

文章编号: 2095-2163(2024)01-0056-08

中图分类号: TP399

文献标志码: A

基于增强隐式神经表示的图像超分辨算法研究

霍旭峰, 张选德

(陕西科技大学 电子信息与人工智能学院, 西安 710021)

摘要: 隐式神经表示为数字图像连续表示提供了一种方法,该方法已成功应用于图像超分辨任务中,并能够取得良好的性能。但是,由于其像素级采样的插值策略,导致权重分配失衡,使得恢复出的高分辨率图像边缘、纹理过平滑;同时由深度网络提取的特征图存在底层特征失真的问题。针对上述问题,本文提出一种基于增强隐式神经表示的图像超分辨重构算法(WCESR)。方法中引入权重修正模块,学习局部面积权重与全局结构权重的关系,缓解权重分配失衡现象;同时引入低分辨率图像的边缘特征,扩展由深度神经网络得到的深层图像特征,产生锐利的边缘。通过大量对比实验和消融实验证明:本方法可以得到与现有算法相当甚至更好的效果。

关键词: 图像超分辨; 隐式神经表示; 权重修正; 边缘特征扩展

Image super-resolution based on enhanced implicit neural representation

HUO Xufeng, ZHANG Xuande

(School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an 710021, China)

Abstract: Implicit neural representation provides a method for continuous representation of digital images. This method has been successfully applied to image super-resolution tasks and can achieve good performance. However, due to the imbalance of weight distribution caused by the interpolation strategy of pixel-level sampling, the edge and texture of the restored high-resolution image are too smooth; At the same time, the feature map extracted from the depth network has the problem of distortion of the underlying features. To solve these problems, this paper proposes an image super-resolution reconstruction algorithm based on enhanced implicit neural representation, called WCESR. This method introduces the weight correction module to learn the relationship between the local area weight and the global structure weight, and alleviate the imbalance of weight distribution. At the same time, the edge feature of low resolution image is introduced to extend the deep image feature obtained by depth neural network to produce sharp edges. Through a large number of comparative experiments and ablation experiments, it is proved that this method can get equivalent or even better results than the existing algorithms.

Key words: image super-resolution; implicit neural representation; weight correction; edge feature extension

0 引言

对于计算机而言,为了提高存储和计算处理的效率,数字图像一般是通过离散像素点存储和计算的。在这种情况下,像素分辨率成为了评估图像质量的一个重要度量标准。因此,旨在提高数字图像分辨率的单帧图像超分辨重构(Single Image Super-Resolution, SISR),已成为计算机视觉领域的一个基本问题。

图像超分辨(Super-Resolution, SR)是指从低分辨率(Low-Resolution, LR)图像得到高分辨率

(High-Resolution, HR)图像的过程,是图像恢复领域中非常重要的任务^[1-2]。随着深度学习的发展,其已成为解决SR问题的主要方法之一。SRCNN^[3-4]采用预上采样超分辨框架,通过3个卷积层,学习从插值的低分辨率图像得到高分辨率图像的端到端映射。EDSR^[5]去除了传统残差网络中不必要的模块,从而得到了较好的图像超分辨效果。然而,这些方法对不同缩放因子需要训练单独的网络,因此在实际应用中受到很大的限制。MDSR^[5]提出了预处理模块和特定缩放因子的上采样模块,并将其放在网络的前后位置,通过单一网络恢复出

作者简介: 霍旭峰(1999-),男,硕士研究生,主要研究方向:图像超分辨。

通讯作者: 张选德(1979-),男,硕士,教授,博士生导师,主要研究方向:图像恢复、图像质量评价、视觉追踪等。Email:love_truth@126.com

收稿日期: 2023-09-24

多个缩放因子的图像超分辨结果,但 MDSR 仍只能生成特定缩放因子的图像超分辨结果。Meta-SR^[6]提出了元上采样模块(Meta-Upscale Module),利用最近邻规则,将 SR 像素映射到 LR 域上,动态的为每一个 SR 像素生成卷积权重。但该方法在推广到非训练缩放因子的 SR 时表现不佳,产生了类似棋盘的伪影。如何更高效的运用单个网络实现任意缩放因子或任意分辨率图像超分辨成为了 SR 的一个热点问题。

受三维视觉领域中广泛应用的隐式函数启发,隐式函数也被应用到了 SISR 任务中。研究人员利用神经网络来逼近这个隐式函数,这种方法被称为隐式神经表示(Implicit Neural Representation, INR),通过 $v = f(x, y)$,将空间坐标 $(x, y) \in R^2$ 映射到对应的 RGB 值 $v \in R^3$ 来表示二维图像。由于 INR 可以将空间坐标映射为 RGB 值,因此可以通过 INR 解除图像分辨率或 SR 缩放因子固定的限制。Dupont 等人^[7-8]提出了利用生成对抗网络学习图像的隐式神经表示,并产生了较为自然和真实的结果,但这些方法只能通过以噪声向量作为输入来得到输出,而不能将真实图像转换为相应的表示。LIIF^[9]用一个 SISR 自监督任务,训练出一种基于编码器的局部隐式图像函数(Local Implicit Image Function),并利用连续坐标采样的灵活性,实现了任意分辨率的图像表示。由于在恢复图像的过程中,LIIF 采用了双线性插值的策略来恢复图像,使得恢复的图像是模糊的,没有较为锐利的边缘。

为了解决这些问题,本文研究了一种基于隐式神经表示的图像超分辨重构方法(WCESR)。方法中设计了一个权重修正模块,以隐式神经表示的双线性插值权重为输入,通过学习局部面积权重与全局结构权重之间的关系,输出与全局结构高度相关的权重信息,缓解由于双线性插值策略导致的权重分配失衡问题。另外,由于深度网络存在底层特征失真的问题,引入 LR 图像的边缘信息,指导恢复图像的整体结构,约束图像的平滑性。经过实验,WCESR 可以得到与 LIIF 相当甚至更好的效果。

1 相关工作

1.1 单帧图像超分辨

图像超分辨重构是指从低分辨率图像中恢复高分辨率图像的过程,是计算机视觉和图像处理中的一类重要的图像处理技术,有着非常广泛的现实应用价值。如:医学成像、监控和安全等等。除了提高

图像感知质量外,其还有助于改善其他计算机视觉任务。一般来说,SR 问题是非常具有挑战性的,并且具有高度的不稳定性,因为总是有多个 HR 图像对应于一个 LR 图像。

传统 SR 方法可以分为基于图像高低分辨率块的^[1-2]、基于边缘的^[10-11]和基于统计的^[12]方法。首个基于卷积神经网络的 SISR 工作是 SRCNN^[3],其中 3 个卷积层被用于特征提取、特征映射和图像重构。随后,又出现了更大的、基于残差结构的 SR 方法,如 VDSR^[13]和 SRResNet^[14]。EDSR^[5]通过去除残差结构中的 BN 层来改进残差块。Yu 等人^[15]在 EDSR 的 ReLU 层之前引入更宽的通道,进一步增强 EDSR。同时,RDN^[16]提出了 3 个卷积层相互密集连接的残差密集块(Residual Dense Block, RDB),将 SR 质量提高到了一个新的水平。然而这些方法只能处理特定的缩放因子,而对于不同的缩放因子需要训练单独的网络,这大大限制了其实际应用。

1.2 任意缩放因子的图像超分辨

随着深度卷积神经网络的发展,近年来对 SR 的研究取得了巨大的突破,但任意缩放因子的 SR 一直被忽略,相关工作十分匮乏。在以往的研究中,研究人员将不同缩放因子的 SR 视为独立的任务,在这种问题下产生的 SR 模型往往为每个缩放因子训练一个特定的模型,这种模型效率较低;同时,这种方法往往只考虑一些整数倍的缩放因子(例如 2、3、4 等),而任意缩放因子的 SR 在时间和方便性方面明显超过了以往的单帧图像超分辨任务。

将一个神经网络用于多个缩放因子 SR 的方法可以追溯到 Lim 等人提出的 MDSR^[5]。为了实现不同尺度的 SR, MDSR 提出了预处理模块和特定缩放因子的上采样模块,并将其放在网络的前后位置,但 MDSR 只能处理缩放因子为 2、3 和 4 的超分辨任务,并不是一个真正意义上的任意缩放因子 SR 方法。

Meta-SR^[6]是第一个基于 CNN 的、实现任意缩放因子的 SR 方法,其所提出的元上采样模块利用最近邻规则,将 SR 像素映射到 LR 域上。在训练过程中,Meta-SR 的缩放因子在 1 到 4 上连续随机采样生成,所有的特征向量都乘以基于缩放因子和坐标动态生成的权重,输出的图像由元上采样模块和一些卷积层生成。但该方法仅可以在训练尺度范围内进行任意缩放因子的超分辨,若推广到非训练缩放因子的超分辨时表现欠佳,产生了类似插值引起的棋盘效果。

1.3 基于隐式神经表示的图像超分辨

对于数字图像而言,其最常见的表示方式为二维空间上的离散像素点。但在真实世界中所看到的世界(图像),可以认为是连续或者近似连续的。于是,研究人员考虑用一个连续函数来表示图像的真实状态,而这个连续函数的准确形式却无法得知,用神经网络来逼近这个连续函数成功地解决了这个问题,这个方法被称为隐式神经表示。图像、视频等都可以用 INR 进行表示。对于图像,INR 函数将二维坐标映射为 RGB 值。总而言之,在 INR 中,一个对象通常被表示为一个多层感知机 (Multi-layer Perception, MLP),其可以将坐标映射为信号。INR 是一个连续的函数,函数(即网络)的复杂程度和信号的复杂程度成正比,但是与信号的分辨率无关。如:一个 16×16 的图像和一个 32×32 的图像,若内容相同,则 INR 也相同。

近来,隐式神经表示方法研究不是为每个对象学习一个独立的 INR,而是为不同对象的 INR 共享一个函数空间。通常情况下,INR 定义一个隐式空间,其中每个对象通过使用自动编码器,得到一个对应的隐式编码。

虽然 INR 已经在 3D 任务中取得了成功,但其

在图像表示方面还相对不足。早期的 INR 通过组合模式生成网络并将 2D 图像参数化。Sitzman 等人^[17]观察到,之前用 ReLU^[18]参数化的 MLP 进行隐式神经无法表示自然图像的细节,因此用周期激活函数取代 ReLU,并证明其可以在更高质量下建模自然图像。然而,这个方法无法共享隐式函数空间,其泛化性能有限。Chen 等人^[9]提出了一种新的连续表示图像超分辨重构算法 (LIIF),算法运用了一种新的框架。在这个框架中,RGB 值由多层感知机计算,缩放因子和 LR 作为其输入。在 LIIF 中通过隐式神经表示将图像映射到一个连续域中,在恢复超分辨过程中,通过在连续域中对坐标进行采样恢复出超分辨结果。LIIF 在训练中仅在 $\times 1 \sim \times 4$ 上进行训练,却可以扩展到更高的倍数,即使在训练过程中没有训练,LIIF 也因其 SR 可以在很高的缩放因子下表现出十分稳健的性能而被认可。

2 基于增强隐式神经表示的图像超分辨

本文提出的 WCESR 方法,是受 LIIF 启发而改进的。通过隐式神经表示,学习一个隐式图像函数,并通过低分辨率图像的坐标信息与特征图生成高分辨率图像。WCESR 的网络结构如图 1 所示。

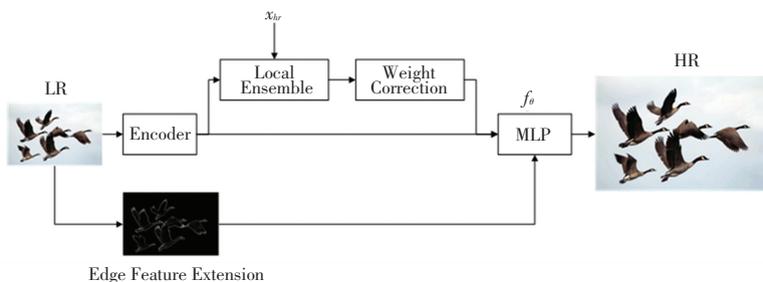


图 1 WCESR 网络架构

Fig. 1 WCESR Network Architecture

图中,LR 表示输入的低分辨率图像,Encoder 是编码器,可以由绝大多数现有的后上采样框架 SR 方法的特征提取模块替代,其以低分辨率图像为输入,输出其对应的特征图;Local Ensemble 是局部集成模块,该模块以局部空间欧氏距离为输入,输出与欧氏距离相关的插值权重;Weight Correction 是权重修正模块;MLP 是隐式神经表示模块,由 5 层多层感知机组成; x_{hr} 是高分辨率图像域的一个二维坐标。

本方法图像隐式形式如式(1)所示:

$$s = f_{\theta}(z, x) \quad (1)$$

其中, z 表示隐式编码,即特征向量; $x \in X$ 表示二维空间内的一个图像坐标; $s \in S$ 表示通过隐式图像函数预测出坐标 x 处的 RGB 值; f_{θ} 是由多层感知机表示的隐式图像函数; $f_{\theta}(z, \cdot)$ 将二维空间坐标映射为 RGB 值。

通过对图像进行像素采样就可以通过 $f_{\theta}(z, \cdot)$ 预测出所有点的 RGB 值。因此,对于一个连续图像 $I^{(i)}$,其坐标为 x_q 处的 RGB 值可以表示为

$$I^{(i)}(x_q) = f_{\theta}(z_q, x_q - v_q) \quad (2)$$

其中, z_q 表示参与预测 x_q 处信号值的特征向量, v_q 表示 z_q 在图像域中的二维坐标。

为了将这个隐式图像函数作用于所有的图像中, 本方法将所有图像映射到高和宽均为 $[-1, 1]$ 的二维连续坐标域中, 如图 2 所示。

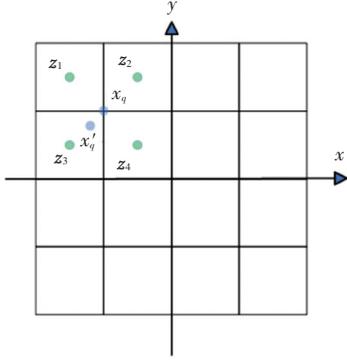


图 2 隐式神经表示的坐标映射

Fig. 2 Coordinate mapping of implicit neural representation

图中每个网格表示离散图像的一个像素点, 将其映射到连续坐标域后, 取每个像素的中心作为这个离散像素点坐标, 如图中绿色点所示。

2.1 权重修正

在隐式神经表示中, 预测一个坐标的 RGB 值时, 该坐标点近邻的特征向量将参与预测。若使用与预测坐标最近邻的特征向量进行预测, 则会出现 z_q 的选择难以确定的情况。如图 2 中, x_q 在二维空间中, 与 z_1, z_2, z_3 和 z_4 的欧式距离大小相等, 此时使用最近邻的策略将出现无法确定使用哪个特征向量进行预测的问题。

为了解决这个问题, LIIF 使用局部集成 (Local Ensemble) 模块, 通过被预测点周围的 4 个特征向量共同参与预测, 如式 (3) 所示:

$$I^{(i)}(x_q) = \sum_t \frac{S_t}{S} f_{\theta}(z_t, x_q - v_t), t \in \{00, 01, 10, 11\} \quad (3)$$

其中, $t \in \{00, 01, 10, 11\}$ 表示与 x_q 最近邻的左上、右上、左下、右下的特征向量; v_t 表示 z_t 的坐标; S_t 表示 x_q 与 v_t 围成矩形的对角矩形面积; S 表示 4 个特征向量围成的矩形面积。

局部集成策略可以解决仅由单一特征向量预测 RGB 值产生的不连续预测, 但以欧氏距离作为权重的方法仍存在权重失衡的问题。如图 2 中点 x_q' 处, 局部集成策略下, z_3 预测的 RGB 值所提供的权重最大, 若在真实 HR 图像中, 图像的一条边缘经过了 z_1, x_q' 和 z_4 而未经过 z_3 , 此时 z_3 则不应提供较大的权重。

为了缓解权重失衡的问题, 本方法提出了权重修正模块, 模块结构如图 3 所示。其中, FC 表示全连接层, ReLU 表示 ReLU 激活函数, Softmax 表示

softmax 激活函数。

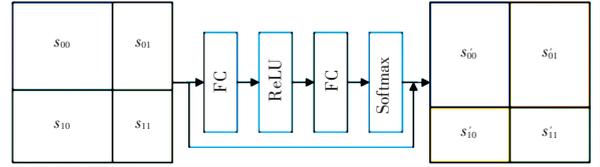


图 3 权重修正网络架构

Fig. 3 Weight correction network architecture

模块以式 (3) 中的 $\frac{S_t}{S}$ 输入, 通过学习局部面积权重与全局结构权重之间的关系, 输出与全局结构高度相关的权重信息, 其表达式如式 (4) 所示:

$$S'_t = (1 - \lambda_c) \frac{S_t}{S} + \lambda_c p\left(\frac{S_t}{S}\right), t \in \{00, 01, 10, 11\} \quad (4)$$

其中, λ_c 为权重因子, 平衡权重修正与局部集成的比例; $p(\cdot)$ 表示权重修正模块; S'_t 表示修正后的权重。

softmax 激活层将网络的输出转换为 $[0, 1]$ 范围内的分布, 使得输出与原本的 $\frac{S_t}{S}$ 具有相同的值域。经过权重修正后, 一张连续图像 $I^{(i)}$ 可以表示为

$$I^{(i)}(x_q) = \sum_t S'_t f_{\theta}(z_t, x_q - v_t), t \in \{00, 01, 10, 11\} \quad (5)$$

对于图像的大部分区域而言 (即图像较为平滑的区域), 局部集成的策略也能提供较为准确的预测, 因此权重修正模块的输出将与局部集成的结果加权求和。这个方法可以有效抑制在图像边缘、纹理较为复杂的区域内由于插值策略导致的权重分配失衡问题, 从而抑制图像的平滑性。

2.2 边缘特征扩展

权重修正的目标是避免预测在信号值时脱离图像全局的结构信息, 仅由欧式距离决定预测权重。本方法通过边缘特征扩展, 进一步解决预测信号值与图像结构信息不相关的问题。具体来说, 本方法将低分辨率图像的边缘特征加入由编码器提取的特征向量中, 如式 (6) 所示:

$$z_t^* = (z_t; \delta_t) \quad (6)$$

其中, z_t 表示低分辨率图像一点处由编码器 (即特征提取模块) 提取的特征向量; δ_t 表示该点低分辨率图像的边缘特征; z_t^* 表示扩展后的特征。

经过边缘特征扩展后, 一个连续图像 $I^{(i)}$ 可表示为

$$I^{(i)}(x_q) = \sum_t S'_t f_{\theta}(z_t^*, x_q - v_t), t \in \{00, 01, 10, 11\} \quad (7)$$

低分辨率图像包含非常丰富的结构信息,而边缘特征能使网络充分地应用图像的底层特征信息。引入边缘特征,可以缓解深度网络对底层特征信息描述较少的问题,从而指导恢复图像的整体结构,使恢复出的高分辨率图像有更锐利的边缘。

3 数值实验

本节将讨论实验数据集、指标和训练细节,同时将通过 WCESR 与其他方法的对比结果来讨论 WCESR 的模型性能。最后将通过消融实验进一步分析 WCESR。

3.1 数据集及评价指标

本实验通过 DIV2K 数据集^[19]训练和评估 WCESR。DIV2K 由 1 000 张 2K 分辨率的图像组成,其×2 缩放因子为 2、3 和 4 的低分辨率图像,是对数据集中高分辨率的图像通过双三次插值下采样得到的。与 LIIF 类似,在实验过程中使用数据集中的 800 张图像作为训练集,100 张图像作为验证集。本方法将在 Set5^[20]、Set14^[21]、B100^[22] 和 Urban100^[23] 这 4 个标准数据集上进行模型性能测试。

实验以峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 作为 WCESR 的评价指标。其计算如式 (8) 所示:

$$PSNR = 10 \log_{10} \left(\frac{1}{H \times W \cdot \|I_{SR} - I_{HR}\|_2^2} \right) \quad (8)$$

3.2 训练细节

为了训练 WCESR,首先从训练集的 HR 图像中随机裁切出大小为 $\{48r_i \times 48r_i\}$ 的块。其中, r_i 表示高分辨率图像块与低分辨率图像块之间的缩放因子,缩放因子是从 1 到 4 均匀采样得到,其对应的 LR 图像由这些 $\{48r_i \times 48r_i\}$ 大小的块通过双三次

插值下采样得到,每个 LR 图像的大小是 $\{48 \times 48\}$ 。将 LR 图像块输入到网络中,恢复出对应的高分辨率图像 SR,并反向传播真实图像 HR 和恢复图像 SR 的 L1 损失。实验选择 Adam 优化器,初始学习率为 $1 \cdot 10^{-4}$,权重因子 λ_c 设置为 0.005。实验训练共 1 000 个 epochs, batch size 设置为 16,学习率每 200 个 epochs 衰减至原先的一半。

3.3 实验结果分析

Meta-SR、LIIF 和 WCESR 在 DIV2K 上不同缩放因子下的性能比较结果见表 1。根据表中数据可知,WCESR 模型的 PSNR 值得到了与 Meta-SR、LIIF 相当甚至更好结果。WCESR 在 Set5、Set14、B100、Urban100 这 4 个标准数据集上的分析结果见表 2,由此可见 WCESR 在大多数实验中得到了优于 LIIF 的结果。在这 4 个数据集上,本实验仅测试了以 RDN 作为编码器的方法。通过表 2 可以看到,WCESR 在缩放因子较大,特别是超出训练范围时表现较好,这是因为在缩放因子较大时,低分辨率图像丢失了大量的边缘细节,而权重预测模块和边缘特征扩展可以帮助恢复的图像产生更锐利的边缘。Set5 和 Set14 两个数据集的图像结构与纹理较为简单,此时采用密集连接的 RDN 方法可以得到更好的结果,但由于 RDN 无法推广到更大的缩放因子,因此 RDN 的实际应用受到了限制。

图 4、图 5 给出了 WCESR 和 LIIF 的超分辨视觉效果对比。图 4 为缩放因子为 8 的重构结果,图 5 为缩放因子为 12 的重构结果。从图例可以看出,在面对较大的缩放因子时,LIIF 会产生图片过于平滑的情况,同时对于边缘处理出现了问题,而 WCESR 由于引入了权重修正模块,对于复杂边缘纹理有较好的处理,同时边缘特征也会指导恢复图像的整体结构。

表 1 Meta-SR、LIIF、WCESR 在 DIV-2k 数据集上的定量分析结果

Table 1 Quantitative analysis results of Meta-SR, LIIF, WCESR on the DIV-2k Dataset

dB

方法	×2	×3	×4	×6	×12	×18	×24	×30
Bicubic ^[5]	31.01	28.22	26.66	24.82	22.27	21.00	20.19	19.59
EDSR ^[5]	34.55	30.90	28.94	-	-	-	-	-
EDSR-MetaSR ^[6]	34.64	20.93	28.92	26.61	23.55	22.03	21.06	20.37
EDSR-LIIF ^[9]	34.67	30.96	29.00	26.75	23.71	22.17	21.18	20.48
WCESR-EDSR (本文)	34.69	30.98	29.03	26.79	23.74	22.20	21.20	20.50
RDN-MetaSR ^[6]	35.00	31.27	29.25	26.88	23.73	22.18	21.17	20.47
RDN-LIIF ^[9]	34.99	31.26	29.27	26.99	23.89	22.34	21.31	20.59
RDN-WCESR (本文)	35.00	31.26	29.29	27.00	23.89	22.36	21.32	20.61

表 2 Set5、Set14、B100、Urban100 数据集上的定量分析结果

Table 2 Quantitative Analysis Results on the Set5, Set14, B100, and Urban100 Datasets

数据集	方法	×2	×3	×4	×6	×8
Set5 ^[20]	RDN ^[16]	38.24	34.71	32.47	-	-
	MetaSR-RDN ^[6]	38.22	34.63	32.38	29.04	29.96
	LIIF-RDN ^[9]	38.17	34.68	32.50	29.15	27.14
	WCESR-RDN(本文)	38.18	34.70	32.51	29.21	27.19
Set14 ^[21]	RDN ^[16]	34.01	30.57	28.81	-	-
	MetaSR-RDN ^[6]	33.98	30.54	28.78	26.51	24.97
	LIIF-RDN ^[9]	33.97	30.53	28.80	26.64	25.15
	WCESR-RDN(本文)	33.96	30.53	28.86	26.70	25.20
B100 ^[22]	RDN ^[16]	32.24	29.26	27.72	-	-
	MetaSR-RDN ^[6]	32.33	29.26	27.71	25.90	24.83
	LIIF-RDN ^[9]	32.32	29.26	27.74	25.98	24.91
	WCESR-RDN(本文)	32.33	29.28	27.77	26.02	24.96
Urban100 ^[23]	RDN ^[16]	32.89	28.80	26.61	-	-
	MetaSR-RDN ^[6]	32.92	28.82	26.55	23.99	22.59
	LIIF-RDN ^[9]	32.87	28.82	26.68	24.20	22.79
	WCESR-RDN(本文)	32.92	28.88	26.72	24.28	22.83

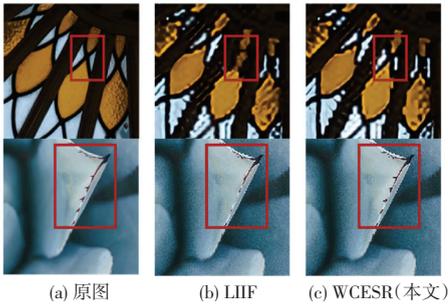


图 4 LIIF 与 WCESR 效果对比 (缩放因子: 8)

Fig. 4 Comparison of LIIF and WCESR effects (Scale: 8)

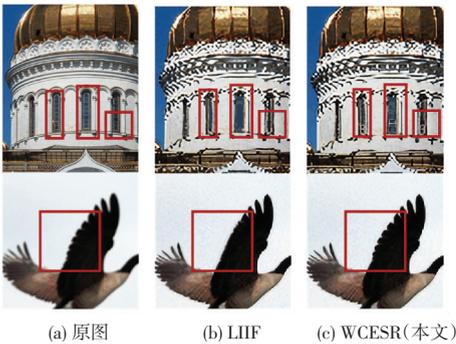


图 5 LIIF 与 WCESR 效果对比 (缩放因子: 12)

Fig. 5 Comparison of LIIF and WCESR effects (Scale: 12)

3.4 消融实验

为了表明本文所提网络架构的有效性,设计以下 4 组实验并对其进行定量分析和比较。

(1) Model_A: EDSR-baseline-LIIF^[9];

(2) Model_B: 表示在(1)的基础上引入权重修正模块;

(3) Model_C: 表示在(1)的基础上引入边缘特征扩展;

(4) Model_D: 表示在(2)的基础上引入边缘特征扩展,其结果见表 3; 实验还测试了修正的权重因子 λ_c 于结果的影响,见表 4。

表 4 的结果仅加入了权重修正模块,而没有引入边缘特征扩展。从表 4 可以看出,随着权重因子 λ_c 逐渐降低,网络表现越来越好,由于插值所带来的权重分配失衡问题逐渐被缓解。这是由于对于图像而言,绝大多数的区域都是较为平滑的区域,局部集成策略在这些区域可以得到很好的效果,然而对于图像的不平滑区域,引入权重修正可以得到更好的结果。

表 3 消融实验定量结果

Table 3 Quantitative results of ablation experiment

	权重修正	边缘特征扩展	× 2	× 3	× 4	× 8	× 12
Model_A	×	×	34.67	30.96	29.00	26.75	23.71
Model_B	√	×	34.68	30.97	29.02	26.77	23.73
Model_C	×	√	34.67	30.96	29.01	26.76	23.71
Model_D	√	√	34.69	30.98	29.03	26.79	23.74

表4 权重因子 λ_c 取值的定量结果

Table 4 Quantitative results of weighting factor values

λ_c	$\times 2$	$\times 3$	$\times 4$	$\times 6$	$\times 12$	$\times 18$	$\times 24$	$\times 30$
0.01	34.676 4	30.965 2	29.008 8	26.753 4	23.714 4	22.175 5	21.181 8	20.481 2
0.05	34.673 4	30.969 0	29.013 4	26.755 0	23.718 0	22.175 0	21.189 0	20.487 2
0.005	34.683 1	30.972 5	29.020 2	26.769 5	23.733 9	22.189 8	21.191 1	20.494 3

4 结束语

本文提出了一种基于隐式神经表示的图像超分辨率重构方法,方法运用隐式神经表示,将二维坐标映射为对应的信号值,实现任意分辨率的超分辨率重构。本方法引入了权重修正网络,通过学习局部面积权重与全局结构权重的关系,修正双线性插值策略带来的权重分配失衡问题。同时引入 LR 图像的边缘信息,指导恢复图像的整体结构,约束恢复图像的平滑性。经过实验证明,在相同设置下与 LIIF 相比 WCESR 取得了更优越的性能。

实验发现:

(1) 基于隐式神经表示学习图像连续表示的方法是通过二维离散图像特征图学习图像局部隐式函数,并通过像素级采样输出新的离散图像,其连续性体现在二维空间坐标采样的连续性,这种策略忽视了原始离散图像与恢复离散图像的耦合关系;

(2) 由图像的连续表示到二维离散图像的过程往往采用像素级采样,这种策略也导致了恢复图像的过平滑现象。因此,如何建立原始离散图像与恢复离散图像的耦合关系,如何利用图像纹理特征的复杂性决定采样率有待后续的进一步研究。

参考文献

- [1] CHANG H, YEUNG D Y, XIONG Y. Super-resolution through neighbor embedding. [C]// Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2004: 275-282.
- [2] FREEDMAN G, FATTAL R. Image and video upscaling from local self-examples[J]. ACM Transactions on Graphics (TOG), 2011, 30(2): 1-11.
- [3] YOON Y G, JEON H G, YOO D, et al. Learning a deep convolutional network for light-field image super-resolution [C]// Proceedings of IEEE International Conference on Computer Vision Workshop. IEEE, 2015: 57-65.
- [4] DONG C, LOY C C, HE K M, et al. Learning a deep convolutional network for image super-resolution [J]. IEEE Trans Pattern Anal Mach Intell, 2016, 38(2): 295-307.
- [5] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE, 2017: 1132-1140.
- [6] HU X, MU H, ZHANG X, et al. Meta-SR: A Magnification-Arbitrary Network for Super-Resolution [C]// Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020: 1575-1584.
- [7] DUPONT E, TEH Y W, DOUCET A. Generative models as distributions of functions[J]. arXiv preprint arXiv: 2012.04776, 2021.
- [8] SKOROKHODOV I, IGNATYEV S, ELHOSEINY M. Adversarial Generation of continuous images [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 10748-10759.
- [9] CHEN Y, LIU S, WANG X. Learning continuous image representation with local implicit image function [C]// Proceedings of 2021 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2021: 8624-8634.
- [10] FATTAL R. Image upsampling via imposed edge statistics [J]. ACM Transactions on Graphics (TOG), 2007, 26(3): 95.
- [11] JIAN S, XU Z, SHUM H Y. Image super-resolution using gradient profile prior [C]// Proceedings of 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008). IEEE, 2008: 24-26.
- [12] ZONTAK M, IRANI M. Internal statistics of a single natural image [C]// Computer Vision and Pattern Recognition. Colorado: IEEE, 2011: 977-984.
- [13] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 1646-1654.
- [14] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [J]. Computer Vision and Pattern Recognition, arXiv:1609.04802, 2016.
- [15] YU J, FAN Y, YANG J, et al. Wide activation for efficient and accurate image super-resolution [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway, NJ: IEEE, 2018: 2621-2624.
- [16] ZHANG Y, TIAN Y, KONG Y, et al. Residual dense network for image super-resolution [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018: 2472-2481.
- [17] SITZMANN V, ZOLLHOFER M, WETZSTEIN G. Scene representation networks: continuous 3D-structure-aware neural scene representations [C]// Advances in Neural Information Processing Systems, 2019: 1121-1132.
- [18] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines vinod nair [C]// International Conference on International Conference on Machine Learning. Omnipress, 2010: 807-814.
- [19] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study [C]// IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2017: 1122-1131.