

## 增强型多尺度残差网络的图像超分辨率重建算法

许娇\*, 袁三男

上海电力大学电子与信息工程学院, 上海 200120

**摘要** 现有的图像超分辨率重建算法大多具有极深的网络结构导致参数量过大,并且不能充分提取特征。为了解决以上问题,提出了一种基于增强型多尺度残差网络(EMSRN)的图像超分辨率重建算法。该网络主要由多个增强型多尺度残差块(EMSRB)组成,通过使用残差块和并行的多空洞率的空洞卷积组构建该模块的骨干结构,获取了图像的局部和全局多尺度特征的同时有效减小了网络参数量。在模块最后使用通道注意力机制自适应地对提取到的特征进行加权,使网络更多地关注高频信息。实验结果表明,比起基础的多尺度残差网络,所提算法将峰值信噪比(PSNR)提升了 0.53 dB,结构相似性(SSIM)达到了 0.9782。相比于增强型深度超分辨率网络,参数量仅为 31.7%,却取得了近似的重建表现。

**关键词** 成像系统; 超分辨率重建; 增强型多尺度残差网络; 增强型多尺度残差块; 通道注意力机制

中图分类号 TP391

文献标志码 A

DOI: 10.3788/LOP212884

## Image Super-Resolution Reconstruction Algorithm Based on Enhanced Multi-Scale Residual Network

Xu Jiao\*, Yuan Sannan

College of Electronics and Information Engineering, Shanghai University of Electric Power,  
Shanghai 200120, China

**Abstract** Most existing image super-resolution reconstruction algorithms have an extremely deep network structure, which leads to excessive parameters and an inability to fully extract features. To solve these problems, this study proposes an image super-resolution reconstruction algorithm based on an enhanced multi-scale residual network (EMSRN). The network consists of serial enhanced multi-scale residual blocks (EMSRB), and the backbone structure of the EMSRB is constructed using a residual block and parallel multi-dilation rate dilated convolution group, which effectively reduce the network parameters while obtaining the local and global multi-scale features of the image. The channel attention mechanism is used at the end of the block to adaptively weight extracted features, which enables the network to pay more attention to high-frequency information. Experiments show that, compared with the basic multi-scale residual network, the proposed algorithm improves the peak signal-to-noise ratio (PSNR) by 0.53 dB, and the structural similarity (SSIM) reaches 0.9782. Compared with the enhanced deep super-resolution network, the proposed algorithm achieves similar reconstruction performance with only 31.7% of its parameters.

**Key words** imaging system; super-resolution reconstruction; enhanced multi-scale residual network; enhanced multi-scale residual block; channel attention mechanism

## 1 引言

图像超分辨率(SR)重建是指从给定输入的一幅或多幅退化的低分辨率(LR)图像中恢复出具有更多高频信息的高分辨率(HR)图像的过程。它作为一种提高图像分辨率的重要手段,解决了实际情况下由于采集设备的性能不足或外界环境的干扰等因素导致无

法获取高分辨率图像的问题,被广泛应用于视频监控、医学成像和卫星遥感等众多领域<sup>[1-3]</sup>。图像超分辨率重建算法主要分为单幅图像超分辨率(SISR)和多幅图像超分辨率(MISR)两类,本文研究的是SISR重建算法。传统的SR算法主要包括基于插值法和基于重建法<sup>[4]</sup>。基于插值法通过在LR图像的相邻像素点进行插值得到HR图像,易于实现但线性模型拟合能

收稿日期: 2021-11-05; 修回日期: 2021-12-02; 录用日期: 2021-12-21; 网络首发日期: 2022-01-04

通信作者: \*xj15240039674@163.com

力有限,重建结果模糊;基于重建法通过求解退化模型的逆过程并结合先验知识来约束重建结果,改善了模糊效应但计算复杂。

近年来,由于深度学习<sup>[5-6]</sup>在计算机视觉领域取得了迅速发展,借助卷积神经网络(CNN)从大量训练数据对中学习LR与HR图像之间的端到端映射关系实现SR重建已经逐渐成为主流。2014年,Dong等<sup>[7]</sup>将CNN用于图像超分辨率重建并提出了超分辨率卷积神经网络(SRCNN),利用3层卷积层实现了LR和HR图像对之间的非线性映射,但网络层数过少,只能提取到非常有限的局部特征。2016年,Dong等<sup>[8]</sup>提出了快速超分辨率卷积神经网络(FSRCNN),将不做任何预处理的原始LR图像作为输入,在网络末端使用反卷积层进行上采样,明显降低了网络计算量。Shi等<sup>[9]</sup>于同年提出了一种高效的亚像素卷积神经网络(ESPCN),亚像素卷积层通过像素重排实现了对LR图像的上采样操作,相比反卷积层大大提高了重建效率,但这些方法重建的图像清晰度较低。随着网络深度的加深,He等<sup>[10]</sup>提出的残差网络(ResNet)缓解了卷积层数增加带来的梯度消失或爆炸问题。基于ResNet结构,Kim等<sup>[11]</sup>于2016年提出了非常深的超分辨率网络(VDSR),通过简单地堆叠20层卷积层并使用跳跃连接,逐层学习高频残差的同时加速了网络收敛,重建效果有所改善。同年Tai等<sup>[12]</sup>提出了深度递归残差网络(DRRN),利用对多个残差单元的递归学习实现参数共享,有效控制了网络参数量。但是,网络层数的增加产生了大量的冗余特征。为了解决该问题,同年Tong等<sup>[13]</sup>提出了超分辨率稠密网络(SRDenseNet),通过引入稠密的跳跃连接串联网络中的所有层,实现了低层和高层特征复用,缓解了特征冗余。2017年,Lai等<sup>[14]</sup>提出了拉普拉斯超分辨率网络(LapSRN),基于拉普拉斯金字塔结构采用渐进式反卷积上采样的思想,使用多阶段分支网络逐步实现SR重建,能够同时处理多种尺度因子图像,但重建精度有待提高。为了进一步改善重建效果,2017年Lim等<sup>[15]</sup>提出了增强型深度超分辨率网络(EDSR),通过删除每个残差块中的批归一化层并添置一个用于稳定网络训练的残差缩放层,使重建性能得到提升。然而,过深的网络层数会导致参数量庞大,难以提取深层特征。2018年,Li等<sup>[16]</sup>提出了多尺度残差网络(MSRN),通过使用不同尺寸的卷积核分别提取不同尺度的局部特征,经过全局特征融合后重建出的HR图像具有更多纹理细节,但同时也降低了网络的运行速度。2019年,Hui等<sup>[17]</sup>提出了信息多蒸馏网络(IMDN),通过构建轻量级的多级信息蒸馏模块逐步提取层次特征,加快了重建速度。2020年,Tian等<sup>[18]</sup>提出了由粗到细的超分辨率卷积神经网络(CFSRCNN),利用多个细化模块对融合后的低频和高频特征进行更为精准的特征学习,增强了SR模型的稳定性。同年Zhao等<sup>[19]</sup>提出了

像素注意力网络(PAN),通过在主模块和重建模块中引入参数量较少的像素注意力机制,改善了轻量型网络的重建性能。

目前大多数算法主要存在两方面问题:一方面盲目增加网络深度导致参数量过大,增加了训练难度;另一方面仅使用单一尺度卷积核构建网络模型,限制了局部感受野作用的范围,导致特征提取不充分。基于以上问题,本文提出了一个新颖的增强型多尺度残差网络(EMSRN)。该网络使用残差块和并行的多空洞率的空洞卷积组构成增强型多尺度残差块(EMSRB),作为网络的主要特征提取模块。在EMSRB模块中,首先使用残差块和 $3 \times 3$ 卷积提取LR图像的局部多尺度特征,相比于直接使用核尺寸为 $5 \times 5$ 的卷积,残差块可以获得同样大小的感受野且参数量更少;其次,为了进一步增强多尺度特征的提取能力,在局部多尺度特征融合的基础上使用并行的多空洞率的空洞卷积组,通过扩大不同范围感受野丰富了同一层特征的多样性;最后,为了加强网络对重要特征的识别能力,在EMSRB模块的最后加入通道注意力机制<sup>[20]</sup>,自适应地对所有通道特征分配缩放权重,使网络专注于获取高频特征。

## 2 算法设计

### 2.1 网络结构

本文提出的EMSRN主要由浅层特征提取模块、深层特征提取模块和重建模块3部分组成,其结构如图1所示。浅层特征提取模块由一个 $3 \times 3$ 卷积层组成,直接在低维空间提取图像特征有效减少了计算量;深层特征提取模块主要由8个EMSRB模块级联组成,EMSRB模块的叠加有利于网络提取深层次特征。为了降低由于深层网络训练带来的信息损失,将所有提取到的分层特征聚合,随后使用 $1 \times 1$ 卷积层进行降维,最后连接浅层和深层特征以促进网络收敛;重建模块由两个 $3 \times 3$ 卷积层和一个亚像素卷积层组成,在网络末端使用亚像素卷积层作为上采样方法,将输入的LR图像放大到对应目标尺寸的HR图像,提高了模型重建效率。

### 2.2 增强型多尺度残差块

为了增强网络的多尺度特征提取能力并忽略冗余特征,本文在EMSRN中设计了一个EMSRB模块作为网络的核心特征提取模块,其结构如图2所示。该模块主要由3部分组成:多尺度特征融合、多空洞率的空洞卷积组和通道注意力机制。

多尺度特征融合:模块的第1层和第2层均由 $3 \times 3$ 卷积层和残差块组成,将输入 $M_{n-1}$ 经第1层得到的两个支路输出 $P_1$ 和 $S_1$ 合并后经过模块的第2层获得对应输出 $P_2$ 和 $S_2$ ,最后使用 $1 \times 1$ 卷积压缩合并后的特征图获得降维后的输出 $O_1$ 。所有的输出可以表示为

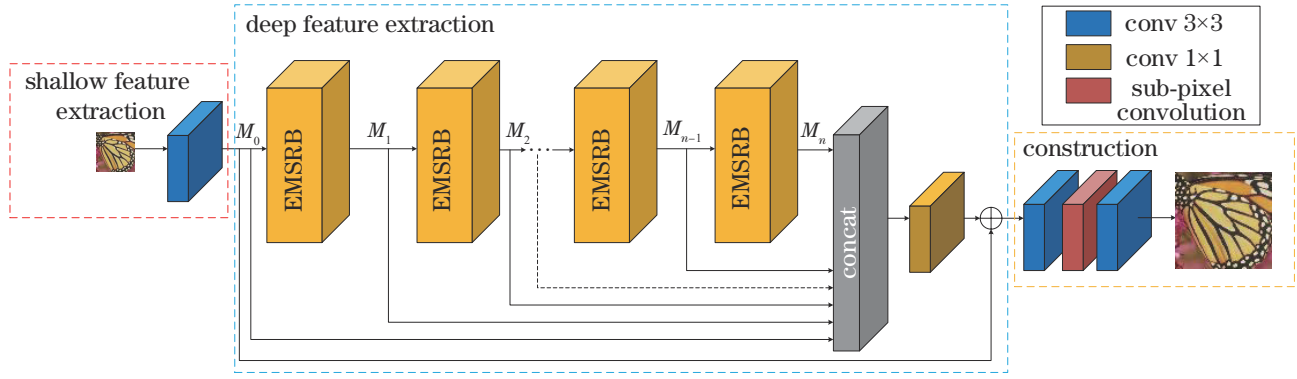


图 1 增强型多尺度残差网络结构

Fig. 1 Structure of enhanced multi-scale residual network

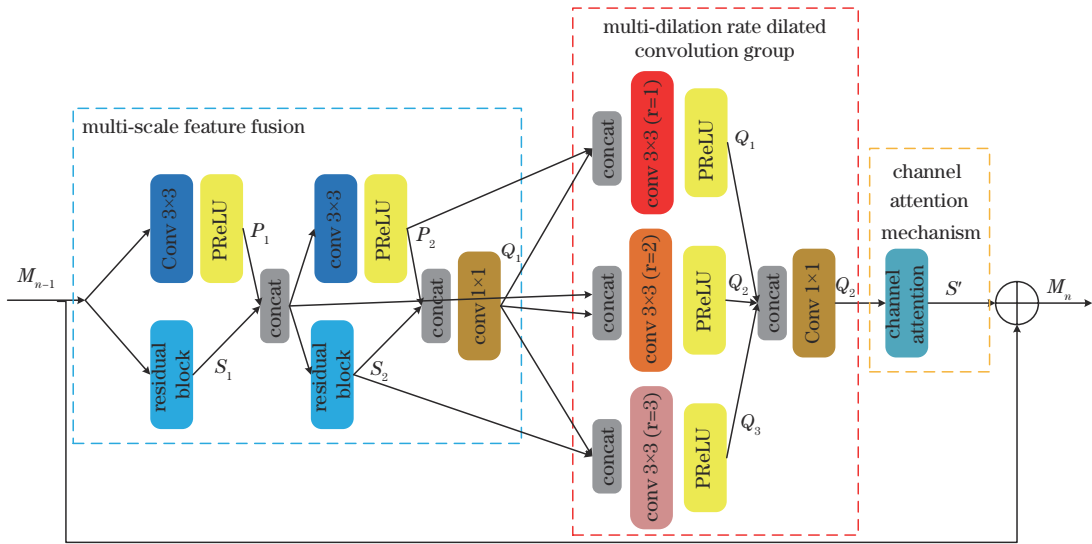


图 2 增强型多尺度残差块结构

Fig. 2 Structure of enhanced multi-scale residual block

$$P_1 = \gamma(\omega_{3 \times 3}^1 * M_{n-1} + b_{3 \times 3}^1), \quad (1)$$

$$S_1 = f_1(M_{n-1}), \quad (2)$$

$$P_2 = \gamma(\omega_{3 \times 3}^2 * [P_1, S_1] + b_{3 \times 3}^2), \quad (3)$$

$$S_2 = f_2([P_1, S_1]), \quad (4)$$

$$O_1 = \omega_{1 \times 1}^3 * [P_2, S_2] + b_{1 \times 1}^3, \quad (5)$$

式中:  $3 \times 3$  和  $1 \times 1$  卷积的权重  $\omega$  和偏置  $b$  的上标表示所处的层数;  $\gamma$  为参数化修正线性单元(PReLU)激活函数。  $f$  表示残差块(residual block)函数, 其中残差块由两个  $3 \times 3$  卷积和修正线性单元(ReLU)激活函数以及残差连接组成, 下标表示所处的层数。  $[P_1, S_1]$  和  $[P_2, S_2]$  表示特征图拼接操作。

多空洞率的空洞卷积组: 在模块的第 4 层使用并行的多空洞率的空洞卷积组。将输出  $O_1$  与  $P_2$  合并作为空洞率为 1 的  $3 \times 3$  卷积的输入, 经过激活函数 PReLU 后得到输出  $Q_1$ ; 将输出  $O_1$  与  $P_1$  和  $S_1$  合并作为空洞率为 2 的  $3 \times 3$  卷积的输入, 经过激活函数 PReLU 后得到输出  $Q_2$ ; 将输出  $O_1$  与  $S_2$  合并作为空洞率为 3 的  $3 \times 3$  卷积的输入, 经过激活函数 PReLU 后得到输出

$Q_3$ 。最后将所有输出进行通道合并并使用  $1 \times 1$  卷积降维, 得到输出  $O_2$ 。所有输出可以表示为

$$Q_1 = \gamma(\omega_{3 \times 3, r=1}^4 * [O_1, P_2] + b_{3 \times 3}^4), \quad (6)$$

$$Q_2 = \gamma(\omega_{3 \times 3, r=2}^4 * [O_1, P_1, S_1] + b_{3 \times 3}^4), \quad (7)$$

$$Q_3 = \gamma(\omega_{3 \times 3, r=3}^4 * [O_1, S_2] + b_{3 \times 3}^4), \quad (8)$$

$$O_2 = \omega_{1 \times 1}^5 * [Q_1, Q_2, Q_3] + b_{1 \times 1}^5, \quad (9)$$

式中:  $r$  表示空洞率, 并分别设置为 1、2、3;  $3 \times 3$  和  $1 \times 1$  卷积的权重  $\omega$  和偏置  $b$  的上标指的是所处层数。

通道注意力机制: 为了使网络更多地关注信息更加丰富的特征, 使用通道注意力机制自适应加权前一层输出的多尺度特征, 用公式表示为

$$S' = H_{CA}(O_2), \quad (10)$$

式中,  $H_{CA}(\cdot)$  表示通道注意力机制函数。

为了提高网络的性能, 实现加速收敛, 最后使用局部残差连接将通道注意力机制输出特征图与模块输入相加得到模块输出:

$$M_n = M_{n-1} + S', \quad (11)$$

式中,  $M_n$  表示 EMSRB 模块的输出。

### 2.3 通道注意力机制

利用前一层每个通道特征之间的相互关联,使用通道注意力机制建立学习模型。根据不同特征输出的

重要程度分别对应赋予不同的权重,从而挖掘出更有利于图像重建的关键特征,其结构如图 3 所示。

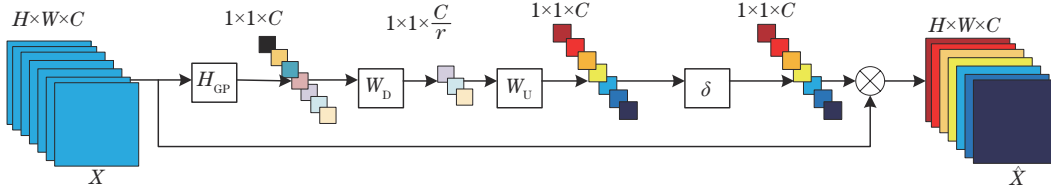


图 3 通道注意力机制结构

Fig. 3 Structure of channel attention mechanism

通道注意力机制的整体映射关系可以表示为

$$\hat{X} = H_{CA}(X), \quad (12)$$

式中:  $H_{CA}(\cdot)$  为通道注意力机制函数;  $X$  表示模型输入特征;  $\hat{X}$  表示模型输出特征。具体来说,对于输入尺寸为  $H \times W \times C$  的卷积层,使用全局平均池化压缩空间维度特征,归一化后得到  $1 \times 1 \times C$  的通道统计量,它是所有通道描述符的集合,有助于描述整个图像的全局特征信息。则第  $k$  个通道描述符可以表示为

$$z_k = H_{GP}(x_k) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_k(i, j), \quad (13)$$

式中:  $H_{GP}(\cdot)$  表示全局平均池化函数;  $x_k(i, j)$  表示第  $k$  个通道特征图上  $(i, j)$  处的像素值。

为了更好地捕获通道特征之间的相关性,突出重要特征的代表能力,使用 Sigmoid 激活函数自适应学习通道间的非线性特征响应,获得通道统计量的权重缩放:

$$s = \delta[W_U \sigma(W_D z)], \quad (14)$$

式中:  $W_D$  和  $W_U$  分别表示使用  $1 \times 1$  卷积对通道统计量  $z$  先后进行通道降维和升维  $r$  倍后的权重集;  $\sigma(\cdot)$  为 ReLU 激活函数;  $\delta(\cdot)$  为 Sigmoid 激活函数;  $s$  为通道统计量经过自适应缩放后得到的权重系数。最后将卷积层所有通道的输入特征图  $X$  乘以权重系数,得到加权后的输出特征图:

$$\hat{X} = X \cdot s. \quad (15)$$

### 2.4 损失函数

为了减小重建误差,利用损失函数对 EMSRN 进行最优化。目前大多数算法使用最多的损失函数为  $L_2$  损失函数<sup>[10-12]</sup>和  $L_1$  损失函数<sup>[14-16]</sup>。由于  $L_2$  损失函数往往会导致重建结果过于平滑,因此本文使用  $L_1$  损失函数。给定包含  $N$  个 LR 和 HR 图像对的训练集  $\{I_{LR}^i, I_{HR}^i\}_{i=1}^N$ , 优化目标可表示为

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|H_{EMSRN}(I_{LR}^i) - I_{HR}^i\|, \quad (16)$$

式中:  $\theta = \{w_i, b_i\}$  表示 EMSRN 的学习参数集;  $H_{EMSRN}(\cdot)$  表示 EMSRN 函数。采用随机梯度下降法优化  $L_1$  损失函数。

## 3 实验结果与分析

### 3.1 实验数据与设置

本文使用公开的分流 2K 数据集<sup>[21]</sup>作为训练集,其中 800 张 HR 的红绿蓝(RGB)图像作为训练图像。在测试阶段,采用 5 个广泛使用的基准数据集 Set5<sup>[22]</sup>、Set14<sup>[23]</sup>、BSD100<sup>[24]</sup>、Urban100<sup>[25]</sup>和 Manga109<sup>[26]</sup>作为测试集,对 EMSRN 模型性能进行评估。

将训练集 DIV2K 中经过双三次插值后的 LR 图像随机裁剪成  $48 \times 48$  大小的图像块,并通过  $90^\circ$ 、 $180^\circ$ 、 $270^\circ$  的随机旋转和水平翻转实现数据增强。在训练阶段,使用 ADAM 优化器<sup>[27]</sup>优化网络参数,其中  $\beta_1 = 0.9$ ,  $\beta_2 = 0.99$ ,  $\epsilon = 10^{-8}$ 。学习率被初始化为 0.0001,并且每隔 200 个周期学习率减半。每个批次输入设置为 8。

本文实验选择 Pytorch 框架构建 EMSRN 模型,并使用 NVIDIA GeForce RTX 2080Ti 进行 GPU 加速。整个网络模型在 RGB 空间进行训练, Y 通道进行测试。为了衡量图像重建效果,本文采用峰值信噪比(PSNR)和结构相似性(SSIM)<sup>[28]</sup>作为客观评价指标。PSNR 指标主要基于均方误差(MSE)计算 SR 图像与 HR 图像之间的对应像素值误差,单位为 dB; SSIM 指标结合亮度、对比度和结构信息度量 SR 图像与 HR 图像之间的相似程度,取值范围为  $[0, 1]$ 。这两种指标的计算公式分别为

$$C_{PSNR} = 10 \cdot \lg \left[ \frac{(2^n - 1)^2}{E_{MSE}} \right] = 20 \cdot \lg \left( \frac{2^n - 1}{\sqrt{E_{MSE}}} \right), \quad (17)$$

$$E_{MSE} = \frac{1}{l \times m} \sum_{i=1}^l \sum_{j=1}^m \|I_{SR}(i, j) - I_{HR}(i, j)\|^2, \quad (18)$$

$$S_{SSIM}(I_{SR}, I_{HR}) = \frac{(2\mu_{I_{SR}}\mu_{I_{HR}} + c_1)(2\sigma_{I_{SR}I_{HR}} + c_2)}{(\mu_{I_{SR}}^2 + \mu_{I_{HR}}^2 + c_1)(\sigma_{I_{SR}}^2 + \sigma_{I_{HR}}^2 + c_2)}, \quad (19)$$

式中:  $n$  为每像素的比特数,一般取 8;  $l, m$  为图像的尺寸;  $I_{SR}(i, j)$  和  $I_{HR}(i, j)$  分别表示 SR 图像和 HR 图像位于  $(i, j)$  处的像素值;  $\mu_{I_{SR}}$  和  $\mu_{I_{HR}}$  分别为 SR 图像和 HR 图像的像素平均值;  $\sigma_{I_{SR}}$  和  $\sigma_{I_{HR}}$  分别为 SR 图像和 HR 图像的像素标准差;  $\sigma_{I_{SR}I_{HR}}$  为 SR 图像与 HR 图像间的协方

差;  $c_1$  和  $c_2$  为常数。

### 3.2 通道注意力机制的有效性

为了充分验证算法中引入通道注意力机制的有效性,通过去除 EMSRN 中的通道注意力机制得到 EMSRN-NOCA 模型,并与原模型进行对比实验。分别对 EMSRN-NOCA 和 EMSRN 训练 1000 个 epoch,在 Manga109 测试集上得到  $\times 2$ 、 $\times 3$ 、 $\times 4$  和  $\times 8$  重建后的 PSNR 和 SSIM 值,实验结果如表 1 所示。从表中可以看出,相比于 EMSRN-NOCA,EMSRN 随着放大倍数的增长均取得更高的 PSNR 和 SSIM,其平均 PSNR 值得到了 0.13 dB 的提升, $\times 8$  时虽然 PSNR 没有明显差异,但 SSIM 有所提高。实验结果表明本文在 EMSRB 模块中引入通道注意力机制是有效的,有助于网络关注和获取更多图像细节信息,提升了算法的

表 1 EMSRN-NOCA 和 EMSRN 的比较

Table 1 Comparison between EMSRN-NOCA and EMSRN

Algorithm	Scale	Manga109 PSNR/SSIM
EMSRN-NOCA	$\times 2$	39.07/0.9777
EMSRN	$\times 2$	39.20/0.9782
EMSRN-NOCA	$\times 3$	34.01/0.9478
EMSRN	$\times 3$	34.16/0.9483
EMSRN-NOCA	$\times 4$	30.85/0.9142
EMSRN	$\times 4$	30.95/0.9148
EMSRN-NOCA	$\times 8$	24.74/0.7852
EMSRN	$\times 8$	24.72/0.7857

表 2 不同算法的 PSNR 和 SSIM 值

Table 2 PSNR and SSIM values of different algorithms

Algorithm	Scale	Set5	Set14	BSD100	Urban100	Manga109
		PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic	$\times 2$	33.69/0.9284	30.34/0.8675	29.57/0.8434	26.88/0.8438	30.82/0.9332
SRCNN	$\times 2$	36.31/0.9535	32.26/0.9053	31.16/0.8859	29.30/0.8939	35.16/0.9663
FSRCNN	$\times 2$	36.78/0.9561	32.57/0.9089	31.38/0.8894	29.74/0.9009	36.26/0.9700
ESPCN	$\times 2$	36.47/0.9544	32.32/0.9067	31.17/0.8867	29.21/0.8925	35.28/0.9666
VDSR	$\times 2$	37.16/0.9582	32.87/0.9126	31.75/0.8951	30.74/0.9146	36.42/0.9730
LapSRN	$\times 2$	36.91/0.9573	32.71/0.9105	31.59/0.8922	30.26/0.9090	36.09/0.9717
EDSR	$\times 2$	38.11/0.9601	<b>33.92/0.9195</b>	<b>32.32/0.9013</b>	-/-	-/-
MSRN	$\times 2$	38.09/0.9607	33.73/0.9182	32.22/0.9001	32.29/0.9300	38.65/0.9771
IMDN	$\times 2$	37.91/0.9594	33.59/0.9169	32.15/0.8987	32.14/0.9274	38.79/0.9764
CFSRCNN	$\times 2$	37.79/0.9591	33.50/0.9165	32.10/0.8987	32.07/0.9273	38.16/0.9751
PAN	$\times 2$	38.00/0.9605	33.59/0.9181	32.18/0.8997	32.01/0.9273	38.70/0.9773
EMSRN (proposed)	$\times 2$	<b>38.19/0.9613</b>	<b>33.78/0.9195</b>	32.30/0.9012	<b>32.74/0.9342</b>	<b>39.20/0.9782</b>
Bicubic	$\times 3$	30.41/0.8655	27.64/0.7722	27.21/0.7344	24.46/0.7411	26.96/0.8555
SRCNN	$\times 3$	32.60/0.9088	29.21/0.8198	28.30/0.7840	26.04/0.7955	30.09/0.9098
FSRCNN	$\times 3$	32.51/0.9054	29.17/0.8181	28.24/0.7821	25.97/0.7917	30.00/0.9051
ESPCN	$\times 3$	32.56/0.9073	29.19/0.8195	28.26/0.7834	25.98/0.7929	30.01/0.9063
VDSR	$\times 3$	33.54/0.9214	29.69/0.8315	28.73/0.7963	27.05/0.8265	31.55/0.9312
EDSR	$\times 3$	<b>34.65/0.9282</b>	<b>30.52/0.8462</b>	<b>29.25/0.8093</b>	-/-	-/-

重建能力。

### 3.3 重建算法对比

为了衡量本文算法的图像重建效果,将 EMSRN 与传统的 Bicubic 算法以及其他 10 种先进的基于 CNN 的 SR 算法 SRCNN<sup>[7]</sup>、FSRCNN<sup>[8]</sup>、ESPCN<sup>[9]</sup>、VDSR<sup>[11]</sup>、LapSRN<sup>[14]</sup>、EDSR<sup>[15]</sup>、MSRN<sup>[16]</sup>、IMDN<sup>[17]</sup>、CFSRCNN<sup>[18]</sup> 和 PAN<sup>[19]</sup> 进行重建结果比较。其中 SRCNN、FSRCNN、ESPCN、VDSR、LapSRN 和 MSRN 的结果均为重新训练所得,IMDN、CFSRCNN 和 PAN 的结果由官方提供预训练模型测试所得,EDSR 的结果引用于文献[16]。

#### 3.3.1 客观定量分析

表 2 比较了本文与对比算法分别在 5 个基准数据集上进行测试得到的  $\times 2$ 、 $\times 3$ 、 $\times 4$  和  $\times 8$  重建后的 PSNR 与 SSIM 值,PSNR 与 SSIM 值越高,表示重建的效果越好。其中黑体数字表示最优值,斜体数字表示次优值。从表中数据可以看出,除了在部分测试集上 EMSRN 的 PSNR 和 SSIM 值略低于 EDSR 外,相比于 MSRN 和其他算法,EMSRN 在所有放大倍数上均取得最优指标,其中 PSNR 值比 MSRN 最多提升了 0.53 dB,SSIM 最高达到 0.9782。虽然本文算法取得了与 EDSR 差不多的重建效果,但其参数量远小于 EDSR,具体对比将在 3.4 节网络参数量对比实验中说明。实验结果表明基于本文算法的图像 SR 重建效果优势更加明显,重建性能显著提高。

表 2 (续)

Algorithm	Scale	Set5	Set14	BSD100	Urban100	Manga109
		PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
MSRN	×3	34.47/0.9276	30.38/0.8434	29.13/0.8062	28.31/0.8558	33.58/0.9452
IMDN	×3	34.32/0.9259	30.31/0.8409	29.07/0.8036	28.15/0.8510	33.58/0.9434
CFSRCNN	×3	34.23/0.9256	30.26/0.8409	29.02/0.8033	28.03/0.8495	33.27/0.9418
PAN	×3	34.40/0.9272	30.36/0.8422	29.10/0.8049	28.11/0.8510	33.58/0.9448
EMSRN (proposed)	×3	<b>34.65/0.9290</b>	<b>30.54/0.8461</b>	29.24/0.8089	<b>28.70/0.8634</b>	<b>34.16/0.9483</b>
Bicubic	×4	28.43/0.8022	26.10/0.6936	25.97/0.6517	23.14/0.6599	24.91/0.7826
SRCNN	×4	30.22/0.8597	27.40/0.7489	26.78/0.7074	24.29/0.7141	27.10/0.8457
FSRCNN	×4	30.44/0.8595	27.51/0.7507	26.85/0.7090	24.44/0.7188	27.42/0.8432
ESPCN	×4	30.25/0.8566	27.37/0.7487	26.77/0.7072	24.26/0.7114	27.00/0.8398
VDSR	×4	31.18/0.8819	28.00/0.7673	27.19/0.7229	25.09/0.7494	28.50/0.8812
LapSRN	×4	31.52/0.8879	28.17/0.7729	27.28/0.7278	25.26/0.7596	28.85/0.8900
EDSR	×4	<b>32.46/0.8968</b>	<b>28.80/0.7876</b>	<b>27.71/0.7420</b>	-/-	-/-
MSRN	×4	32.19/0.8952	28.63/0.7837	27.61/0.7377	26.16/0.7894	30.53/0.9093
IMDN	×4	32.19/0.8936	28.57/0.7803	27.54/0.7342	26.03/0.7829	30.44/0.9065
CFSRCNN	×4	32.06/0.8920	28.57/0.7801	27.52/0.7331	26.02/0.7823	30.30/0.9048
PAN	×4	32.13/0.8948	28.61/0.7822	27.59/0.7363	26.11/0.7854	30.51/0.9095
EMSRN (proposed)	×4	32.43/ <b>0.8980</b>	28.77/0.7863	27.69/0.7408	<b>26.51/0.7995</b>	<b>30.95/0.9148</b>
Bicubic	×8	24.40/0.6045	23.19/0.5110	23.67/0.4808	20.74/0.4841	21.46/0.6138
SRCNN	×8	24.10/0.6599	23.02/0.5704	23.53/0.5479	20.68/0.5157	21.41/0.6470
FSRCNN	×8	25.20/0.6832	23.77/0.5879	24.08/0.5599	21.21/0.5342	22.21/0.6630
ESPCN	×8	25.41/0.6895	23.92/0.5923	24.13/0.5629	21.31/0.5382	22.32/0.6677
VDSR	×8	25.79/0.7183	24.20/0.6084	24.37/0.5744	21.58/0.5614	22.82/0.7056
LapSRN	×8	25.96/0.7276	24.35/0.6168	24.49/0.5811	21.74/0.5727	23.12/0.7202
MSRN	×8	26.87/0.7714	24.83/0.6381	24.74/0.5950	22.32/0.6122	24.39/0.7743
EMSRN (proposed)	×8	<b>26.88/0.7759</b>	<b>24.99/0.6434</b>	<b>24.79/0.5987</b>	<b>22.55/0.6230</b>	<b>24.72/0.7857</b>

3.3.2 主观视觉效果

图 4~7 分别展示了不同算法在 ×2、×3、×4 和 ×8 图像 SR 重建后的视觉效果对比。从图中可以观察到,经过 Bicubic 重建得到的 HR 图像较为模糊,视觉效果较差;相比于 MSRN 等算法重建后图像局部细节失真较为严重,本文算法有效还原了边缘和纹理等细节信息。如放大倍数为 ×3 时矩形框区域内重建对比图所示,EMSRN 几乎准确恢复出了条纹的形状,而

MSRN 等却重建出了完全错误的方向。实验结果表明本文所提算法能够更好地表征 HR 特征空间,重建后的图像恢复了更多高频信息,更接近于原始 HR 图像。

3.4 网络参数量比较

将本文与对比算法应用于 Set5 测试集,以 ×2 为例,进行参数量与 PSNR 对应关系的比较,如图 8 所示。其中 EMSRN 的参数量为  $13 \times 10^6$ ,仅为 EDSR 的

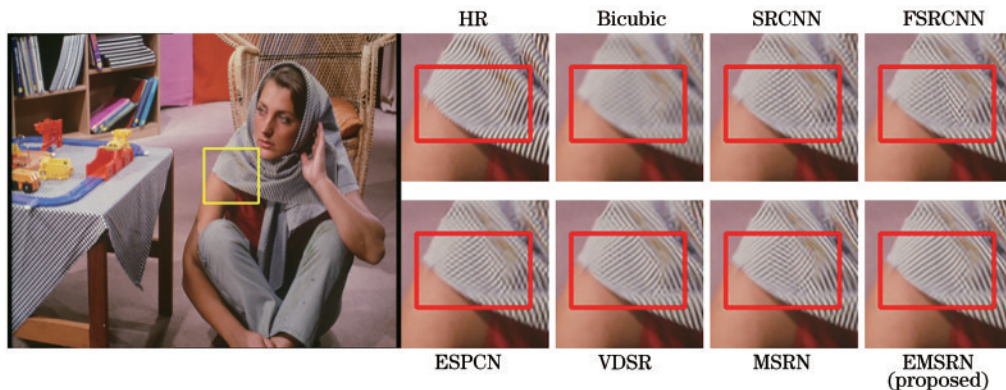


图 4 放大因子为 ×2 时不同算法的图像视觉效果

Fig. 4 Image visual effects of different algorithms with scale factor ×2

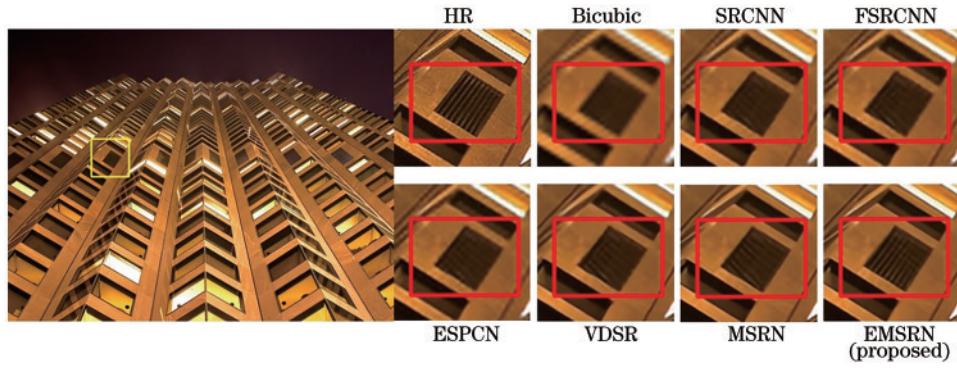


图 5 放大因子为  $\times 3$  时不同算法的图像视觉效果

Fig. 5 Image visual effects of different algorithms with scale factor  $\times 3$

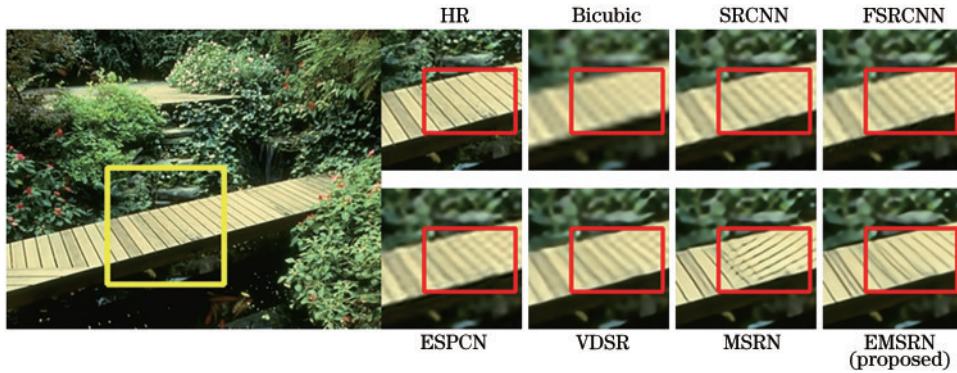


图 6 放大因子为  $\times 4$  时不同算法的图像视觉效果

Fig. 6 Image visual effects of different algorithms with scale factor  $\times 4$

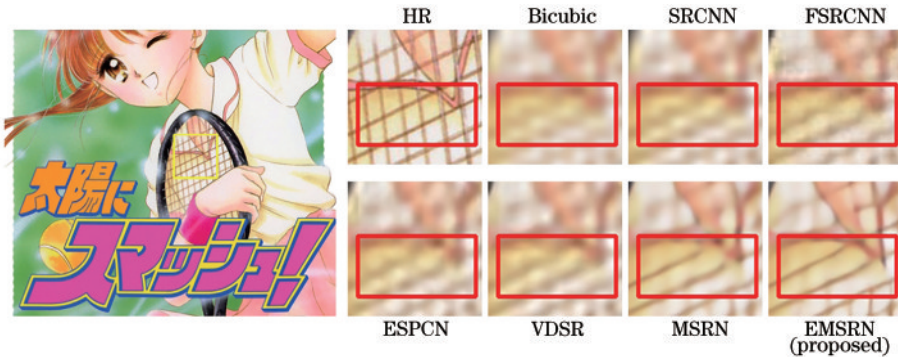


图 7 放大因子为  $\times 8$  时不同算法的图像视觉效果

Fig. 7 Image visual effects of different algorithms with scale factor  $\times 8$

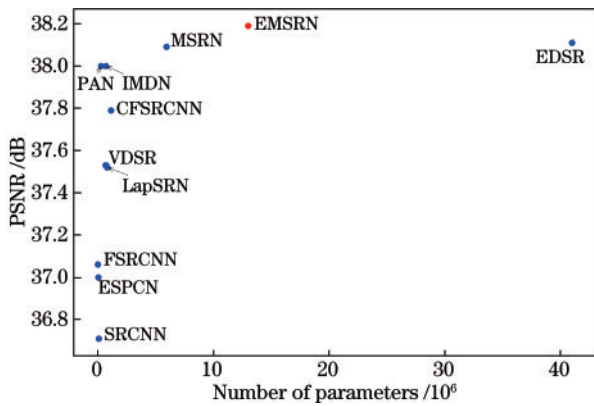


图 8 不同算法的网络参数与 PSNR 对应关系比较

Fig. 8 Comparison of network parameters and PSNR correspondence for different algorithms

31.7%，但却取得更高的 PSNR；同时 EMSRN 的 PSNR 在远高于 MSRN 的情况下，网络参数量只有略微增长。实验结果表明，本文算法在提高重建性能的同时有效降低了网络参数量，二者实现了良好平衡。

#### 4 结 论

针对现有超分辨率重建算法对图像特征利用不足和参数量庞大等问题，提出了一种 EMSRN，通过级联多个 EMSRB 充分提取图像特征。该模块使用残差块和并行的多空洞率的空洞卷积组捕获了图像在不同尺度上的上下文信息，增强了多尺度特征的提取能力。同时引入通道注意力机制在模块末端以加强网络对重要特征的判别能力，有助于待重建图像恢复更多高频

信息。最后在重建模块通过亚像素卷积实现 HR 图像重建。在基准数据集上的实验结果表明,所提算法在客观评价指标和主观视觉效果上均取得更好的图像重建结果,并且网络参数量显著减少。但是,这仍然无法满足实际应用的需求,在以后的工作中,将继续朝着轻量化方向进行研究。

## 参 考 文 献

- [1] 陈星宇, 张伟劲, 孙伟智, 等. 基于多尺度与多重残差网络的图像超分辨率重建[J]. 激光与光电子学进展, 2020, 57(18): 181009.  
Chen X Y, Zhang W J, Sun W Z, et al. Super-resolution reconstruction of images based on multi-scale and multi-residual network[J]. Laser & Optoelectronics Progress, 2020, 57(18): 181009.
- [2] 席志红, 袁昆鹏. 基于残差通道注意力和多级特征融合的图像超分辨率重建[J]. 激光与光电子学进展, 2020, 57(4): 041504.  
Xi Z H, Yuan K P. Super-resolution image reconstruction based on residual channel attention and multilevel feature fusion[J]. Laser & Optoelectronics Progress, 2020, 57(4): 041504.
- [3] Xu C, Xu T F, Yan G, et al. Super-resolution compressive spectral imaging via two-tone adaptive coding[J]. Photonics Research, 2020, 8(3): 395-411.
- [4] 曲海成, 唐博文, 袁贵森. 改进的超分辨率图像重建算法[J]. 激光与光电子学进展, 2021, 58(2): 0210018.  
Qu H C, Tang B W, Yuan G S. Improved super-resolution image reconstruction algorithm[J]. Laser & Optoelectronics Progress, 2021, 58(2): 0210018.
- [5] Gao L, Chai Y, Zibar D, et al. Deep learning in photonics: introduction[J]. Photonics Research, 2021, 9(8): DLP1-DLP3.
- [6] Lim J, Ayoub A B, Psaltis D. Three-dimensional tomography of red blood cells using deep learning[J]. Advanced Photonics, 2020, 2(2): 026001.
- [7] Dong C, Loy C C, He K M, et al. Learning a deep convolutional network for image super-resolution[M]// Fleet D, Pajdla T, Schiele B, et al. Computer vision-ECCV 2014. Lecture notes in computer science. Cham: Springer, 2014, 8692: 184-199.
- [8] Dong C, Loy C C, Tang X O. Accelerating the super-resolution convolutional neural network[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9906: 391-407.
- [9] Shi W Z, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1874-1883.
- [10] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [11] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1646-1654.
- [12] Tai Y, Yang J, Liu X M. Image super-resolution via deep recursive residual network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2790-2798.
- [13] Tong T, Li G, Liu X J, et al. Image super-resolution using dense skip connections[C]//2017 IEEE International Conference on Computer Vision, October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 4809-4817.
- [14] Lai W S, Huang J B, Ahuja N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5835-5843.
- [15] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1132-1140.
- [16] Li J C, Fang F M, Mei K F, et al. Multi-scale residual network for image super-resolution[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11212: 527-542.
- [17] Hui Z, Gao X B, Yang Y C, et al. Lightweight image super-resolution with information multi-distillation network[C]//MM '19: Proceedings of the 27th ACM International Conference on Multimedia, October 21-25, 2019, Nice, France. New York: ACM Press, 2019: 2024-2032.
- [18] Tian C W, Xu Y, Zuo W M, et al. Coarse-to-fine CNN for image super-resolution[J]. IEEE Transactions on Multimedia, 2021, 23: 1489-1502.
- [19] Zhao H Y, Kong X T, He J W, et al. Efficient image super-resolution using pixel attention[M]//Bartoli A, Fusiello A. Computer vision-ECCV 2020 workshops. Lecture notes in computer science. Cham: Springer, 2020, 12537: 56-72.
- [20] Zhang Y L, Li K P, Li K, et al. Image super-resolution using very deep residual channel attention networks[M]//Ferrari V, Hebert M, Sminchisescu C, et al. Computer vision-ECCV 2018. Lecture notes in computer science. Cham: Springer, 2018, 11211: 294-310.
- [21] Agustsson E, Timofte R. NTIRE 2017 challenge on single image super-resolution: dataset and study[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1122-1131.
- [22] Bevilacqua M, Roumy A, Guillemot C, et al. Low-



- complexity single-image super-resolution based on nonnegative neighbor embedding[C]//Proceedings of the British Machine Vision Conference 2012, September 3-7, 2012, Surrey, UK. London: British Machine Vision Association, 2012: 135.1-135.10.
- [23] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[M]//Boissonnat J D, Chenin P, Cohen A, et al. Curves and surfaces. Lecture notes in computer science. Heidelberg: Springer, 2012, 6920: 711-730.
- [24] 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.
- [25] Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 5197-5206.
- [26] 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.
- [27] Kingma D P, Ba J L. Adam: a method for stochastic optimization[C]//Proceedings of the 3rd International Conference on Learning Representations(ICLR), May 7-9, 2015, San Diego, CA, USA. [S.l.: s.n.], 2015.
- [28] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.