

一种轻量级的多尺度通道注意图像 超分辨率重建网络

周登文, 李文斌, 李金新, 黄志勇
(华北电力大学控制与计算机工程学院, 北京 102206)

摘要: 近年来, 基于深度卷积神经网络的图像超分辨率技术取得了突出进展, 并主导了当前的超分辨率技术的研究. 但是, 性能的改进, 往往以参数数量的急剧增加为代价, 这限制了超分辨率方法的实际应用. 本文设计了一个轻量级单图像超分辨率深度卷积网络, 主要贡献包括: 提出了一个多尺度的特征融合模块, 使用不同感受野的卷积核, 提取多种尺度的特征; 提出了一个通道搅乱注意力模块, 促进特征通道之间的信息流动, 并增强特征选择能力; 提出了一个全局特征融合连接模块, 提高浅层特征的利用率. 实验证明, 本文方法与当前代表性的方法 MSRN (Multi-Scale Residual Network) 相比, 参数量减少了 3/4, 重建的高分辨率图像的主观和客观质量均显著更好.

关键词: 超分辨率; 深度学习; 卷积神经网络; 注意力机制; 多尺度特征

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2022)10-2336-11

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20201089

Image Super-Resolution Reconstruction Based on Lightweight Multi-Scale Channel Attention Network

ZHOU Deng-wen, LI Wen-bin, LI Jin-xin, HUANG Zhi-yong

(School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China)

Abstract: Recently, image super-resolution technology based on deep convolutional neural network has made remarkable achievements and has become popular in the current super-resolution technology. However, superior performance is often at the expense of the large number of parameter amounts, which limits the real-world applications for single image super-resolution. In this paper, a lightweight single image super-resolution deep convolutional network is proposed. The main contributions of this paper are as follows: a multi-scale feature fusion block is proposed to extract multiple features via convolution kernels with different receptive fields; the channel shuffle attention mechanism we designed promotes the flow of the information across feature channels, which enhances the ability of feature selection; a global feature fusion connection is proposed to improve the feature utilization. Extensive experiments demonstrate that the parameter amounts of our method reduced by 3/4 compared with the current state-of-the-art MSRN method, while subjective visual and objective quality of the reconstructed high-resolution image are perform significantly better.

Key words: super-resolution; deep learning; convolutional neural network; attention mechanism; multi-scale feature

1 引言

单图像超分辨率 (Single Image Super-Resolution, SISR) 技术是从给定的低分辨率图像 (Low-Resolution, LR) 恢复出对应的高分辨率图像 (High-Resolution, HR), 是计算机视觉领域的经典问题, 被广泛用于遥感成像^[1]、视频监控^[2]和医学图像处理^[3]等领域.

单图像超分辨率是一个病态逆问题: 一个 LR 图像

可对应多个 HR 图像^[4], 重建的 HR 图像往往会出现细节丢失、边缘走样和模糊等缺陷. 深度学习技术极大地推动了计算机视觉领域的迅速发展^[5,6], 当前绝大多数 SISR 算法都是基于端到端的深度学习技术的, 即直接学习 LR 和 HR 之间的映射. 基于深度学习的 SISR 方法虽然取得了很大的进步, 但是在实际应用中依然存在问题: 更好的结果往往依赖更深的网络, 更多的参数也需要更多的训练数据. 这需要更长的训练和推理时

间,以及更强的计算能力和存储能力.因此,其实用性受到了很大的限制,尤其是在资源受限的移动设备中.例如,EDSR(Enhanced Deep Super-Resolution)^[7]通过堆叠残差块的方式构建了一个宽且深的网络,赢得了2017年NTIRE(New Trends in Image Restoration and Enhancement)超分辨率比赛冠军,其参数量有43 MB之多.RDN(Residual Dense Network)^[8]将稠密连接嵌入残差块中,构建了一个更深的网络,其性能超过了EDSR,参数量有所降低,也约有22 MB之多.当前也有一些使用轻量级网络的SISR方法.Dong等人^[9]首次将卷积神经网络用于SISR重建,他们的方法被称为SRCNN(Super-Resolution Convolutional Neural Network).SRCNN只有3个卷积层,但是,取得了较好的结果.Kim等人^[10]提出的VDSR(Super-Resolution using Very Deep convolutional networks),使用了残差网^[11]和梯度裁剪技术,扩展到20个卷积层,进一步提升了性能.DRRN(Deep Recursive Residual Network)^[12]结合递归学习、局部残差学习和全局残差学习,连续递归残差块,进一步加深了网络,效果优于VDSR.MemNet^[13]提出了以记忆模块和门控单元为基本单元的持续记忆网络,同时使用递归的技术,其SISR性能优于DRRN.CARN(Cascading Residual Network)^[14]构建了一个级联的残差网,其SISR性能优于DRRN.CBPN(Compact Back-Projection Network)^[15]通过级联上/下采样层,在LR和HR空间中提取特征信息,其SISR性能优于CARN.Hui等人^[16]提出的IMDN(Information Multi-Distillation Network)方法在残差块内逐步提取特征,通过通道注意力机制对特征进行选择,其SISR性能超过了CBPN.Li等人^[17]提出的MSRN(Multi-Scale Residual Network)方法,运用了全局融合层特征,其SISR性能超越IMDN.在这些轻量级SISR方法中,MSRN性能最好,参数量也更大(6 MB);SRCNN参数量最小(57 KB),性能也是最低的.

本文提出了一个新的轻量级SISR网络,主要贡献包括:(1)提出了一个多尺度特征融合块(MultiScale Feature Fusion Block, MSFFB),使用不同感受野的卷积核(5×5, 3×3和1×1),提取和融合多种尺度的特征;(2)提出了通道搅乱注意力模块(Channel Scramble Attention Module, CSAM),促进通道之间信息流动,增强特征选择能力;(3)提出了全局特征融合连接(Global Feature Fusion Connection, GFFC)模块,提高浅层特征的利用率,增强图像细节信息.

2 相关工作

2.1 注意力机制

注意力机制源于人类视觉系统的研究.人类视觉倾向于重点关注显著区域,忽略无用的信息,这提高了

大脑处理信息的效率^[18].Hu等人^[19]利用注意力机制提出的压缩-扩张(Squeeze-and-Excitation, SE)模块,可以显式建模特征通道之间的相互依赖关系,通过学习,自动获取每个特征通道的重要程度,然后依据重要性抑制用处不大的特征,提高了图像分类网络的性能.Zhang等人^[20]提出的RCAN(Residual Channel Attention Networks)方法,引入了Hu等人^[19]的注意力模型.其通道注意力模块,可自适应地选择信息更丰富的特征通道,改进了SISR的性能.不过,该注意力机制的特征通道之间是相互独立的,这限制了特征信息在通道之间的流动.

2.2 递归学习

增加网络深度,一般可以提升SISR方法的性能.但是,更深的网络,一般有更多的参数,需要更多的训练数据.实际应用中,训练数据的获取往往受限,网络训练过拟合的风险也在增加.递归的网络一个优势是:可以增加网络深度,但不增加参数量.DRCN(Deeply-Recursive Convolutional Network)^[21]首次在SISR方法中引入了递归技术:使用一个卷积层作为递归单元,权值在该递归单元之间共享.DRRN^[12]改进了递归技术的使用:以残差块作为递归单元,参数在残差块之间共享,也改进了DRCN的性能.类似地,Tai等人^[13]提出的MemNet,也以残差块作为递归单元构建网络.

3 本文方法

通常,更深的网络,能获得更好的SISR性能^[7,8].残差网络缓解了深度网络的梯度消失/爆炸问题,但是简单地堆叠残差块^[7,8]会使参数量急剧增加,这是一个两难的选择.

增加网络深度,但不增加参数的一种可行的方式是:采用递归的结构,在残差块之间参数共享^[12].但是,由于严重的梯度消失/爆炸问题^[21],训练深度递归网络非常困难.本文提出了一个新的递归网络——多尺度通道注意超分辨率网络(Multi-scale Channel attention Super-resolution Network, MCSN):MCSN把递归的残差块嵌套在一起,形成一个嵌套块构件,再递归这个嵌套块.MCSN包括3个部分:浅层特征提取(Shallow Feature Extraction Block, SFEB)、特征串联组(Features Connection Group, FCG)和重建模块(Reconstruction Block, RB),如图1(a)所示.FCG包含3个嵌套块(Nested Unit, NU),每个NU包含3个递归的层次特征融合块(Hierarchical Feature Fusion Block, HFFB),HFFB的结构参见图1(b).HFFB主要由3个基本嵌套单元(Basic Nested Cell, BNC)组成,BNC的结构参见图1(c).BNC主要由多尺度特征融合块(MSFFB)[图1(c)]和通道搅乱注意力模块(Channel Scramble Attention Module, CSAM)[图1(d)].下面详细介绍MCSN的各个部分.

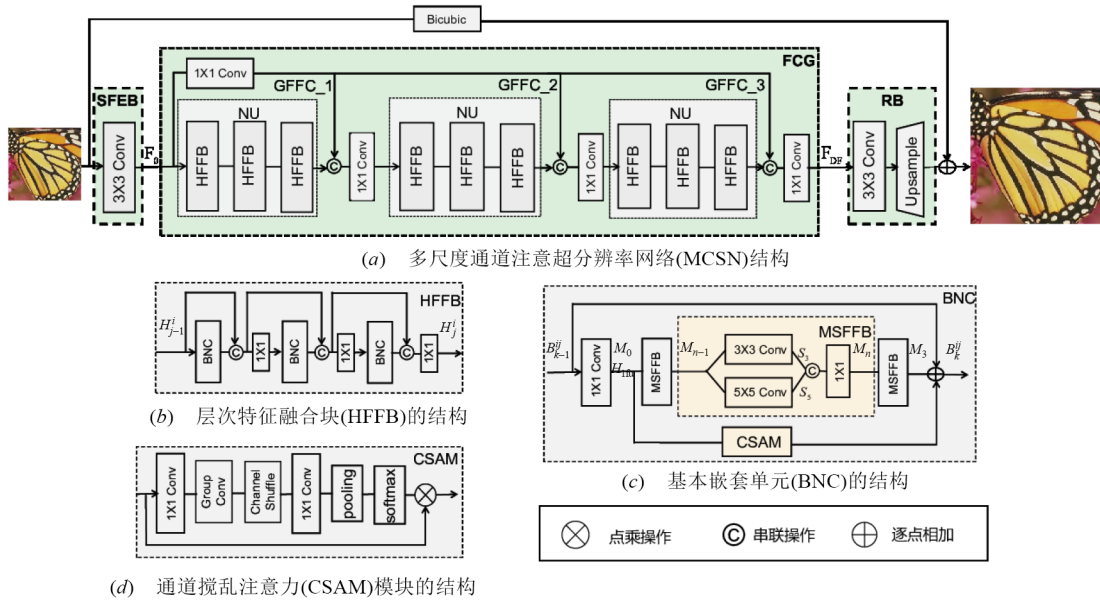


图1 递归网络

3.1 多尺度通道注意超分辨率网络(MCSN)结构

假定 I_{LR} 和 I_{SR} 分别是输入的 LR 图像和输出的 SR 图像。 I_{LR} 首先输入到 SFEB 模块(SFEB 只是一个简单的卷积层),取浅层特征 F_0 为

$$F_0 = G_{\text{ext}}(I_{LR}) \quad (1)$$

其中, $G_{\text{ext}}(\cdot)$ 表示 3×3 卷积。 F_0 输入到 FCG 模块,进行深层特征的提取和全局特征的融合,即

$$F_{DF} = G_{\text{fcg}}(F_0) \quad (2)$$

其中, $G_{\text{fcg}}(\cdot)$ 是 FCG 模块函数; F_{DF} 是其输出的特征。 F_{DF} 再输入到 RB 模块,重建输出的高分辨率图像 I_{SR} 为

$$I_{SR} = G_{\text{up}}(F_{DF}) + G_{\text{bic}}(I_{LR}) \quad (3)$$

其中, $G_{\text{bic}}(\cdot)$ 是双三次插值函数; $G_{\text{up}}(\cdot)$ 是重建模块的函数。把原输入的低分辨率图像 I_{LR} 进行双三次上采样,叠加到 $G_{\text{bic}}(\cdot)$ 的输出,获得重建的高分辨率图像 I_{SR} 。在基于卷积神经网络的单图像超分辨率方法中,重建模块大都采用文献[7, 17, 22]的方法(Conv-Pixel Shuffle-Conv)。为了减少参数量,本文去掉了上述方法中的最后一个 3×3 卷积层。

3.2 特征串联组

如图 1(a) 所示, FCG 主要包括 3 个嵌套块(NU)和 3 个全局特征融合连接(GFFC)。1 个嵌套块包含 3 个层次特征融合块(HFFB); 1 个 HFFB 主要包含 3 个基本嵌套单元(BNC)[参见图 1(b)]; 1 个 BNC 主要包含 3 个多尺度特征融合块(MSFFB)和 1 个通道搅乱注意力(CSAM)模块[参见图 1(c)]; 同一 BNC 内的 3 个 MSFFB 之间参数共享。

3.2.1 全局特征融合连接(GFFC)

浅层特征包含大量低频信息。为充分利用这些低

频信息,全局特征融合连接将浅层特征融合到网络的每一个阶段。不同于一般的全局特征融合(如逐点相加或串联操作),使用逐点卷积对浅层特征实现跨通道交互和信息融合,以实现特征信息的流动,减少图像信息的丢失,提高浅层特征的利用率。GFFC 对 F_0 执行逐点卷积(即 1×1 的卷积),输出特征级联到 FCG 模块内 3 个 NU 模块的输出特征之中,即

$$F_{\text{gc}} = f_{1 \times 1}(F_0, W_1) \quad (4)$$

其中, F_0 表示浅层特征; $f_{1 \times 1}(\cdot)$ 和 W_1 分别表示逐点卷积和相应权重。 F_{gc} 最后融合到网络的每一阶段。

3.2.2 嵌套块(NU)

如图 1 所示, 3 个层次特征融合块以递归的方式组成一个嵌套块(NU)。本文的嵌套块类似于残差中残差结构(Residual In Residual, RIR)^[23], 差别在于: 没有采用跳跃连接方式, 而是模块嵌套模块的方式, 不仅有多层次特征融合模块, 也有多尺度特征提取模块。这种方式不仅能增加网络深度, 还有利于提取更丰富的深层特征。可用下式表示:

$$H_i^0 = F_0 \quad (5)$$

$$H_i^3 = G_{\text{nu}}(H_i^0) = R_i^j(\cdots(R_i^1(R_i^0(H_i^{j-1})))) \quad (6)$$

$$H_i^j = R_i^j(H_i^{j-1}) \quad (7)$$

其中, H_i^0 和 H_i^3 ($i=1, 2, 3$) 分别表示第 i 个嵌套块的输入和输出; $G_{\text{nu}}(\cdot)$ 和 $R_i^j(\cdot)$ 分别表示嵌套块和层次特征融合块。在式(3)~式(5)中, R_i^j ($j=1, 2, 3$) 表示第 i 个嵌套块中的第 j 个层次特征融合块; H_i^j 和 H_i^{j-1} 表示相应的输入和输出。此外, 本文将第一个嵌套块的输入 H_i^0 初始化为 F_0 。然后, 将嵌套块和全局特征融合连接的输出融合在一起形成一个新的输入, 传递给下一个嵌套

块,即

$$H_i^0 = F_{1 \times 1}([H_{i-1}^3, F_{gc}], W_2) \quad (8)$$

其中, $[H_{i-1}^3, F_{gc}]$ 表示将嵌套块和全局特征融合连接的输出级联在一起; $F_{1 \times 1}(\cdot)$ 和 W_2 分别表示 1×1 的卷积和相应的权重. 式(8)的作用是将特征通道数目压缩到和嵌套块的输入通道数一样. 最终的深层特征 F_{DF} 可以表示为

$$F_{DF} = F_{1 \times 1}([H_3^3, F_{gc}], W_2) \quad (9)$$

3.2.3 层次特征融合块(HFFB)

一个层次特征融合块(HFFB)包括3个不同的基本嵌套单元和压缩单元. 层次特征融合块内部融合不同层次的多样化特征,是一种局部与全局特征的融合,能有效提高特征使用率,减少图像信息的丢失. 首先,将第一个基本嵌套单元的输出与输入合并在一起,之后通过压缩单元将特征通道数压缩到与基本嵌套单元的输入一样. 接着将压缩的结果输入第二个基本嵌套单元中,它的输出与之前合并后的结果再次合并在一起,经过压缩单元后的输出作为第三个基本嵌套单元的输入. 最后将第三个基本嵌套单元的输出与之前第二个合并后的结果再次合并在一起. 这一过程可表示为

$$C_0^j = B_0^j = H_i^{j-1} \quad (10)$$

$$B_k^j = G_{ij}^k(C_{k-1}^j) \quad (11)$$

$$C_k^j = F_{1 \times 1}([B_0^j, B_1^j, \dots, B_k^j], W_3) \quad (12)$$

$$H_i^j = C_3^j \quad (13)$$

3.2.4 基本嵌套单元(BNC)

基本嵌套单元(BNC)主要包括局部融合单元、多尺度特征融合块和通道搅乱注意力模块3部分. 对提取的多尺度特征和经过通道搅乱注意力模块选择后的特征以及局部残差连接的输出3部分特征,通过逐点相加操作实现多样化特征融合,有助于增强特征表达能力,提供丰富的细节信息. 此外,对输入到多尺度特征融合块和通道搅乱注意力模块的特征,局部融合单元首先使用一个 1×1 卷积使特征通道之间进行交互,便于特征信息的融合. 如式(14)所示:

$$H_{ifu} = F_{1 \times 1}(B_{k-1}^j, W_4) \quad (14)$$

$$B_k^j = G_{bnc}^k(C_{k-1}^j) = G_{csam}(H_{ifu}) + G_{msfb}(H_{ifu}) + B_{k-1}^j \quad (15)$$

其中, $G_{bnc}(\cdot)$, $G_{csam}(\cdot)$ 和 $G_{msfb}(\cdot)$ 分别表示基本嵌套单元、多尺度特征融合块和通道搅乱注意力模块. W_4 表示局部融合单元的权重; B_{k-1}^j 和 C_{k-1}^j 如式(11)和式(12)所示.

3.2.5 多尺度特征融合块(MSFFB)

许多轻量级 SISR 方法仅使用一个 3×3 卷积提取特征, MSRN 方法首次使用多尺度卷积提取特征. MSRN 中的多尺度残差块(MultiScale Residual Block, MSRB)包含多个卷积核(5×5 和 3×3), 可提取更丰富的多尺

度特征,参数也显著更多;MSRN 中的两次级联(concat)运算也有较多的特征信息冗余. 本文提出了多尺度特征融合块(MSFFB),每个支路分别使用一个 5×5 和 3×3 的卷积,提取多种尺度的特征. MSFFB 采用递归的结构增加了网络深度,大幅度减少了参数量. 每个 MSFFB 的 2 个支路提取的特征,仅级联(concat)一次,减少了特征信息冗余. 多尺度特征融合块主要包括 3 个尺寸不同的卷积核,其中, 5×5 和 3×3 的卷积核用于提取特征, 1×1 的卷积核用于融合这 2 种不同尺度的特征. 这里 1×1 的卷积不仅能实现特征之间的融合,也能压缩输出,保证输出与输入之间具有相同的通道数,便于 3 个共享多尺度特征融合块实现递归. 这一过程可用下式表示:

$$S_3 = \sigma(F_{3 \times 3}(M_{n-1}, W_5)) \quad (16)$$

$$S_5 = \sigma(F_{5 \times 5}(M_{n-1}, W_6)) \quad (17)$$

$$M_n = F_{1 \times 1}([S_3, S_5], W_7) \quad (18)$$

其中, σ 表示 ReLU 激活函数; $F_{5 \times 5}(\cdot)$, $F_{3 \times 3}(\cdot)$ 和 $F_{1 \times 1}(\cdot)$ 分别表示 3 个不同感受野大小的卷积核; W_5 , W_6 和 W_7 表示相应的权重; S_3 和 S_5 表示提取到的不同尺度的特征; 此外, M_{n-1} 和 M_n ($n=1, 2, 3$) 表示第 n 个共享多尺度特征融合块的输入和输出. 3 个共享多尺度特征融合块的最终输出 M_3 可表示为

$$M_0 = H_{ifu} \quad (19)$$

$$M_3 = G_{msfb}(G_{msfb}(G_{msfb}(M_0))) \quad (20)$$

其中, M_0 表示第一个多尺度特征融合块的输入,将其初始化为 H_{ifu} .

3.2.6 通道搅乱注意力模块(CSAM)

近年来,许多基于 CNN 的 SISR 方法都忽略了特征通道之间的相互依赖性,阻碍了网络的表示能力,受文献[19, 20, 24]等方法的启发,结合通道搅乱^[24]和注意力机制,本文提出了通道搅乱注意力模块(CSAM). 通道搅乱运算能够有效促进通道之间信息的流动,而注意力机制能够增强网络对特征的选择能力,进而提高网络的表示能力. 如图 2 所示,通道搅乱注意力模块主要由 3 部分组成: 2 个 1×1 之间的瓶颈块、全局平均池化和 softmax 函数.

首先, H_{ifu} 经过第一个 1×1 卷积层压缩通道减少参数量. 然后经过一个组卷积对通道分组,分完组的特征通道经过通道搅乱操作,实现通道之间的信息流动. 最后经过第二个 1×1 卷积层将输出通道数恢复到输入通道数. 这个过程如下式所示:

$$X = F_{r2c}(F_{cf}(F_{gc}(F_{c2r}(H_{ifu}, W_c), W_g)), W_r) \quad (21)$$

其中, $F_{cf}(\cdot)$ 表示通道搅乱操作; $F_{r2c}(\cdot)$, $F_{c2r}(\cdot)$ 和 $F_{gc}(\cdot)$ 分别表示 1×1 卷积和逐点组卷积; W_r , W_c 和 W_g 是其相应权重. 然后,假定 $X = [x_1, x_2, \dots, x_c, \dots, x_C] \in \mathbb{R}^{H \times W \times C}$, 通过全局平均池化操作,得到维度为 C 的通道描述子

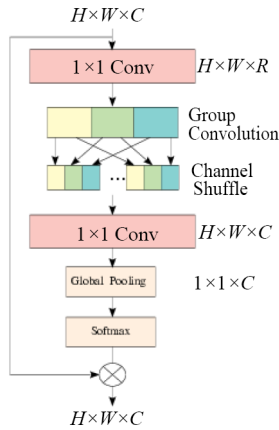


图2 通道搅乱注意力模块结构

(Channel Descriptor)^[19], z 中第 c 个通道描述子可表示为

$$z_c = G_{\text{gap}}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (22)$$

其中, $G_{\text{gap}}(\cdot)$ 表示全局平均池化, 通道描述子 z 可以看作对整个图像信息的一个集中表示^[19], 因为平均操作将图像中每个像素都考虑在内。最后, 使用 softmax 函数计算这些通道描述子之间的相关权重, 即

$$w = \tau(z) = \frac{e_i^z}{\sum_{i=1}^C e_i^z} \quad (23)$$

其中, τ 表示 softmax 函数, $w \in \mathbb{R}^C$ 表示维度为 C 的相关权重。将这些相关权重乘以相对应的通道, 就得到了通道搅乱注意力模块的最终输出, 即

$$\hat{x} = w_c \cdot x_c \quad (24)$$

其中, w_c 和 x_c 表示第 c 个通道的相关权重和相应特征通道。

4 实验细节

4.1 实验设置

本文训练和验证图像采用高质量 DIV2K^[25] 数据集。前 800 个图像用于训练, 第 801~805 个图像用于验证。共同使用的 5 个基准测试集是 Set5^[26], Set14^[27], B100^[28], Urban100^[29] 和 Manga109^[30]。原 HR 训练图像, 进行双三次 (Bicubic, BI) 下采样, 获得配对的 LR 和 HR 训练图像对。类似其他方法, 通过 90°, 180° 和 270° 旋转, 以及水平翻转, 增强了训练图像。生成的超分辨率图像转换到 YCbCr 空间, 在亮度 (Y) 通道上计算峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和结构相似性 (Structural Similarity Index, SSIM)^[31]。

本文采用与文献 [7] 相同的参数设置, 训练图像裁剪为 48×48 的图像片, 每批次随机输入 16 个图像片。使用 Adam^[32] 优化器且设置 $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$, L_1 损失函数、通道数 (滤波器个数) $C = 64$, 初始学习率设

置为 $L_r = 10^{-4}$, 每 200 个迭代周期 (epoch) 学习率减半。表 1、表 2 分别表示多尺度特征融合块 (MSFFB) 和通道搅乱注意力模块 (CSAM) 的具体参数设置。本文方法在训练 ×2 模型时, 总共运行 1 000 个迭代周期 (epoch), 用时 3 天左右, 每个迭代周期用时 5 min。在 Set14 数据集 ×4 放大倍数下, 单幅图像上实现超分辨率重建平均用时 0.12 s。模型使用 Pytorch^[33] 框架, 在 NVIDIA 2080Ti GPU 上进行训练与测试。

表 1 通道搅乱注意力模块 (CSAM) 参数设置

层名称	输入尺寸	输出尺寸	卷积核大小
1×1 卷积层	$H \times W \times 64$	$H \times W \times 48$	1×1
组卷积层	$H \times W \times 48$	$H \times W \times 48$	1×1
通道搅乱层	$H \times W \times 48$	$H \times W \times 48$	—
1×1 卷积层	$H \times W \times 48$	$H \times W \times 64$	1×1
平均池化层	$1 \times 1 \times 64$	$1 \times 1 \times 64$	—

表 2 多尺度特征融合块 (MSFFB) 参数设置

层名称	输入尺寸	输出尺寸	卷积核大小
3×3 卷积层	$H \times W \times 64$	$H \times W \times 64$	3×3
Relu 激活	$H \times W \times 64$	$H \times W \times 64$	5×5
5×5 卷积层	$H \times W \times 64$	$H \times W \times 64$	—
Relu 激活	$H \times W \times 64$	$H \times W \times 64$	1×1
合并 (concat)	$H \times W \times 64$	$H \times W \times 128$	—
1×1 卷积层	$H \times W \times 128$	$H \times W \times 64$	1×1

4.2 模型分析

为了验证多尺度特征融合块、通道搅乱注意力模块和全局特征融合连接的有效性, 本文设计了 8 组对比实验, 每个实验使用相同的压缩率^[24] ($R = 48$)、分组数 ($G = 3$)、多尺度特征融合块 ($M = 3$) 和层次特征融合块 ($H = 3$) 数目。如表 3 所示, 不包含这 3 个模块的网络称为基准网络, 记为 MCSN_G0_C0_M0, 在 DIV2K 验证集上 PSNR 值只有 32.14 dB。

在基准网络上分别增加 3 个模块之一, 记为 MCSN_G1_C0_M0, MCSN_G0_C1_M0 和 MCSN_G0_C0_M1。如表 3 中第 3 列至第 5 列所示, PSNR 值均比基准网络高。只包含多尺度特征融合块的网络比基准网络 PSNR 值高 0.06 dB, 说明多尺度特征融合块是有效的。

表 3 多尺度特征融合块、通道搅乱注意力模块和全局特征融合连接在 200 个迭代周期、DIV2K 验证集 ×4 上的 PSNR 值

	×	√	×	×	√	√	×	√
GFFC	×	√	×	×	√	√	×	√
CSAM	×	×	√	×	√	×	√	√
MSFFB	×	×	×	√	×	√	√	√
PSNR	32.14	32.17	32.15	32.20	32.22	32.26	32.27	32.32

接下来, 在基准网络上增加其中 2 个模块, 分别记为 MCSN_G1_C1_M0, MCSN_G1_C0_M1 和 MCSN_G0_C1_M1, 如表 3 中第 6 列至第 8 列所示, PSNR 值逐渐增

高,且包含任意 2 个模块的网络都比仅包含 1 个模块的网络效果好. 其中 MCSN_G0_C0_M1 网络和 MCSN_G0_C1_M1 网络相比,即在多增加一个通道搅乱注意力模块的情况下,PSNR 值高出 0.06 dB. 通道搅乱注意力模块能充分利用特征通道之间的相关性,促进特征之间信息的流动,进而增强网络对特征的选择能力,提升网络的性能.

最后,包含全部 3 个模块的网络记为 MCSN_G1_C1_M1,如表 3 最后一列所示,PSNR 值进一步提高,达到了 32.32 dB. 在包含通道搅乱注意力模块和多尺度特征融合块网络的基础上,再增加全局特征融合连接,PSNR 值增加了 0.06 dB (对比表 3 中最后 2 列),由此可见,全局特征融合连接在网络中能提高浅层特征的利用率,减少图像低频信息的丢失,改进模型的性能.

4.3 递归性能分析

为了验证递归块数对网络性能的影响,本文做了 3 组对比实验. 选取 DIV2K 验证集中 5 个图像,200 个迭代周期和 $\times 4$ 放大倍数. 假定 H 表示层次特征融合块

数目, M 表示多尺度特征融合块数目,且 H 和 M 同时增大,即 H 和 M 同时等于 1,2 或 3,如表 4 所示. $H=M=1$ 表示没有使用递归层次特征融合块和多尺度特征融合块,其平均 PSNR 性能最低(32.14 dB). H 和 M 增大,即使用递归方式,其平均 PSNR 性能也增加(表 4 中第 3 和 4 行). 递归数增多,网络越深,平均 PSNR 性能也越好.

表 4 DIV2K 验证集中递归块对网络性能影响

H 和 M 数目	PSNR/dB
1	32.14
2	32.25
3	32.32

4.4 R, G, H 和 M 的设置

本节将分析一些超参数的设置. 假定 R 和 G 分别表示通道搅乱注意力模块中的压缩率和分组数, H 表示层次特征融合块数目, M 表示多尺度特征融合块数目. 超参数的设置对网络的影响,如图 3 所示.

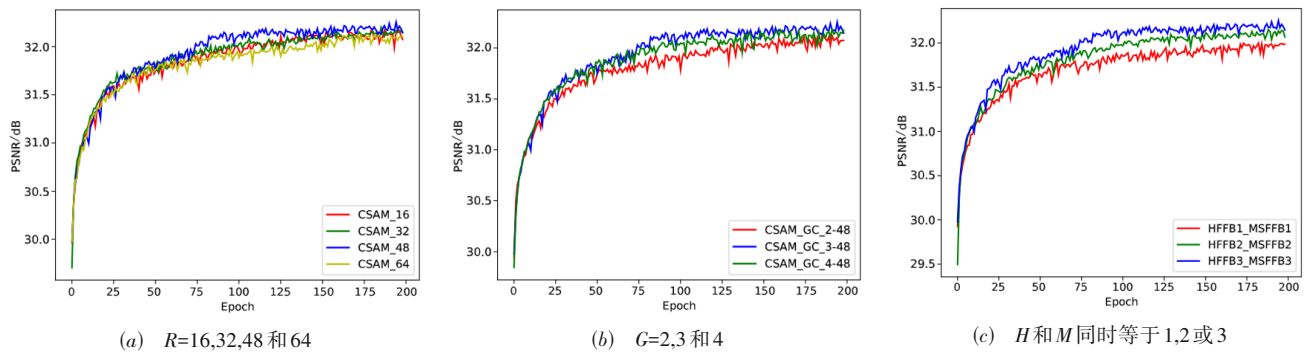


图 3 在 200 个迭代周期、DIV2K 验证集 $\times 4$ 放大倍数下,超参数 R, G, H 和 M 的设置对网络的影响

为验证通道搅乱注意力模块中第一个 1×1 卷积压缩率 R ,本文做了 4 组实验,分别设 $R=16, 32, 48$ 和 64 ,如图 3(a) 所示. 可以看出, $R=48$ 时的曲线位于其他 3 条曲线之上,即 PSNR 更高. 合适的压缩率 R 能减少参数量,过滤掉一些无效信息,为后续通道搅乱操作提供更为有效的信息.

对于通道搅乱注意力模块中分组卷积的组数 G ,本文做了 3 组实验. 分别设 $G=2, 3$ 和 4 ,如图 3(b) 所示,可以看出, $G=3$ 时网络性能表现最好. 一般情况下,分组越多,参数量越少,但是会影响网络性能,分组数 $G=3$ 是参数量和性能之间较好的平衡点.

最后,网络中的递归结构在层次特征融合块和多尺度特征融合块之间共享参数,在不增加参数量的情

况下,可增加网络的深度. 为验证递归块个数 H 和 M ,本文做了 3 组对比实验,分别假定 H 和 M 同时增大,即 H 和 M 同时等于 1,2 或 3,如图 3(c) 所示. 可以看出, $H=M=3$ 的时效果最好. 考虑到计算量的约束,没有试验更大的 H 和 M .

4.5 与先进方法比较

为了验证本文 MCSN 网络的有效性,同其他先进方法进行了定量客观比较和主观视觉效果比较. 其他代表性的方法包括: SRCNN^[9], VDSR^[10], SRMDNF^[34], DRRN^[12], MemNet^[13], CARN^[14], CBPN^[15], IMDN^[16], MSRN^[17], DRCN^[21], FSRCNN^[35], LapSRN^[36], LFFN^[37], IMDN^[16], SAN^[38], EBPN^[39], OISR^[40] 和 ERN^[41]. 放大倍数分别为 $\times 2, \times 3, \times 4$ 和 $\times 8$,如表 5 和图 4~9 所示.

表 5 各种 SISr 方法的平均 PSNR 值与 SSIM 值

放大 倍数	模型	参数量	Set5	Set14	B100	Urban100	Manga109	
			PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	
×2	SRCNN	57K	36.66/0.952 4	32.42/0.906 3	31.36/0.887 9	29.50/0.894 6	35.74/0.966 1	
	FSRCNN	12K	37.00/0.955 8	32.63/0.908 8	31.53/0.892 0	29.88/0.902 0	36.67/0.969 4	
	VDSR	665K	37.53/0.958 7	33.03/0.912 4	31.90/0.896 0	30.76/0.914 0	37.22/0.972 9	
	DRCN	1774K	37.63/0.958 8	33.04/0.911 8	31.85/0.894 2	30.75/0.913 3	37.63/0.972 3	
	LapSRN	813K	37.52/0.959 0	33.08/0.913 0	31.80/0.895 0	30.41/0.910 0	37.27/0.974 0	
	DRRN	297K	37.74/0.959 1	33.23/0.913 6	32.05/0.897 3	31.23/0.918 8	37.92/0.976 0	
	MemNet	677K	37.78/0.959 7	33.28/0.914 2	32.08/0.897 8	31.31/0.919 5	37.72/0.974 0	
	SRMDNF	1513K	37.79/0.960 0	33.32/0.915 0	32.05/0.898 0	31.33/0.920 0	38.07/0.976 1	
	CARN	1592K	37.76/0.959 0	33.52/0.916 6	32.09/0.897 8	31.92/0.925 6	38.36/0.976 5	
	CBPN	1036K	37.90/0.959 0	33.60/0.917 1	32.17/0.898 9	32.14/0.927 9	—	
	IMDN	694K	38.00/0.960 5	33.63/0.917 7	32.19/0.899 6	32.17/0.928 3	38.88/0.977 4	
	LFFN	1522K	37.95/0.959 7	—	32.20/0.899 4	32.39/0.929 9	38.73/0.976 5	
	MSRN	5930K	38.08/0.960 7	33.70/0.918 6	<u>32.23/0.900 2</u>	32.29/0.930 3	38.69/0.977 2	
	MCSN	1561K	<u>38.09/0.960 8</u>	<u>33.76/0.918 9</u>	32.22/0.900 1	<u>32.41/0.930 8</u>	<u>38.89/0.977 5</u>	
	MCSN+	1561K	38.14/0.961 0	33.79/0.919 5	32.26/0.900 6	32.58/0.932 1	39.07/0.977 9	
×3	SRCNN	57K	32.75/0.909 0	29.28/0.820 9	28.41/0.786 3	26.24/0.798 9	30.59/0.910 7	
	FSRCNN	12K	33.16/0.914 0	29.43/0.824 2	28.53/0.791 0	26.43/0.808 0	30.98/0.921 2	
	VDSR	665K	33.66/0.921 3	29.77/0.831 4	28.82/0.797 6	27.14/0.827 9	32.01/0.931 0	
	DRCN	1774K	33.82/0.922 6	29.76/0.831 1	28.80/0.796 3	27.15/0.827 6	32.31/0.932 8	
	DRRN	297K	34.03/0.924 4	29.96/0.834 9	28.95/0.800 4	27.53/0.837 8	32.74/0.939 0	
	MemNet	677K	34.09/0.924 8	30.00/0.835 0	28.96/0.800 1	27.56/0.837 6	32.51/0.936 9	
	SRMDNF	1530K	34.12/0.925 0	30.04/0.837 0	28.97/0.803 0	27.57/0.840 0	33.00/0.940 3	
	CARN	1592K	34.29/0.925 5	30.29/0.840 7	29.06/0.803 4	27.38/0.840 4	33.50/0.944 0	
	IMDN	703K	34.36/0.927 0	30.32/0.841 7	29.09/0.804 6	28.17/0.851 9	33.61/0.944 5	
	LFFN	1534K	34.43/0.926 6	—	29.13/0.805 9	28.34/0.855 8	33.65/0.944 5	
	MSRN	6114K	34.46/0.927 8	30.41/0.843 7	<u>29.15/0.806 4</u>	28.33/0.856 1	33.67/0.945 6	
	MCSN	1569K	<u>34.53/0.928 3</u>	<u>30.43/0.843 7</u>	29.15/0.806 2	28.45/0.858 1	33.95/0.946 6	
	MCSN+	1569K	34.59/0.928 7	30.49/0.844 8	29.20/0.807 2	28.61/0.860 4	34.18/0.947 8	
	×4	SRCNN	57K	30.48/0.862 8	27.49/0.750 3	26.90/0.710 1	24.52/0.722 1	27.66/0.850 5
		FSRCNN	12K	30.71/0.865 7	27.59/0.753 5	26.98/0.715 0	24.62/0.728 0	27.90/0.851 7
VDSR		665K	31.35/0.883 8	28.01/0.767 4	27.29/0.725 1	25.18/0.752 4	28.83/0.880 9	
DRCN		1774K	31.53/0.885 4	28.02/0.767 0	27.23/0.723 3	25.14/0.751 0	28.98/0.881 6	
LapSRN		813K	31.54/0.885 0	28.19/0.772 0	27.32/0.728 0	25.21/0.756 0	29.09/0.884 5	
DRRN		297K	31.68/0.888 8	28.21/0.772 0	27.38/0.728 4	25.44/0.763 8	29.46/0.896 0	
MemNet		677K	31.74/0.889 3	28.26/0.772 3	27.40/0.728 1	25.50/0.763 0	29.42/0.894 2	
SRMDNF		1555K	31.96/0.893 0	28.35/0.777 0	27.49/0.734 0	25.68/0.773 0	30.09/0.902 4	
CARN		1592K	32.13/0.893 7	28.60/0.780 6	27.58/0.734 9	26.07/0.783 7	30.47/0.908 4	
CBPN		1197K	32.21/0.894 4	28.63/0.781 3	27.58/0.735 6	26.14/0.786 9	—	
IMDN		715K	32.21/0.894 8	28.58/0.781 1	27.56/0.735 3	26.04/0.783 8	30.45/0.907 5	
LFFN		1531K	32.15/0.894 5	—	27.52/0.737 7	26.24/0.790 2	30.66/0.909 9	
MSRN		6078K	32.26/0.896 0	28.63/0.783 6	27.61/0.738 0	26.22/0.791 1	30.57/0.910 3	
MCSN		1581K	<u>32.43/0.897 2</u>	<u>28.73/0.784 7</u>	<u>27.65/0.738 8</u>	<u>26.35/0.793 9</u>	<u>30.81/0.912 5</u>	
MCSN+		1581K	32.50/0.898 3	28.81/0.786 1	27.70/0.739 8	26.50/0.796 7	31.06/0.914 8	

续表

放大倍数	模型	参数量	Set5	Set14	B100	Urban100	Manga109
			PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
×8	SRCNN	57K	25.34/0.647 1	23.86/0.544 3	24.14/0.504 3	21.29/0.513 3	22.46/0.660 6
	FSRCNN	12K	25.42/0.644 0	23.94/0.548 2	24.21/0.511 2	21.32/0.509 0	22.39/0.635 7
	VDSR	655K	25.73/0.674 3	23.20/0.511 0	24.34/0.516 9	21.48/0.528 9	22.73/0.668 8
	DRCN	1774K	25.93/0.674 3	24.27/0.551 0	24.49/0.516 8	21.71/0.528 9	23.20/0.668 6
	LapSRN	813K	26.15/0.702 8	24.45/0.579 2	24.54/0.529 3	21.81/0.555 5	23.39/0.706 8
	MSRN	6226K	26.59/0.725 4	24.88/0.596 1	24.70/0.541 0	22.37/0.597 7	24.28/0.751 7
	MCSN	1664K	<u>27.03/0.775 0</u>	<u>24.95/0.640 6</u>	<u>24.79/0.596 5</u>	<u>22.41/0.614 8</u>	<u>24.52/0.777 1</u>
	MCSN+	1664K	<u>27.12/0.778 6</u>	<u>25.05/0.643 2</u>	<u>24.84/0.598 2</u>	<u>22.55/0.620 2</u>	<u>24.68/0.782 3</u>

注:最好结果与次好结果分别用加粗和下划线标出.

(1)定量客观比较. 由表5可知,本文MCSN方法比其他方法PSNR值和SSIM值显著要高. 与本文方法参数量相当的CARN和SRMDNF相比,在Set5数据集×2,×3,×4放大倍数下,本文的PSNR值都超过了0.2 dB. 图4是各个SISR方法×4模型,在Set5数据集上的结果,本文MCSN方法具有更好的PSNR性能和参数量平衡. 与MSRN相比,本文的PSNR值高约0.2 dB,而参数量只有它的1/4. 虽然一些方法比MCSN参数少,但是,有很显著的性能差距. 使用自集合(self-ensemble)^[42]的方法,记为MCSN+,可以进一步提高性能,如表5所示. 本文的PSNR值和SSIM值超过了其他所有比较的SISR方法.

在Set5×4放大倍数下,本文MCSN方法与其他轻量级模型相比,在运行时间、参数量和性能上取得了更好的平衡. MCSN与比较的方法分别多用时0.027 s,0.022 s和0.021 s,但是平均PSNR性能分别提高了0.30 dB,0.17 dB和0.22 dB(表6). 此外,与当前最先进的SISR方法相比(表7),在参数量和性能之间,本文方法似乎取得了更好的平衡. 与SAN方法相比,本文方法的平均PSNR值略低,而参数量只有其1/10(参见表中最后2列). ERN和EBPN方法相比,本文的参数量只相当其1/6,但是,平均PSNR性能相当(EBPN略高0.06 dB,ERN略低0.04 dB).

表6 Set5×4放大倍数下网络性能与运行时间对比

模型	CARN	MSRN	IMDN	MCSN
时间/s	0.040	0.045	0.046	0.067
PSNR/dB	32.13	32.26	32.21	32.43

表7 Set5×4放大倍数下最先进方法性能与参数量对比

模型	SAN	EBPN	ERN	MCSN
参数量(M)	15.7	8.18	9.53	1.58
PSNR/dB	32.64	32.79	32.39	32.43

(2)主观效果比较. 图10是本文方法与其他先进方法在×4放大倍数下重建的超分辨率(Super Resolution, SR)图像,本文方法比其他方法重建结果显著更好. 以

图像062为例,Bicubic,SRCNN,CARN,MSRN等方法恢复出来的玻璃窗格子模糊不清,条纹产生畸变,图像失真严重,本文方法恢复的图像纹理清晰、细节丰富,更加接近真实图像. 图像063情况相似. 再以B100测试集中的图148026为例,其他方法恢复的图像木板条纹严重模糊,本文方法恢复的木板条纹接近真实图像. 图5和图6分别是在×2和×3放大倍数下重建的SR图像,本文方法能重建出高质量的图像,是清晰可视的. 以Set14中“barbara”图像为例,在×2放大倍数下,CARN和

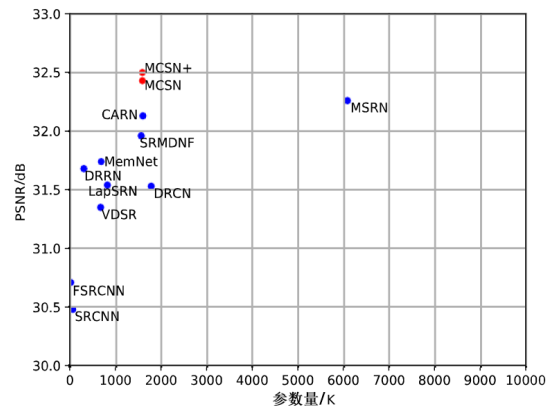


图4 各种SISR方法的×4模型在Set5数据集上的平均PSNR值及参数量



图5 Set14测试集中图像barbara的×2视觉效果比较

MSRN方法恢复出的围巾模糊不清,存在严重失真,本文方法几乎完美恢复了图像(图5). $\times 3$ 放大倍数下,本文方法恢复结果类似(图6). 图7是与当前最先进的、更轻量级网络,在Set14数据集 $\times 4$ 放大倍数的主观视觉效果比较,本文方法更好地恢复出了“with”的轮廓. 图8、图9分别是真实图像“Hilbert”与飞机,在 $\times 4$ 放大倍数下不同方法的重建结果(左下角红色方框是局部放大图). 很容易观察到,本文的重建结果,也显著优于其他2个尖端的方法.



图6 Manga109测试集中图像YumeiroCooking的 $\times 3$ 视觉效果比较



图7 Set14测试集中图像ppt3的 $\times 4$ 视觉效果比较



图8 真实Hilbert图像 $\times 4$ 放大倍数下视觉效果比

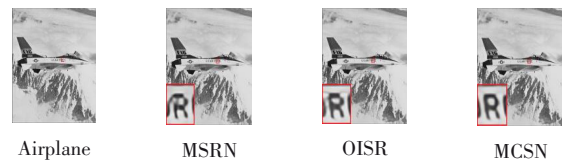


图9 真实飞机图像 $\times 4$ 放大倍数下视觉效果比较

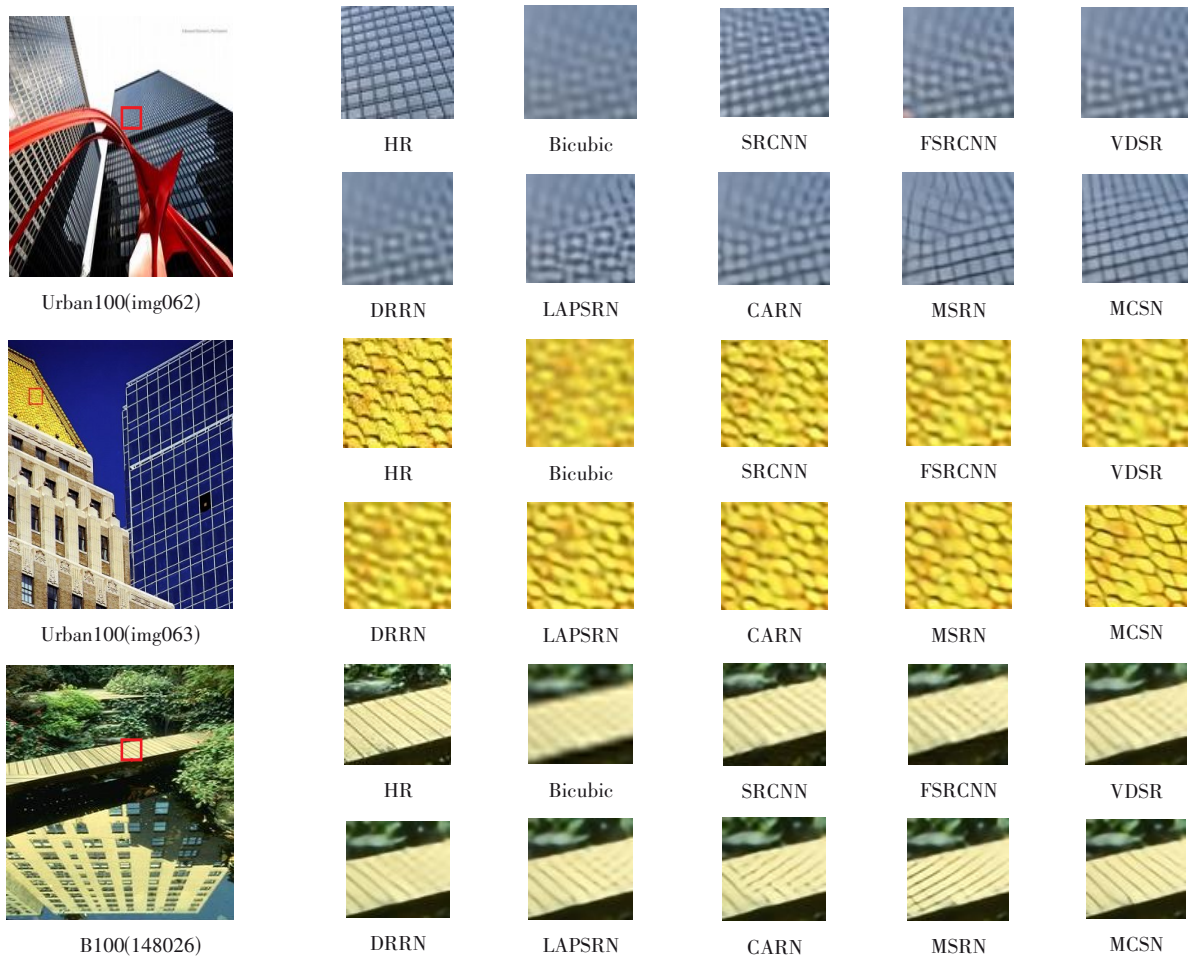


图10 基准测试集Urban100和B100中图像img062,img063和148026的 $\times 4$ 视觉效果比较

5 结论

本文提出了一种轻量级、多尺度通道注意力超分辨率网络,不仅能有效提取各种特征,还能充分利用图像的低频信息,重建出细节丰富的高质量图像. 多尺度特征融合块使用尺寸不同的卷积核提取多种尺度特征,确保提取的特征信息更加多样化;通道搅乱注意力模块有效地促进通道之间信息的流动,增强网络的特征选择能力;全局特征融连接充分利用浅层特征包含的低频信息,减少图像信息的丢失. 实验证明:本文方法具有良好的性能,无论是客观 PSNR/SSIM 值,还是主观视觉效果,均比当前最先进的方法要好,且参数量较少,在资源受限的移动等设备上具有潜在应用价值.

参考文献

- [1] THORNTON M W, ATKINSON P M, HOLLAND D A. Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping[J]. *International Journal of Remote Sensing*, 2006, 27(3): 473-491.
- [2] ZOU W W W, YUEN P C. Very low resolution face recognition problem[J]. *IEEE Transactions on Image Processing*, 2012, 21(1): 327-340.
- [3] SHI W Z, CABALLERO J, LEDIG C, et al. Cardiac image super-resolution with global correspondence using multi-atlas patchmatch[C]//*International Conference on Medical Image Computing and Computer-Assisted Intervention*. Berlin, Heidelberg: Springer, 2013: 9-16.
- [4] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 105-114.
- [5] WANG F, JIANG M Q, QIAN C, et al. Residual attention network for image classification[C]//*2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017: 6450-6458.
- [6] LI Z, YANG J L, LIU Z, et al. Feedback network for image super-resolution[C]//*2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR)*. Long Beach: IEEE, 2019: 3862-3871.
- [7] 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*. Honolulu: IEEE, 2017: 1132-1140.
- [8] ZHANG Y L, TIAN Y P, KONG Y, et al. Residual dense network for image super-resolution[C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 2472-2481.
- [9] DONG C, LOY C C, HE K M, et al. Learning a Deep Convolutional Network for Image Super-Resolution[C]//*European Conference on Computer Vision*. Cham: Springer, 2014: 184-199.
- [10] 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*. Las Vegas: IEEE, 2016: 1646-1654.
- [11] 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*. Las Vegas: IEEE, 2016: 770-778.
- [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*. Honolulu: IEEE, 2017: 2790-2798.
- [13] TAI Y, YANG J, LIU X M, et al. MemNet: A persistent memory network for image restoration[C]//*2017 IEEE International Conference on Computer Vision*. Honolulu: IEEE, 2017: 4549-4557.
- [14] AHN N, KANG B, SOHN K A. Fast, accurate, and lightweight super-resolution with cascading residual network [C]//*European Conference on Computer Vision*. Cham: Springer International Publishing, 2018: 256-272.
- [15] ZHU F Y, ZHAO Q J. Efficient single image super-resolution via hybrid residual feature learning with compact back-projection network[C]//*2019 IEEE/CVF International Conference on Computer Vision Workshop*. Seoul: IEEE, 2019: 2453-2460.
- [16] 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*. New York: ACM, 2019: 2024-2032.
- [17] LI J C, FANG F M, MEI K F, et al. Multi-scale residual network for image super-resolution[C]//*European Conference on Computer Vision*. Cham: Springer International Publishing, 2018: 527-542.
- [18] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, 20(11): 1254-1259.
- [19] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//*2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 7132-7141.
- [20] ZHANG Y L, LI K P, LI K, et al. Image super-resolution using very deep residual channel attention networks[C]//*European Conference on Computer Vision*. Cham: Springer, 2018: 294-310.
- [21] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution[C]//*2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 1637-1645.
- [22] 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. Las Vegas: IEEE, 2016: 1874-1883.
- [23] ZHANG Y L, LI K P, LI K, et al. Image super-resolution using very deep residual channel attention networks[C]// European Conference on Computer Vision. Cham: Springer, 2018: 294-310.
- [24] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 6848-6856.
- [25] 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. Honolulu: IEEE, 2017: 1122-1131.
- [26] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[C]//Proceedings British Machine Vision Conference. Surrey: BMVA Press, 2012: 135.1-135.10.
- [27] ZEYDE R, ELAD M, PROTTER M. On Single Image Scale-up Using Sparse-Representations[M]//Curves and Surfaces. Berlin, Heidelberg: Springer, 2012: 711-730.
- [28] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]//Proceedings Eighth IEEE International Conference on Computer Vision. Vancouver: IEEE, 2001: 416-423.
- [29] 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. Boston: IEEE, 2015: 5197-5206.
- [30] 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.
- [31] 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: A Publication of the IEEE Signal Processing Society, 2004, 13(4): 600-612.
- [32] KINGMA D P, BA J. Adam: A method for stochastic optimization[EB/OL]. (2014)[2020]. <https://arxiv.org/abs/1412.6980>.
- [33] PASZKE A, GROSS S, CHINTALA S, et al. Automatic differentiation in pytorch[OL]. (2017)[2020]. <https://openreview.net/forum?id=BJJsrmfCZ>.
- [34] ZHANG K, ZUO W M, ZHANG L. Learning a single convolutional super-resolution network for multiple degradations[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 3262-3271.
- [35] DONG C, LOY C C, TANG X O. Accelerating the super-resolution convolutional neural network[C]//European Conference on Computer Vision. Cham: Springer, 2016: 391-407.
- [36] 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. Honolulu: IEEE, 2017: 5835-5843.
- [37] YANG W M, WANG W, ZHANG X C, et al. Lightweight feature fusion network for single image super-resolution[J]. IEEE Signal Processing Letters, 2019, 26(4): 538-542.
- [38] DAI T, CAI J R, ZHANG Y B, et al. Second-order attention network for single image super-resolution[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019: 11057-11066.
- [39] QIU Y J, WANG R X, TAO D P, et al. Embedded block residual network: A recursive restoration model for single-image super-resolution[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 4179-4188.
- [40] HE X Y, MO Z T, WANG P S, et al. ODE-inspired network design for single image super-resolution[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach: IEEE, 2019: 1732-1741.
- [41] LAN R S, SUN L, LIU Z B, et al. Cascading and enhanced residual networks for accurate single-image super-resolution[J]. IEEE Transactions on Cybernetics, 2020, 51(1): 115-125.
- [42] TIMOFTE R, ROTHE R, GOOL L VAN. Seven ways to improve example-based single image super resolution[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1865-1873.

作者简介



周登文 男, 1965年出生, 湖北人. 华北电力大学控制与计算机工程学院教授. 主要研究方向为神经网络和深度学习在图像处理和计算机视觉中的应用, 特别是图像超分辨率技术.

E-mail: zdw@ncepu.edu.cn



李文斌 男, 1996年出生, 甘肃人. 2017年在上海电力学院获得学士学位, 现为华北电力大学硕士研究生. 主要研究方向为计算机视觉和深度学习.

E-mail: 1182227108@ncepu.edu.cn