



基于多窗口残差网络的单图像超分辨率重建

肖雅敏, 张家晨, 冯 铁

(吉林大学 计算机科学与技术学院, 长春 130012)

摘 要: 基于卷积神经网络的单图像超分辨率模型网络结构过深, 导致高频信息丢失以及模型体积庞大等问题。提出一种由多个残差模块构成的多窗口残差网络优化模型, 通过使用多个不同尺寸的窗口对同一特征图进行提取, 获取更丰富的高频与低频信息, 并过滤出深层网络的所需特征。残差模块中较大尺寸的窗口采用较小尺寸的滤波器和多层映射层叠加组成, 可在减少参数总量的同时增强网络的非线性表达能力。实验结果表明, 与A+、SRCNN、ESPCN等模型相比, 该模型可有效利用原始图像信息生成细节更清晰的超分辨率图像, 且在主观视觉效果与客观评价指标上均有所提升。

关键词: 单图像超分辨率重建; 多窗口残差网络; 卷积神经网络; 深度学习; 特征融合

开放科学(资源服务)标志码(OSID):



中文引用格式: 肖雅敏, 张家晨, 冯铁. 基于多窗口残差网络的单图像超分辨率重建[J]. 计算机工程, 2021, 47(2): 293-299, 306.

英文引用格式: XIAO Yamin, ZHANG Jiachen, FENG Tie. Single image super-resolution reconstruction based on multi-windows residual network[J]. Computer Engineering, 2021, 47(2): 293-299, 306.

Single Image Super-Resolution Reconstruction Based on Multi-Windows Residual Network

XIAO Yamin, ZHANG Jiachen, FENG Tie

(College of Computer Science and Technology, Jilin University, Changchun 130012, China)

[Abstract] The single image super-resolution model based on Convolutional Neural Network (CNN) is faced with problems including the loss of high-frequency information and the large model size caused by the depth of deep network structure. To address the problems, this paper proposes an optimization model of multi-windows residual network composed of multiple residual modules. By using multi-windows with different sizes to extract the same feature map, more abundant high-frequency and low-frequency information can be obtained, and the features required by deep network are filtered out. The larger window in the residual module is composed of the superposed smaller filter and multiple mapping layers, which can reduce the total number of parameters and enhance the nonlinear expression ability of the network. Experimental results show that compared with A+, SRCNN, ESPCN and other models, this model can effectively use the original image information to generate super-resolution images with clear details, and improve the subjective visual effect and objective evaluation index.

[Key words] single image super-resolution reconstruction; multi-windows residual network; Convolutional Neural Network (CNN); deep learning; feature integration

DOI: 10.19678/j.issn.1000-3428.0057224

0 概述

单图像超分辨率重建是计算机视觉领域中重要的图像处理技术, 广泛应用于医学图像、视频监控、

公共安全与遥感侦察等领域。在实际应用场景中, 现有硬件条件的限制造成图像的分辨率质量较低, 如以数字摄像机为监控设备时, 其获取的图像通常缺少关键场景或人物细节信息。因此, 为克服现有

基金项目: 国家重点研发计划(2018YFC1315604); 国家自然科学基金面上项目(61872164); 赛尔网络下一代互联网技术创新项目(NGII20180701)。

作者简介: 肖雅敏(1994—), 女, 硕士研究生, 主研方向为计算机视觉、机器学习; 张家晨, 教授、博士; 冯 铁, 副教授、博士。

收稿日期: 2020-01-15 **修回日期:** 2020-02-17 **E-mail:** xymfree@icloud.com

硬件系统的分辨率限制,使用单图像超分辨率重建技术来增强图像的空间分辨率。该技术的核心思想是通过分析低分辨率图像的关键语义信息或信号信息,推理出欠缺的真实细节并重建成具有高像素密度的超分辨率图像。

目前,单图像超分辨率重建的研究主要分为3个阶段。基于采样理论的插值方法^[1-2]是出现较早且直观的方法,该类方法的优点是运行速度快且适合并行计算,但是存在不能引入额外有用的高频信息,从而难以得到锐化的高清图像的问题。基于此,研究人员提出利用低分辨率图像推测出对应的高分辨率部分信息的映射函数算法,该算法主要依赖邻域嵌入^[3-4]、稀疏编码^[5-7]等技术,但是当图像中未包含较多的重复样式时,其推断出的图像通常会产生非细节性的锐利边缘。

基于深度学习的方法已成为超分辨率算法中的热点方向。文献[8]提出的SRCNN模型采用卷积神经网络(Convolutional Neural Network, CNN)技术重建生成了清晰度更高的图像,它采用大量外部的高分辨率图像构造学习库,并经过训练生成神经网络模型,在对低分辨率图像重建过程中引入模型获取的先验知识,以此得到图像的高频细节信息,从而达到良好的图像重建效果。FSRCNN^[9]、ESPCN^[10]以及文献[11-13]模型均是在SRCNN的基础上对网络结构进行改进而得到的,这些模型增加了网络层数,并专注于学习低分辨率图像到高分辨率图像的端到端映射关系。由于随着构建网络层数的加深,训练模型的花销逐渐增大,同时由于通道数、过滤器尺寸与步长等超参数的增加,导致设计合理的网络结构更加困难,因此,文献[14]提出利用ResNet模型解决上述问题,虽然该模型适用于图像分类,但其残差思想和重复堆叠模块的策略能够适用于所有计算机视觉任务。此外,ResNet还证明了跳跃连接和递归卷积可有效缓解神经网络携带大量关键信息的负担。

基于残差网络的超分辨率重建模型DCRN^[15]、DRNN^[16]、LapSRN^[17]、SRResNet^[18]与EDSR^[19]等被相继提出,这些模型都是通过单尺寸卷积模块的线性叠加来实现网络的纵向加深,以追求更高的表达能力和抽象能力。但对于超分辨率技术而言,从原始图像中提取到丰富完整的特征信息至关重要。如果继续纵向加深网络,在逐层卷积和过滤计算过程中将会出现高频信息丢失的问题,这对最终映射生成超分辨率图像的真实程度造成影响,且模型参数量

也会呈指数级增长。如果训练数据集有限则易产生过拟合,且模型规格与计算量也随之增大,造成不易重构和移植以及训练难度成倍增加的问题,从而难以在实际过程中应用。

本文提出一种基于多窗口残差网络的单图像超分辨率重建模型MWSR。该模型的横向网络结构使用多个不同尺寸的窗口,同时提取同一特征图的关键信息,纵向网络结构将大尺寸窗口重构为多个串联的小窗口和非线性ReLU激活函数,并对输出的多尺寸特征图进行跨窗口融合,以优化通过网络的信息流。

1 MWSR实现方法

本文提出的基于多窗口残差网络的单图像超分辨率MWSR模型的目标是学习从低分辨率图像 I^{LR} 到真实的高分辨率图像 I^{HR} 的端到端映射,并生成一张清晰的超分辨率图像 I^{SR} 。其中, I^{LR} 是 I^{HR} 经过双三次插值方法下采样得到的低分辨率图像,且大小为 $H \times W \times C$ 。此外, I^{HR} 和 I^{SR} 大小为 $rH \times rW \times C$ 。其中, H 表示纵向像素数量, W 表示横向像素数量, r 为目标放大倍数, C 为颜色通道。本文实验采用RGB格式的图像,因此 $C=3$ 。

1.1 MWSR网络架构

MWSR的网络架构如图1所示,该架构主要分为特征提取、全局特征融合以及高倍数重建3个部分,具体步骤如下:

步骤1 将 I^{LR} 作为神经网络的输入,先经过一个 3×3 的卷积层提取初始特征,再使用激活函数ReLU调整神经元的活跃度,以增强网络的非线性,具体的特征提取函数可表示为:

$$Q_n^l = P(I^{\text{LR}}) = \sigma(w_{n \times n}^l \otimes I^{\text{LR}} + b^l) \quad (1)$$

其中,“ $\otimes$ ”表示卷积操作,“+”表示逐像素相加操作,变量的上标 l 指它所在的网络层,下标 n 表示卷积核的大小,变量 w 和变量 b 分别表示神经网络中的权重矩阵和偏倚矩阵, w 是一个大小为 $f^{l-1} \times f^l \times n \times n$ 的张量。 f^l 是第 l 层特征图的数量($f^0 = C$), $\sigma(x) = \max(0, x)$ 表示非线性激活函数ReLU。 Q_n^l 和 $P(x)$ 均表示第 l 层卷积核大小为 n 的输出。

步骤2 通过 m 个重复连续的多窗口残差模块逐层提取图像的高频与低频信息,并对初始特征图和上述 m 个残差模块的输出进行全局特征融合,对通过网络的信息流进行汇合并输入到重建模块中,具体表示方法为:

$$T = [Q_3^1, M_1, M_2, \dots, M_m] \quad (2)$$

其中, $[\]$ 表示特征之间的连接操作, Q_3^1 为初始特征图, M_m 表示第 m 个多窗口残差模块的输出, T 表示全局特征融合后的输出。

步骤3 在高倍数重建部分, 首先使用 3×3 卷积整理过滤冗余信息, 重构出最优的稀疏网络结构, 然后使用亚像素卷积操作^[10]将特征图

T 上采样至目标倍数 r 。最后, 通过一层 3×3 卷积完成 I^{LR} 到 I^{SR} 的映射, 生成清晰的超分辨率图像, 具体表示方法为:

$$I^{SR} = \sigma(w_{3 \times 3}^l \times \text{SF}(\sigma(w_{3 \times 3}^{l-2} \times T + b^{l-2})) + b^l) \quad (3)$$

其中, $\text{SF}(x)$ 表示重新排列组合像素的亚像素卷积操作, 变量上标中的 l 表示网络中的最后一个卷积层, $l-2$ 表示重建部分的第一个卷积层。

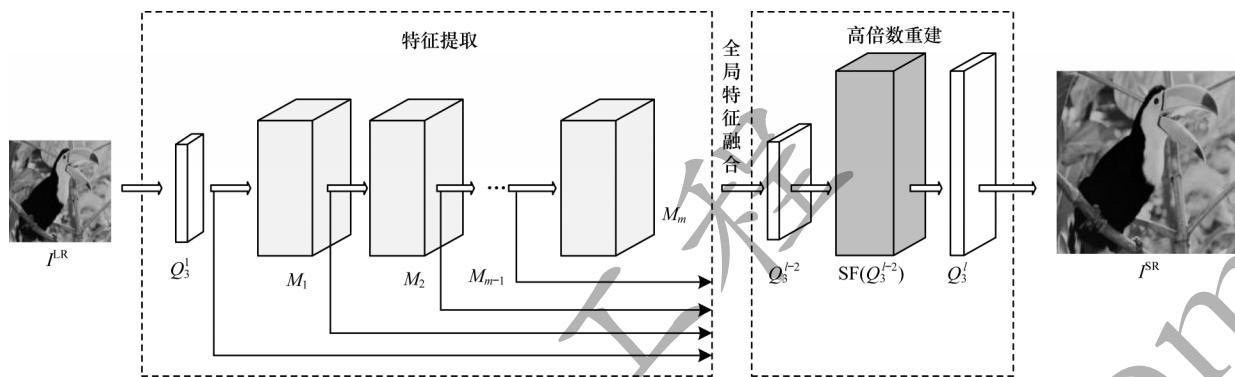


图1 MWSR模型网络架构

Fig.1 Network architecture of the MWSR model

1.2 多窗口残差模块

将 I^{LR} 输入到网络中后, 每经过一层纵向卷积操作提取特征信息的同时也会丢失相关的高频信息, 进而造成完整的原始图像语义会随着网络层次的加深而逐渐缺失。因此, 本文从浅层网络开始, 当每次特征提取时尽可能提取出更丰富的原始图像特征。本文提出一种横向拓展网络结构的多窗口残差模块, 如图2所示。同时使用 3×3 、 5×5 、 7×7 三种尺寸的窗口作为同一特征域的局部关键信息检测器, 大幅提高检测到有效特征信息的概率。

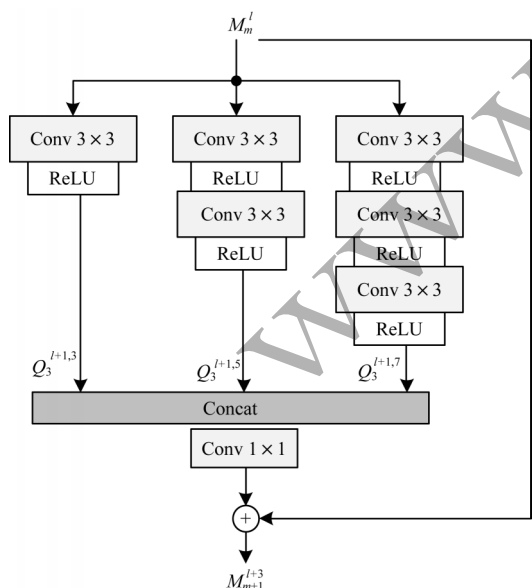


图2 MWSR模型的残差模块

Fig.2 Residual module of MWSR model

与单尺寸卷积模块堆叠而成的模型相比, 本文所提 MWSR 模型使用更少的特征提取模块即可提取更全面的特征信息, 从而避免了网络结构过深, 同时训练过程中可以更早达到饱和的准确度。

为进一步优化模型规模大小, 在保持模型的准确度和复杂度的前提下, 本文使用 k 个串联的 3×3 小卷积核构成相应的大尺寸窗口。如图3所示, 2个 3×3 的卷积核串联相当于一个 5×5 的卷积核, 即一个像素会与周围 5×5 个像素产生关联。两者感受野大小相同, 而前者的参数量为 $2 \times 3 \times 3 = 18$, 后者的参数量为 $5 \times 5 = 25$, 这说明前者的参数量更少。

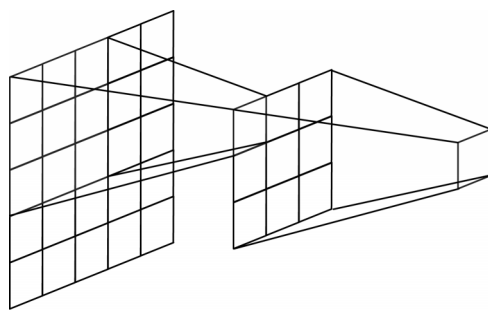


图3 两个串联的 3×3 卷积核感受野与一个 5×5 卷积核感受野示意图

Fig.3 Schematic diagram of receptive field of two 3×3 filters in series and a 5×5 filter

除此之外, 还可以在2个小卷积核之间增加一个非线性激活函数 ReLU, 使得整个网络变得更复杂, 拥有更强的学习能力。以此类推, 3个 3×3 的卷

积层串联相当于1个7×7的卷积层,且还可以增加2个ReLU函数,具体表述方法为:

$$Q_n^{l,s} = \underbrace{P_n^{l,s,k} (P_n^{l,s,k-1} P_n^{l,s,k-2} \dots (P_n^{l,s,1}(M)))}_{k\text{个连续的特征提取操作}} \quad (4)$$

其中, $P_n^{l,s,k}(x)$ 表示特征提取函数,参数 l, s, k 依次表示第 l 层的 $s \times s$ 大小的窗口中第 k 个特征提取结果, n 表示卷积核大小, M 表示该窗口的输入, $Q_n^{l,s}$ 表示该窗口的最终输出。

在获取输出的三种尺寸特征图后,将其与矩阵的第1维度连接融合,然后通过1×1滤波器逐维度过滤筛选出深层网络需要的特征信息,同时维持该模型输入输出维度的一致性以平衡模块规模大小。

为缓解随着网络加深造成准确度降低的问题,本文在特征提取模块中应用 ResNet^[11] 模型的残差思想:

$$H(x) = F(x) + x \quad (5)$$

将浅层网络的输出 x 直接传递给深层网络 $H(x)$ 作为输入,并在该过程中增加跳跃连接,使得无需将 x 映射成一个新的 $H(x)$,只需要学习 x 和 $H(x)$ 的差距 $F(x)$,显著减小需要学习的参数值,使得网络对反向传播时的损失值更加敏感,起到一定的正则化作用。

多窗口残差模块的具体操作表示方法为:

$$M_{m+1}^{l+3} = w_{1 \times 1}^{l+2} \times [Q_3^{l+1,3}, Q_3^{l+1,5}, Q_3^{l+1,7}] + b^{l+2} + M_m^l \quad (6)$$

其中, $[Q_3^{l+1,3}, Q_3^{l+1,5}, Q_3^{l+1,7}]$ 表示3种尺寸窗口的输出结果连接操作,变量 $w_{1 \times 1}^{l+2}$ 和 b^{l+2} 分别表示模块中使用1×1权重张量以及同一层的偏倚张量,变量 M_m^l 表示第 m 个残差模块的输出结果。

实验结果表明,本文提出的多窗口残差模块简洁且高效。当设置MWSR的残差模块数量为24时,重建成的超分辨率图像质量已经超越了当前流行的多数模型,此时MWSR的参数量仅为5.9 M, FLOPS低至13.64×10⁹次的浮点计算。该轻量网络模型可以在几乎所有服务端内部署,包括移动端和嵌入式芯片等对模型大小有严格限制的服务端。此外,经实验对比可得,3×3、5×5、7×7三种尺寸的窗口组合为最优解,模型性能优于3×3、5×5组合。如果再增加一个9×9的窗口,则参数量会高达50.91M,计算量更会提高到117.37×10⁹次的浮点计算。

1.3 亚像素卷积上采样层

在相机成像过程中,由于硬件方面的限制,生成的图像上每个像素都代表附近的一整块颜色,其实在微

观上,实际物理像素之间还存在许多像素,即亚像素。在超分辨率领域中,无法被传感器检测出来的亚像素可以通过算法近似计算出来,相当于推理出图像缺失的纹理细节等高频信息。亚像素卷积在MWSR模型的高倍数重建部分用于完成低分辨率图像到高分辨率图像的映射,具体上采样过程如图4所示。假设目标倍数为 r , 输入的低分辨率特征图大小为 $H \times W$, 将其与通道数为 r^2 的 $H \times W$ 亚像素卷积核进行卷积,得到 $H \times W \times r^2$ 个像素值,再将其重新排列组合成大小为 $rH \times rW$ 的目标图像。

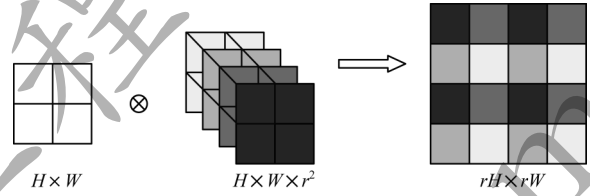


图4 亚像素卷积层上采样过程

Fig.4 Upsampling process on sub-pixel convolution layer

2 实验与结果分析

2.1 数据集

DIV2K数据集^[20]是广泛运用在超分辨率领域的分辨率为2K的高质量数据集,其包含800张训练图像、100张验证图像以及100张测试图像。本文选择DIV2K作为模型的训练数据集,并在Set5^[21]、Set14^[22]、BSDS100^[23]、Urban100^[24]和Manga109^[25]5个公开的基准数据集上进行测试与对比实验。这些数据集包含大量来自各个领域的图像,能够对模型性能进行有效验证。其中,Set5、Set14与BSDS100数据集主要由自然景观图像组成,Urban100数据集中包括不同频带细节的城市场景图像,Manga109为日本漫画数据集。

2.2 实验参数设置

本文使用从 I^{LR} 裁剪出分辨率为48×48的RGB图像作为输入,并通过目标放大倍数的 I^{HR} 评估所生成超分辨率图像的质量。为了提高模型的泛化能力和鲁棒性,通过随机使用水平翻转180°、垂直翻转180°与顺时针旋转90°3种方法操作将训练数据量增加至原来的8倍,且每种方法执行概率均为0.5。另外,使用Adam优化器^[24],分别设置 $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$ 。本文设置每一次迭代的批大小为64,每迭代2500次验证一次模型。学习速率初始化为0.0002,且每迭代500000次将其减少至原来的一半。

本文选择 L_1 范数作为损失函数训练模型, 与 L_2 相比其具有的稀疏性可以实现特征的自动化选择, 且参数量更少。同时, 本文结合通用并行计算架构 CUDA10.0 与深度学习框架 PyTorch1.0, 使用 Python 编码实现 MWSR 的算法, 并在 NVIDIA GeForce RTX 2080Ti GPU 和 Ubuntu16.04 操作系统上, 通过大量实验对算法进行训练和评估。

2.3 评价指标

本文采用 2 个常见的客观评价指标来评估生成的超分辨率图像的质量, 该指标分别为峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 和结构相似度 (Structural SIMilarity, SSIM)。

1) PSNR 主要基于误差敏感来评估图像质量, 计算方法如式(7)所示, 单位为 dB, 且其数值越大, 则说明图像失真程度越低。

$$\text{PSNR} = 10 \lg \left(\frac{(2^t - 1)^2}{\frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2} \right) \quad (7)$$

其中, $H \times W$ 为图像的大小, X 表示真实的高分辨率图像, Y 表示模型生成的超分辨率图像, 变量 t 表示每个像素占用的比特数。

2) SSIM 通过比较图像结构之间的差异来评估图像质量, 其结果更符合人眼视觉系统, 计算方法如

式(8)所示, 且其值越接近 1, 则说明两个图像之间的结构越相似, 重建质量越高。

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (8)$$

其中, 变量 x 和 y 分别为真实的高分辨率图像和模型生成的超分辨率图像, μ_x, μ_y 表示图像的灰度平均值, σ_x, σ_y 表示图像的方差, σ_{xy} 表示图像的协方差。 C_1 和 C_2 是避免公式分母为 0 的常数。

2.4 算法对比分析

实验对 MWSR 模型与目前主流的双三次插值 (Bicubic)、A+^[6]、SRCNN^[8]、ESPCN^[10]、DRNN^[13]、LapSRN^[14] 6 种超分辨率模型进行对比分析。与其他 6 种模型一致, MWSR 模型生成的所有超分辨率图像转换成 YCbCr 色彩模式, 仅在 Y 通道上计算 PSNR 和 SSIM。

表 1 展示了 7 种超分辨率模型在 5 个公开测试数据集上分别放大 2 倍~4 倍的评估结果。其中, 最优结果加粗表示。从表 1 可以看出, 本文提出的 MWSR 模型的性能均优于其他 6 种主流模型, 且该模型不仅改善了图像感知质量, 而且实现了模型轻量化及运行效率的优化。

表 1 7 种模型的 PSNR 和 SSIM 评估结果对比

Table 1 Comparison of PSNR and SSIM evaluation results of seven models

数据集	放大倍数	Bicubic 模型		A+模型		SRCNN 模型		ESPCN 模型		DRRN 模型		LapSRN 模型		MWSR 模型	
		PSNR/ dB	SSIM	PSNR/ dB	SSIM	PSNR/ dB	SSIM	PSNR/ dB	SSIM	PSNR/ dB	SSIM	PSNR/ dB	SSIM	PSNR/ dB	SSIM
Set5	2	33.66	0.929 9	36.54	0.954 4	36.66	0.954 2	37.00	0.955 9	37.74	0.959 1	37.52	0.958 1	38.23	0.961 6
	3	30.39	0.868 2	32.58	0.908 8	32.75	0.909 0	33.02	0.913 5	34.03	0.924 4	33.82	0.920 7	34.65	0.929 1
	4	28.42	0.810 4	30.28	0.860 3	30.48	0.862 8	30.66	0.864 6	31.68	0.888 8	31.54	0.881 1	32.41	0.898 2
Set14	2	30.24	0.868 8	32.28	0.905 6	32.42	0.906 3	32.75	0.909 8	33.23	0.913 6	33.08	0.910 9	33.96	0.920 1
	3	27.55	0.774 2	29.13	0.818 8	29.28	0.820 9	29.49	0.827 1	29.96	0.834 9	29.89	0.830 4	30.54	0.845 6
	4	26.00	0.702 7	27.32	0.749 1	27.49	0.750 3	27.71	0.756 2	28.21	0.772 0	28.19	0.763 5	28.73	0.785 2
BSDS100	2	29.56	0.843 1	31.21	0.886 3	31.36	0.887 9	31.51	0.893 9	32.05	0.897 3	31.80	0.894 9	32.30	0.901 8
	3	27.21	0.738 5	28.29	0.783 5	28.41	0.786 3	28.50	0.793 7	28.95	0.800 4	28.82	0.795 0	29.18	0.807 5
	4	25.96	0.667 5	26.82	0.708 7	26.90	0.710 1	26.98	0.712 4	27.38	0.728 4	27.32	0.716 2	27.66	0.739 6
Urban100	2	26.88	0.840 3	29.20	0.893 8	29.50	0.894 6	29.87	0.906 5	31.23	0.918 8	30.41	0.911 2	32.69	0.933 2
	3	24.46	0.734 9	26.03	0.797 3	26.24	0.798 9	26.41	0.816 1	27.53	0.837 8	27.07	0.829 8	28.62	0.857 7
	4	23.14	0.657 7	24.32	0.718 3	24.52	0.722 1	24.60	0.736 0	25.44	0.763 8	25.21	0.756 4	26.46	0.797 3
Manga109	2	30.82	0.933 2	35.37	0.966 3	35.74	0.966 1	36.21	0.969 4	—	—	37.27	0.985 5	38.99	0.997 6
	3	26.96	0.855 5	29.93	0.908 9	30.59	0.910 7	30.79	0.918 1	—	—	32.21	0.931 8	33.96	0.949 7
	4	24.91	0.782 6	27.03	0.843 9	27.66	0.850 5	27.70	0.856 0	—	—	29.09	0.884 5	30.92	0.918 5

表1展示的MWSR模型仅使用了24个多窗口残差模块,经实验表明,当MWSR中的模块数量从24增加至32或64时,训练效果仍继续提高。如图5所示,MWSR模型在DIV2K数据集上评估的PSNR值可高达37.390 dB,且即使有64个残差模块,参数量也仅为15.4M。

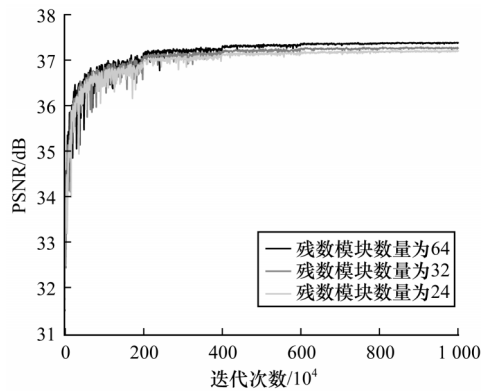


图5 残差模块数量分别为24、32、64的MWSR在DIV2K数据集上训练的PSNR收敛曲线

Fig.5 PSNR convergence curves of MWSR with 24, 32, 64 residual modules trained on DIV2K dataset

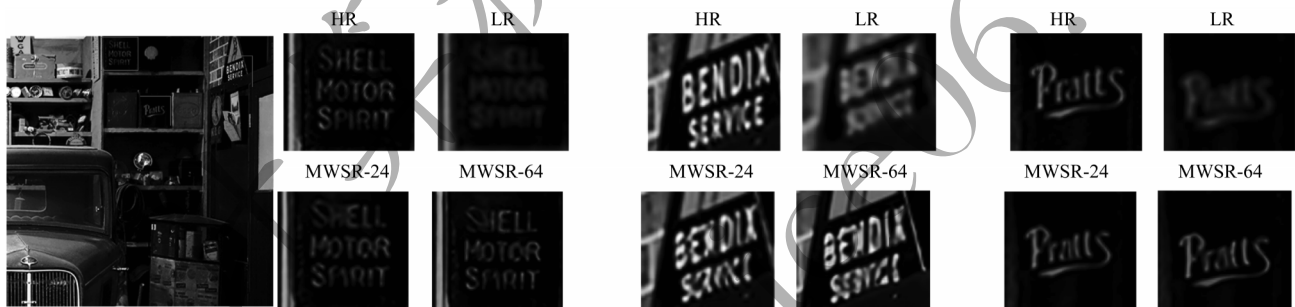


图6 MWSR-24、MWSR-64在DIV2K数据集上的超分辨率重建结果

Fig.6 Super-resolution reconstruction results of MWSR-24 and MWSR-64 on DIV2K dataset

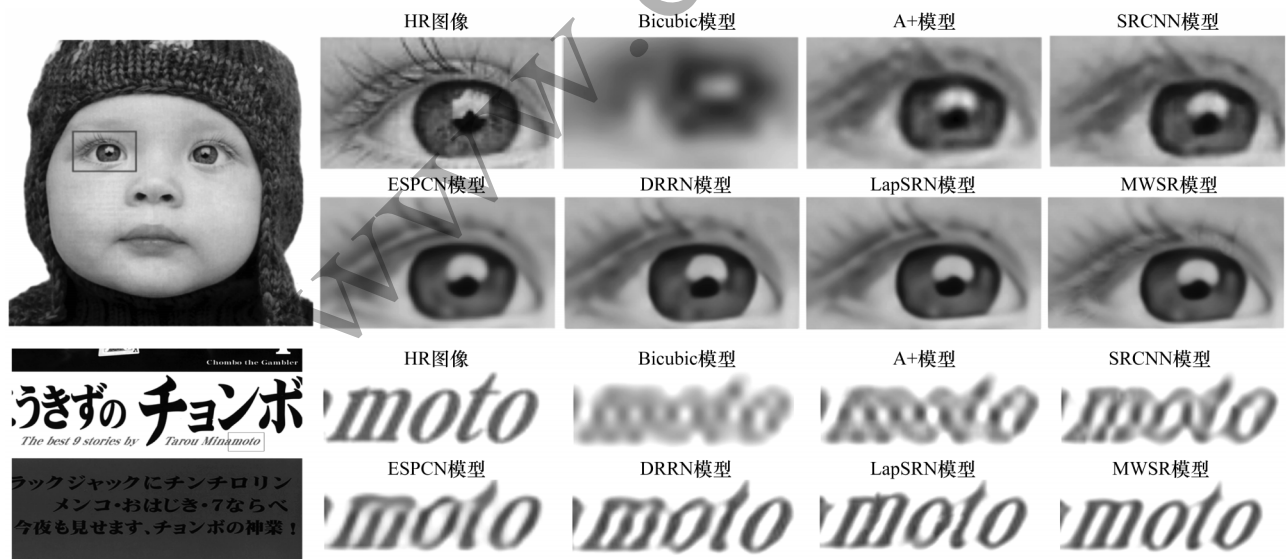


图7 7种模型的超分辨率重建结果对比

Fig.7 Comparison of super-resolution reconstruction results of seven models

从图6中可观察到MWSR模型中的残差模块数量分别为24(MWSR-24)和64(MWSR-64)时的重建图像,证明在合理的范围内增加模块数量可有效提升图像质量。然而,增加模块数量的同时也成倍增加了运算次数和内存读写的时间开销,如MWSR-24的训练时间为1.5天,而MWSR-64的训练时间为3.5天。另外在Set5数据集上测试可得,MWSR-24的平均运行时间为0.23 s,MWSR-64的平均运行时间为0.38 s。如果再继续增加残差模块以加深网络,仍会改善重建效果,但会造成训练成本和运行成本的快速增加。图7展示了各模型在不同数据集上进行4倍超分辨率重建的视觉效果对比。从图7可以看出,相比其他模型,MWSR模型重建出的图像更加准确且完整地补全了高频信息。无论是线形细节,结构细节还是文字细节,都可以根据图像的整体语义预测出其放大之后更真实的新像素值。

综合对比分析可知,在实际应用场景中,先根据需求对训练难度、模型大小、运行效率和重建质量等多种因素进行综合考虑,再对特征提取模块的数量进行选择,利用合适的网络深度使多窗口结构发挥更大的作用。

3 结束语

本文提出一种多尺寸窗口残差网络优化模型MWSR。该模型结合不同尺寸的窗口同时提取同一张特征图的关键信息,从浅层网络开始高效利用每一层的特征图,以提高高频信息的检测概率。这种横向扩展网络结构的方式与仅纵向加深网络结构相比,能够更快速地获取完整的目标特征。此外,针对较大尺寸窗口进行的分解及重构使得模型更易重构和移植。实验结果表明,与目前主流的图像超分辨率重建模型相比,本文模型能够有效提升生成超分辨率图像的清晰度。下一步将通过引入注意力机制,优化高倍数重建部分的上采样操作过程,使生成的图像更加逼真和自然。

参考文献

- [1] LI X, ORCHARD M T. New edge-directed interpolation[J]. IEEE Transactions on Image Processing, 2001, 10(10): 1521-1527.
- [2] ZHANG Lei, WU Xiaolin. An edge-guided image interpolation algorithm via directional filtering and data fusion [J]. IEEE Transactions on Image Processing, 2006, 15(8): 2226-2238.
- [3] ROWEIS S T. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.
- [4] GAO Xinbo, ZHANG Kaibing, TAO Dacheng, et al. Image super-resolution with sparse neighbor embedding[J]. IEEE Transactions on Image Processing, 2012, 21(7): 3194-3205.
- [5] YANG J C, WANG Z W, LIN Z, et al. Coupled dictionary training for image super-resolution[J]. IEEE Transactions on Image Processing, 2012, 21(8): 3467-3478.
- [6] TIMOFTE R, DE SMET V, VAN GOOL L. A+: adjusted anchored neighborhood regression for fast super-resolution[M]. Berlin, Germany: Springer, 2015: 111-126.
- [7] YANG J C, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19(11): 2861-2873.
- [8] DONG C, LOY C C, HE K M, et al. Learning a deep convolutional network for image super-resolution[M]. Berlin, Germany: Springer, 2014: 184-199.
- [9] DONG C, LOY C C, TANG X O. Accelerating the super-resolution convolutional neural network[M]. Berlin, Germany: Springer, 2016: 391-407.
- [10] SHI W Z, CABALLERO J, HUSZAR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 1874-1883.
- [11] LIAN Yiya, WU Xiaojun. Research on image super-resolution reconstruction of super deep convolutional neural network[J]. Computer Engineering, 2019, 45(1): 217-220. (in Chinese)
连逸亚, 吴小俊. 超深卷积神经网络的图像超分辨率重建研究[J]. 计算机工程, 2019, 45(1): 217-220.
- [12] HUANG Wei, FENG Jingjing, HUANG Yao. Super-resolution algorithm for images based on multi-channel extremely deep convolutional neural network[J]. Computer Engineering, 2020, 46(9): 242-247, 253. (in Chinese)
黄伟, 冯晶晶, 黄遥. 基于多通道极深卷积神经网络的图像超分辨率算法[J]. 计算机工程, 2020, 46(9): 242-247, 253.
- [13] HU Xiaohui, ZHANG Jianguo. Research on image super-resolution algorithm based on improved convolutional neural network [J]. Application Research of Computers, 2020, 37(3): 947-950, 956. (in Chinese)
胡晓辉, 张建国. 基于改进卷积神经网络的图像超分辨率算法研究[J]. 计算机应用研究, 2020, 37(3): 947-950, 956.
- [14] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition [C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 770-778.
- [15] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution [C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2016: 1637-1645.
- [16] TAI Ying, YANG Jian, LIU Xiaoming. Image super-resolution via deep recursive residual network [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 2790-2798.
- [17] LAI W S, HUANG J B, AHUJA N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 5835-5843.
- [18] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2017: 105-114.
- [19] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Washington D. C., USA: IEEE Press, 2017: 136-144.

(上接第 299 页)

- [20] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study [C]// Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Washington D. C. , USA; IEEE Press, 2017; 1122-1131.
- [21] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[EB/OL]. [2019-12-13]. <https://www.ixueshu.com/document/7e12bde83a725a9f318947a18e7f9386.html>.
- [22] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[M]. Berlin, Germany: Springer, 2012; 711-730.
- [23] 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.
- [24] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C. , USA; IEEE Press, 2015; 5197-5206.
- [25] MATSUI Y, ITO K, ARAMAKI Y, et al. Sketch-based manga retrieval using Manga109 dataset[J]. Multimedia Tools and Applications, 2017, 76(20): 21811-21838.

编辑 刘继娟