进一步增强重构质量.

表 4 0.4 采样率下使用不同 Transformer 模块在 Set11 上的效果对比

Tab. 4 Comparison of results when using different Transformer modules at 0.4 sampling rate on Set11 test datasets

自注意力	跨注意力	PSNR	SSIM
×	√	37.19	0.9713
√	√	37.43	0.9714

4 结语

本文提出了一种基于 Transformer 多分辨率特征融合的图像压缩感知重构方法,该方法可以较好的改善压缩感知重构性能.具体地,采样阶段通过卷积采样来获取测量值,初始重构阶段则根据测量值重构出一组低分辨率图像,在深度重构阶段使用 Transformer 对初始重构图像进行多分辨率特征提取并进行交叉融合,以进一步提升 CS 重构效果.实验表明,与现有的基于深度学习的方法相比,本文方法能够较好平衡复杂度与重构质量.

参 考 文 献

[1] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289-1306.
 [2] DUARTE M F, DAVENPORT M A, TAKHAR D, et al. Single-pixel imaging via compressive sampling[J]. IEEE

- Signal Processing Magazine, 2008, 25(2): 83-91.
- [3] LI Y, DAI W, ZOU J, et al. Structured sparse representation with union of data-driven linear and multilinear subspaces model for compressive video sampling [J]. IEEE Transactions on Signal Processing, 2017, 65(19): 5062-5077.
- [4] ZHANG J, ZHAO D, ZHAO C, et al. Image compressive sensing recovery via collaborative sparsity [J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2012, 2(3): 380-391.
- [5] ZHANG J, ZHAO C, ZHAO D, et al. Image compressive sensing recovery using adaptively learned sparsifying basis via L0 minimization [J]. Signal Processing, 2014, 103: 114-126.
- [6] AHSEN M, VIDYASAGAR M. Error bounds for compressed sensing algorithms with group sparsity: A unified approach[J]. Applied and Computational Harmonic Analysis, 2017, 43(2): 212-232.
- [7] ZHANG J, GHANEM B. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 1828-1837.
- [8] 熊承义, 李世宇, 高志荣, 等. 级联模型展开与残差学习的压缩感知重构[J]. 中南民族大学学报(自然科学版), 2019, 38(2): 265-272.
- [9] ZHANG Z, LIU Y, LIU J, et al. AMP-Net: Denoising-based deep unfolding for compressive image sensing [J]. IEEE Transactions on Image Processing, 2020, 30: 1487-1500.
- [10] ZHANG J, ZHAO C, GAO W. Optimization-inspired compact deep compressive sensing [J]. IEEE Journal of Selected Topics in Signal Processing, 2020, 14(4): 765-774.
- [11] SHI W, JIANG F, LIU S, et al. Image compressed sensing using convolutional neural network [J]. IEEE Transactions on Image Processing, 2019, 29: 375-388.
- [12] YAO H, DAI F, ZHANG S, et al. DR²-Net: Deep Residual Reconstruction Network for image compressive sensing [J]. Neurocomputing, 2019, 359: 483-493.
- [13] KULKARNI K, LOHIT S, TURAGA P, et al. ReconNet: Non-iterative reconstruction of images from compressively sensed measurements [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 449-458.
- [14] 熊承义, 秦鹏飞, 高志荣, 等. 基于双路信息互补增强的压缩感知深度重构[J]. 中南民族大学学报(自然科学版), 2020, 39(6): 619-624.
- [15] SHI W, JIANG F, LIU S, et al. Scalable convolutional neural network for image compressed sensing [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 12282-12291.
- [16] SUN Y, CHEN J, LIU Q, et al. Dual-path attention network for compressed sensing image reconstruction [J]. IEEE Transactions on Image Processing, 2020, 29: 9482-9495.
- [17] ASHISH V, NOAM S, NIKI P, et al. Attention is all you need [C]//Conference on Neural Information Processing Systems. Long Beach: NIPS, 2017: 1, 2, 4, 5.
- [18] 熊承义, 郑瑞华, 高志荣, 等. 结合多尺度多注意力的遥感图像超分辨率重构[J]. 中南民族大学学报(自然科学版), 2024, 43(5): 692-700.
- [19] 熊承义, 陈文旗, 高志荣, 等. 局部特征增强的磁共振图像Transformer重构[J]. 中南民族大学学报(自然科学版), 2025, 44(1): 50-57.
- [20] 高志荣, 孙清清, 熊承义, 等. 基于ViT-CNN特征增强的图像超分辨率[J]. 中南民族大学学报(自然科学版), 2025, 44(2): 253-259.
- [21] LIANG J, CAO J, SUN G, et al. SwinIR: Image restoration using swin transformer [C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Montreal: IEEE, 2021: 1833-1844.
- [22] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.
- [23] GAN L. Block compressed sensing of natural images [C]//2007 15th International Conference on Digital Signal Processing. Cardiff: IEEE, 2007: 403-406.
- [24] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x6 words: Transformers for image recognition at scale [J]. arXiv Preprint arXiv: 2010.11929, 2020.
- [25] ARBELÁEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.

(责编&校对 刘钊)