

# 多跳连接残差注意网络的图像超分辨率重建

刘遵雄 朱成佳 黄 稷 蔡体健

华东交通大学信息工程学院 南昌 330013

**摘 要** 随着卷积神经网络深度的不断增加,深度卷积神经网络的训练会变得更加困难。此外,在图像超分辨率中,低分辨率图像的通道特征和输入通常在不同的通道中被平等对待,这就导致了卷积神经网络的表征能力被弱化。为了解决这些问题,提出了一种多跳连接残差注意网络,该网络利用多跳连接中的残差(Residual in Multi-skip Connection, RIMC),构造了具有多个残差组的深度网络。每个残差组包含了一定数量的短跳连接和多跳连接。在 RIMC 的基础上,主网络被允许穿过多跳连接来绕过丰富的低频信息,同时高频信息也可以被主网络集中地学习。另外,考虑到通道和空间维度的相互依赖关系,提出了注意机制块(Attention Mechanism Block, AMBlock)来关注信息的位置,并自适应地调整通道特征尺度,其中通道注意机制和空间注意机制被应用在这种方式中。实验结果表明,该网络可以更好地恢复图像细节,获得更高的图像质量和网络性能。

**关键词:** 图像超分辨率; 注意机制块; 残差网络; 多跳连接中的残差; 跳连接

中图分类号 TP391.41

## Image Super-resolution by Residual Attention Network with Multi-skip Connection

LIU Zun-xiong, ZHU Cheng-jia, HUANG Ji and CAI Ti-jian

School of Information Engineering, East China Jiaotong University, Nanchang 330013, China

**Abstract** Deep convolutional neural networks (Deep CNNs) are difficult to train as they become deeper. Moreover, in image super-resolution, channel-wise features and inputs of the low-resolution (LR) image are treated equally between different channels, resulting in the deficiency of the representational ability of the CNNs. To resolve these issues, residual attention network with multi-skip Connection (RANMC) is proposed for single-image super resolution (SISR), which employs residual in multi-skip connection (RIMC) structure, then a very deep network is formulated with several residual groups. Each residual group (RG) contains a certain number of short skip connections (SSC) and multi-skip connections (MC). Based on RIMC, rich low-frequency (LF) information is allowed to be bypassed through multi-skip connection, and high-frequency (HF) information is focused on learning by the principal network. Furthermore, considering interdependencies in channel and spatial dimension, attention mechanism block (AMBlock) is proposed to focus on the location of the information and adaptively readjust channel-wise features, where the spatial attention (SA) mechanism and channel attention (CA) mechanism are taken in the approach. Experiments indicate that RANMC can not only recover image details better, but also obtain higher image quality and network performance.

**Keywords** Image super-resolution, Attention mechanism block, Residual network, Residual in multi-skip connection, Skip connection

## 1 引言

最近,单幅图像超分辨率(Single-image Super-resolution, SISR<sup>[1]</sup>)在计算机视觉领域<sup>[2-5]</sup>很受欢迎,如在监视成像<sup>[6]</sup>、图像生成<sup>[7]</sup>、医学成像<sup>[8]</sup>、安全等方面。SISR的目的是把低分辨率(Low-Resolution, LR)视觉图像转换成高分辨率(High-Resolution, HR)视觉图像,但是在把单幅图像从 LR 空间映射到 HR 空间的过程中存在多个解,这就造成了一个病态的问题,并且当更大的上采样因子(up-sampling factors)

被应用在这一过程中时,这个过程将很难使用更大的上采样因子。为了解决这个问题,很多图像超分辨率(Super-Resolution, SR)技术被应用在 SISR 中,如基于学习的方法<sup>[9-15]</sup>、基于重构的方法<sup>[16]</sup>和基于插值的方法<sup>[17]</sup>。

目前,在图像 SR 中,基于深度卷积神经网络的技术相比传统的技术有更大的进步,同时在解决许多 SISR 问题时获得了非常明显的效果。Chao 等<sup>[18]</sup>设计了三层卷积网络 SRCNN 模型,将其首先引入到图像 SR 中,用于解决图像 SR 问题。在 SISR 中,通过使用跳连接和梯度裁剪,可以很成功地

到稿日期:2020-10-08 返修日期:2021-01-31

基金项目:国家自然科学基金(61861017);江西省青年科学基金(20181BAB211013)

The work was supported by the National Natural Science Foundation of China(61861017) and Science Foundation for Young Scientist of Jiangxi Province, China(20181BAB211013).

通信作者:刘遵雄(zhujsq@163.com)

训练更深的网络模型,同时获得了比 SRCNN 效果更好的结果。大量的研究工作表明,网络模型的深度对各种计算机视觉任务模型的性能起到了非常重要的作用,特别是在提出了带有残差块的残差网络(ResNet<sup>[19]</sup>)以后,这种作用更加明显,其中 ResNet 网络模型的深度达到了 1000 层。在这之后,有许多基于卷积网络的图像 SR 技术都引入了这种有效的残差学习策略。其中,Lim 等<sup>[20]</sup>提出了更深的 MDSR 模型和带有 EDSR 块的更宽的 EDSR 模型,EDSR 和 MDSR 中带有残差块的层数约为 165。

另一方面,在处理丰富的低频(Low-Frequency, LF)和高频(High-Frequency, HF)信息时,LR 图像的输入和通道特征被平等对待,导致卷积神经网络的表征能力减弱,这样就会造成网络模型不能充分地利用网络的 LF 和 HF 信息。如果将丰富的 LF 信息从 LR 图像的输入直接转换成 HR 图像的输出,那么该网络模型就不需要执行太多复杂的计算操作。但是,由于以往基于卷积神经网络的技术在处理各种信息(如 LF 和 HF 信息)方面缺乏灵活性,因此对于 LR 空间中的网络来说,直接提取各个卷积层的输出也是很困难和不切实际的。

基于以上问题,RANMC 被提出并应用于 SISR。RANMC 充分利用来自原始 LR 的所有分层特征,使模型性能更好。为了能够在更深的网络模型中简化训练(如超过 500 层),RIMC 被提出,同时残差组(Residual Group, RG)作为 RIMC 的基本模块。在每个 RG 中,短跳连接被用来堆叠几个简单的残差块,同时在短跳连接和多跳连接中以粗级的方式进行残差学习,这样 RIMC 结构的提出对网络模型起到了很大的改善作用。更进一步地,考虑了通道与空间维度的特

征相关性,提出了 AMBlock,其是由通道(Channel Attention, CA)机制和空间(Spatial Attention, SA)机制组成的,分别对每个通道上的特征进行自适应地缩放并关注信息的位置,这样网络性能就得到了极大的提高。由于 RANMC 模型采用了跳连接策略和注意机制块,因此能够很好地利用 LF 和 HF 信息,并且增强了其鲁棒性。如图 1 所示,在 RANMC 模型中,视觉 SR 的结果明显得到很大的改善。本文的贡献如下:

(1)针对 SISR 问题,在 RANMC 中提出了 RG 这一结构。该结构可以训练更深的网络模型,使网络模型获得更好的效果和性能。

(2)在 RANMC 网络模型中提出了短跳连接和 RIMC 结构。同时,在 RIMC 的结构中,提出跳连接残差注意块(Skip Connection Residual Attention Block, SCRAB)。这些结构使该网络模型可以训练更深的网络。

(3)在 RANMC 中引入 AMBlock,其是由 CA 机制和 SA 机制组成的。由于模型跨特征通道之间的相互依赖关系,AMBlock 可以帮助模型自适应地重新调节特征,提取出更关键、更重要的信息,并且还可以专注于信息的位置,相比以往使用注意机制的网络模型,AMBlock 显示出更大的优越性。

(4)提出了简单的分层特征融合和图像重建方法,以提高网络的泛化能力。在分层特征融合中采用了  $1 \times 1$  卷积运算,避免引入许多参数。在图像重建中采用亚像素卷积运算,降低了网络的复杂性,避免了过多的人为因素干扰。这样,任何一种放大因子(upscaling factors)都可以很容易地被用于分层特征融合和图像重建中。

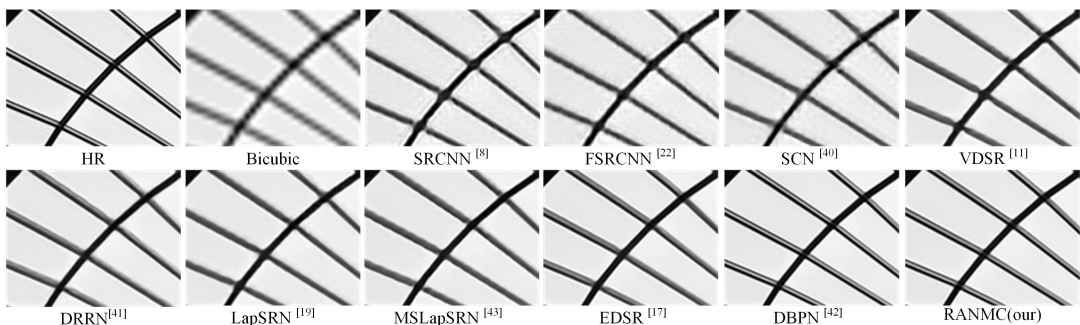


图 1 在“Urban100 的 img\_072”中具有双三插值降解的  $4x$  超分辨图像的视觉图像效果图

Fig. 1 Effect of visual images for  $4x$  super-resolution with bicubic degradation on “img\_072” of Urban100

## 2 相关工作

### 2.1 跳连接

在 SISR 中,许多模型由于没有充分利用 LF 和 HF 信息,造成大量信息丢失,性能不佳,为了解决上述问题,最近提出了很多方法。例如,提出了层间短跳连接和长跳连接的密集跳连接<sup>[21]</sup>(dense skip connections)网络来充分地利用 LF 和 HF 信息。另外,在层与层之间把旁路应用于 ResNets 来有效地训练更深的模型。所有这些网络模型都体现了一个重要的思想:对于图像 SR 来说,在层与层之间建立大量的跳连接很重要。但是,在这些方法中,许多网络模型很难训练超过 1000 层的网络。

### 2.2 注意机制

每个通道都被视为特征图中的一个特征检测器。CA 的主要任务是关注输入图像中什么是有意义的。众所周知,许多通道对信息的传输没有意义,因此利用 CA 对这些通道进行了筛选以获得最优的特征。总的来说,它可以作为将现有处理资源分配给信息最丰富层面的指南。近年来,为了使卷积神经网络在各种模型中的性能得到改善,一些实验将注意力处理应用于各种模型中。例如:Hu 等<sup>[22]</sup>设计了压缩提取块(squeeze-and-excitation block),用于建立模型通道间的关系,从而获得了显著的图像分类性能改进;Wang 等<sup>[23]</sup>引入了残差注意网络结构,将注意力集中在主要信息上。但是,许多方法都没有研究低水平视觉任务中注意力的表现,而且许多

方法也不关注输入图像中什么是有意义的。此外,正如在前面的SR方法中看到的那样,CA机制很少用于图像的SR,CA机制和SA机制共同应用于图像SR中的模型更少,这些方法也很少关注信息的位置。

### 2.3 基于CNN的方法

在之前的工作中,Dong等设计了用于图像超分辨率的SRCNN,并取得了很好的效果,其中SRCNN的上采样利用传统的插值方法,如双三次插值(Bicubic, BI)方法,来获得目标空间的尺寸。在SRCNN的基础上,VDSR<sup>[24]</sup>和DRCN<sup>[25]</sup>得到了很大的改进,但也带来了许多缺点,如丢失了一些细节,增加了计算量。为了使SRCNN的训练和测试速度更快,提出了FSRCN<sup>[26]</sup>。但在这些网络模型中,很多方法限制了网络深度,而深度是视觉识别的关键。

为了深入研究这些方法,本文提出了一个RANMC模型(见图2)。图2给出了精简的RANMC模型,可供读者深入理解。本文第3节介绍了提出的网络结构;第4节给出了实

验结果和与更好结果的可视化比较;第5节讨论了未来工作,以及本文模型需要改进的地方;第6节总结全文。

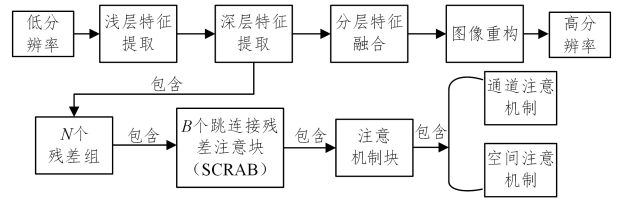


图2 被精简的RANMC模型的流程图

Fig. 2 Flow chart of the simplified RANMC model

## 3 RANMC方法

如图3所示,RANMC模型主要由4个部分组成:浅层特征提取(shallow feature extraction)、深层特征提取(deep feature extraction)、分层融合结构(hierarchical fusion structure)以及图像重构模块(image reconstruction structure)。

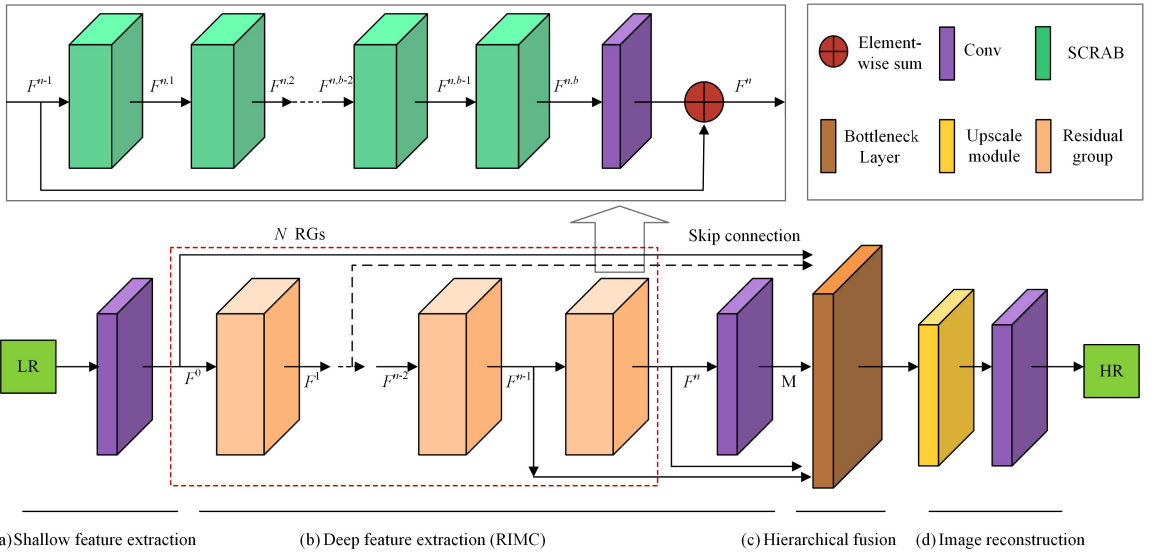


图3 RANMC模型的完整网络结构

Fig. 3 Complete network architecture of RANMC model

本文的主要目的是从LR图像 $I^{LR}$ 中重构HR图像 $I^{HR}$ 。对 $I^{HR}$ 图像执行双三次插值操作,获得 $I^{LR}$ 图像。图像只在Y通道上训练,然后转换到YCbCr颜色空间。本文使用一个张量来描述LR的图像,这个张量的大小被设置为 $W \times H \times C$ 。 $I^{LR}$ 和 $I^{SR}$ 分别表示RANMC的输入和输出,并且它们是已知的。Y通道表示为 $C$  ( $C=1$ )且这个放大因子被表示为 $r$ 。如文献[20,27]中的研究,下文将对RANMC的每个模块进行详细描述。

### 3.1 网络结构

本节只使用了一个卷积层从LR输入中提取浅层特征(Shallow Features, SF) $F^0$ 。

$$F^0 = H^{SF}(I^{LR}) \quad (1)$$

其中,卷积操作表示为 $H^{SF}(\cdot)$ 。然后将所提取的浅层特征 $F^0$ 作为深度特征提取(Deep Feature Extraction, DFE)的输入。为了进行深度特征提取,RANMC网络模型使用了RIMC结构。因此,深度特征提取表示为:

$$F^{DFE} = H^{DFE}(F^0) \quad (2)$$

其中,深度特征提取模块表示为 $H^{DFE}(\cdot)$ 。它是由 $N$ 个残差组(Residual Groups, GRs)组成的(见图3)。每个RG由1个AMBlock组成,在深度特征提取中,跳连接被映射在 $1 \times 1$ 的卷积层上。换句话说,多个跳连接是被同时拼接在一个 $1 \times 1$ 的卷积层上。这里将提取的深度特征 $F^{DFE}$ 作为下一个卷积层的输入。因此,产生的操作如下:

$$F^{BL} = H^{BL}(F^{DFE}) \quad (3)$$

其中,把 $H^{BL}(\cdot)$ 表示为瓶颈层(Bottleneck Layer, BL),瓶颈层由 $1 \times 1$ 的卷积层组成。此外,卷积层的特征 $F^{BL}$ 的作用是用于构建重构层(Reconstruction Layer, RL)。重构层由 $3 \times 3$ 的卷积层和1个上采样层(up-sampling layer)组成,并且上采样层是由ESPCN<sup>[28]</sup>构成,因此重构层的操作表示为:

$$I^{SR} = H^{RL}(F^{BL}) = H^{RANMC}(I^{LR}) \quad (4)$$

其中, $H^{RL}(\cdot)$ 和 $H^{RANMC}(\cdot)$ 分别表示重构层和RANMC的函数。

本文将使用损失函数来优化RANMC,其中MSE函数和 $L_2$ 函数是应用较为广泛的目标优化函数。虽然这些损失函

数可以实现较高的 PSNR/SSIM<sup>[29,30]</sup>,但  $L_2$  优化问题和 MSE 优化解决方案往往产生过于光滑的纹理。另外,当前也有很多损失函数,可是它们的改变非常小。基于上述问题, $L_1$  损失函数被应用在 RANMC 模型中,且效果很好。本文给出的训练集被表示为  $\{I_i^{LR}, I_i^{HR}\}_{i=1}^Z$ ,该训练集包含  $Z$  个 LR 图像和对应的 HR 图像, $L_1$  损失函数的优化如下:

$$L(\theta) = \frac{1}{Z} \sum_{i=1}^Z \| H^{\text{RANMC}}(I_i^{\text{LR}} - I_i^{\text{HR}}) \|_1 \quad (5)$$

其中,RANMC 模型中的参数集表示为  $\theta$ 。为了优化损失函数,在该模型中使用了随机梯度下降算法(stochastic gradient descent algorithm)。由于很多模块的提出与之前有相似之处,本文重点关注 RIMC, AMBlock 和 SCRAB。

### 3.2 多跳连接中的残差(RIMC)

由于 SISR 的 LF 和 HF 信息没有得到充分利用,且 HF 信息并不是集中从原始的 LR 图像中得来,因此在 RANMC 模型中提出了 RIMC 架构(见图 3),其中 RIMC 含有  $N$  个残差组和多跳连接,每一个残差组由 1 个 AMBlock 组成,且 RIMC 可以训练超过 400 层的深度网络。基于 RIMC,非常深的网络模型均可以被训练。

之前的工作已经证明,使用多跳连接和堆积残差块可以重构更深层次的卷积神经网络。但是,在图像 SR 中,用这种方法建立更深层次的网络来训练网络模型并获得良好的性能是非常困难的。为了解决这个问题,提出了一个 RG,在 RANMC 中把它作为深度网络模型的基本模块,该灵感来自于 EDSR<sup>[31]</sup>。第  $n$  组中残差组的方程表示为:

$$F^n = H^n(F^{n-1}) = H^n(H^{n-1}(\dots H^1(F^0))) \quad (6)$$

其中,第  $n$  个残差组的函数表示为  $H^n$ 。针对第  $n$  个残差组来说,其输入和输出分别表示为  $F^n$  和  $F^{n-1}$ 。可以发现,简单地叠加残差组并不能获得好的视觉效果。此外,随着网络深度的增加,虽然语义表达能力增强,但网络的空间表达能力变弱。因此,网络模型几乎不可能获得性能和质量更好的重建图像。在此基础上,提出了一个带有  $1 \times 1$  卷积核的瓶颈层来解决这些问题,并自适应地从这些分层特征中提取有用的信息。此外,许多多跳连接和短跳连接也被应用于 RIMC 结构的残差组中。这个操作表示为:

$$F^{\text{LR}} = \omega * [F^0, F^1, F^2, \dots, M] + b \quad (7)$$

在第一个卷积层中, $F^0$  表示输出, $F^i (i \neq 0)$  是第  $n$  个残差组的输出,同时  $[F^0, F^1, F^2, \dots, M]$  表示拼接操作。 $F^{\text{LR}}$ ,  $\omega$  和  $b$  分别为瓶颈层的输入、权值和偏移量(bias)。

正如前文讨论的那样,在 LR 输入中有丰富的信息,且 SR 的本质就是恢复更多的有用信息,对此本文对残差学习进行了更深入的研究。在每个 RG 中,堆叠  $B$  个 SCRAB。在第  $n$  个残差组中,第  $b$  个 SCRAB 被表示为:

$$\begin{aligned} F^{n,b} &= H^{n,b}(F^{n,b-1}) \\ &= H^{n,b}(H^{n,b-1}(\dots H^{n,1}(F^{n-1})\dots)) \end{aligned} \quad (8)$$

其中,在第  $n$  个残差组中,第  $b$  个 SCRAB 的输入和输出分别表示为  $F^{n,b-1}$  和  $F^{n,b}$ , $H^{n,b}$  表示对应的函数。为了让该模型的主网络关注信息有更加丰富的特征,引入短跳连接(Short Skip Connection, SSC)来获得输出块。

$$F^n = F^{n-1} + W^n F^{n,B} = F^{n-1} + W^n H^{n,B}(H^{n,B-1}(\dots H^{n,1}(F^{n-1})\dots)) \quad (9)$$

其中, $W^n$  是在第  $n$  个 RG 的尾部设置卷积层的权值。基于此操作,该模型可以进一步学习残差信息,同时,在训练中更多的 LF 信息更容易被绕过。

### 3.3 注意机制块(AMBlock)

在本实验中,AMBlock(见图 4)用于重构特征,并且以学习的方式对特征重新进行权重分配,权重高的特征就是要关注的点。如图 4 所示,AMBlock 是由 CA 机制和 SA 机制组成,且可以自适应地重构中间映射。 $\otimes$  表示按元素的方式进行乘积。

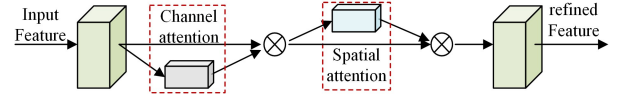


图 4 AMBlock 结构示意图  
Fig. 4 Schematic of AMBlock

(1)通道注意机制。在 RANMC 的模块中,虽然 CA 被视为一个卷积层,但是它在 AMBlock 中起着重要作用,其增加有效的通道权重,减少无效通道权重,因此在此对其进行详细阐述。众所周知,以卷积神经网络为基础的 SR 方法可以平等地对待 LR 通道特征,但在实际应用中却相反。在这种情况下,提出了 CA 机制(见图 5(a)),使 RANMC 网络模型重新分配特征权重,其中 CA 机制利用了特征通道之间的依赖关系。

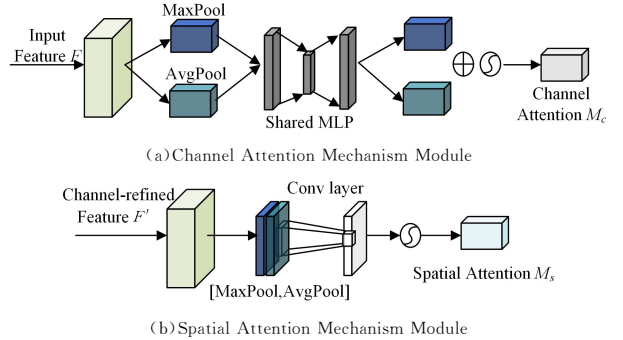


图 5 通道注意和空间注意图

Fig. 5 Diagram of channel attention and spatial attention

为了生成 CA 机制映射,在 AMBlock 中使用了特征的通道间关系。此外,将每个通道作为特征映射中的特征检测器,使通道注意力聚焦于输入图像的主要信息。与此同时,为了提高 CA 的计算效率,本文采用压缩技术对输入的特征映射空间的维度进行压缩,利用平均池化(average-pool)对空间信息进行聚合。基于之前的研究,本文认为最大池化(max-pooling)可以收集到其他的重要线索,通过其他的重要线索可以推断出更好的通道注意力。因此,在实验中,同时将平均池化和最大池化集成到 CA 机制模块中(见图 5(a))。当然,许多经验已经证明,同时使用两种方法,网络的表征能力可以得到极大的提高。

在实验中,使用平均池化和最大池化方法首先对特征映射的空间信息进行了聚合。在这里,两种不同的空间上下文描述符(context descriptors)被表示为平均池和最大池。这样可以通过转换描述符(converting descriptors)得到一个共享的网络(shared network)。基于此共享网络,最大池输出和平均池输出均被应用于通道子模块。空间子模块(Spatial Sub-

module, SS)中使用了类似的两个输出(Similar Two Outputs, STO), SS沿着通道的轴线存放 STO 输出,并最终转发到卷积层。多层感知器被表示为 MLP。基于上述操作,通道注意映射  $M_c \in R^{C \times 1 \times 1}$  就产生了。共享网络由一个多层感知器(Multi-layer Perceptron, MLP)组成,其后跟着一个隐藏层。为了便于计算,隐藏层中激活函数的大小设为  $R^{C/r \times 1 \times 1}$ ,其中  $r$  为缩减比。将共享网络应用到每个描述符(descriptor)之后,使用按元素进行的求和(element-wise summation)函数来合并输出特征向量。具体计算操作如下:

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ = \sigma(W_1(W_0(F_{\text{AVG}})) + W_1(W_0(F_{\text{MAX}}))) \quad (10)$$

其中, sigmoid 函数被表示为  $\sigma$ ,  $W_0 \in R^{C/r \times C}$ ,  $W_1 \in R^{C \times C/r}$ 。值得注意的是,两个输入共享 MLP 权重  $W_0$  和  $W_1$ ,其中  $W_0$  放置在 ReLU 激活函数之后。

(2)空间注意机制。与 CA 机制不同, SA 机制的主要作用是关注特征的位置。为了计算 SA 机制以及产生两个二维映射,即  $F_{\text{avg}}^s \in R^{1 \times H \times W}$  和  $F_{\text{max}}^s \in R^{1 \times H \times W}$ ,两个池化操作被用于聚合特征映射(feature map)的通道信息。这里,首先使用最大池和平均池在通道维度上产生两个不同的特征,然后使用拼接操作合并这两个特征描述,并使用卷积操作生成空间注意机制映射(见图 5 (b))。具体操作如下:

$$M_s(F) = \sigma(f^{3 \times 3}([\text{AvgPool}(F); \text{MaxPool}(F)])) \\ = \sigma(f^{3 \times 3}([F_{\text{avg}}^s; F_{\text{max}}^s])) \quad (11)$$

其中,用  $\sigma$  表示 sigmoid 函数,用  $f^{3 \times 3}$  表示卷积运算,卷积运算的滤波器大小设置为  $3 \times 3$ 。

### 3.4 跳连接残差注意块

如上文所述, RG 和多跳连接专注于 LR 特征的主要信息。此外,由于受到文献中的工作的激励,本文使用 AMBlock 来形成 SCRAB(见图 6)。图 6 中,不同颜色的方块代表不同的操作,在图片的右边给出了具体的描述。第  $n$  个残差组中的第  $b$  个残差块有如下操作:

$$F^{n,b} = F^{n,b-1} + R^{n,b}(X^{n,b}) \cdot X^{n,b} \quad (12)$$

其中, SCRAB 的输入和输出分别表示为  $F^{n,b}$  和  $F^{n,b-1}$ ,且残差  $X^{n,b}$  由 SCRAB 从输入中学习得来。 $R^{n,b}$  是 AMBlock 的函数, AMBlock 由 SA 机制和 CA 机制组成。这里,残差成分由两个叠加的卷积层构成。

$$X^{n,b} = W_2^{n,b} \delta(W_1^{n,b} F^{n,b-1}) \quad (13)$$

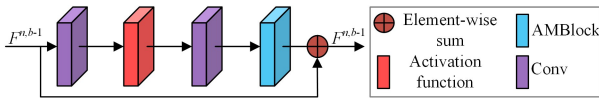


图 6 跳连接残差注意块图(SCRAB)(电子版为彩色)

Fig. 6 Diagram of SCRAB

在 SCRAB 中,用  $W_1^{n,b}$  和  $W_2^{n,b}$  表示两个叠加的卷积层的权重。在此基础上,根据 SCRAB 和 RIMC 结构,提出了一种新且深的 RANMC 模型来获得更高的图像超分辨率和更好的性能。

### 3.5 图像重构

以往的许多 SR 方法主要学习 LR 和 HR 图像之间的映射,通过双三次插值将 LR 图像上采样到与 HR 图像相同的维度。然而,这些方法存在着许多问题,例如它会带来大量的冗余信息,最终导致计算复杂度的增加。此外,几乎没有发现

一个 SR 模型可以被迁移到任何比例放大的因子上,并且只对网络结构做了微小的调整。其中,许多网络都是一个混合的比例因子(fixed up-scaling factor)(如  $\times 4$ ),而没有提供迁移到其他比例因子的具体指令。受这种情况的影响,许多图像 SR 的重建模型使用 Pixel-Shuffle<sup>[32]</sup> 和反卷积(deconvolutional)层。因此,本文也决定考虑使用 Pixel-Shuffle 和反卷积层作为模型的上采样部分。在此操作中,使用比例因子 2 作为模型重构的基础因子。但是,当放大因子(如  $\times 8$ )继续保持增加时,网络模型将变得更深,并且更深的网络也会带来不确定的训练问题。另外,奇数放大因子不适合这些方法,同时希望放大因子缓慢增长(如  $\times 2, \times 3, \times 4, \times 8$  的形式),而不是呈指数增长。

基于上述情况,本文提出了一种新的重构模块(见图 3(d)),该重构模块的结构简单、快捷、灵活。通过比较 Pixel-Shuffle 与反卷积层之间的差异,最终选择 Pixel-Shuffle 作为模型的重建部分。这将帮助 RANMC 模型在微小调整后适应任何比例因子,且在最终的测试中,该结构在不同比例的放大因子下表现出了很好的性能。

## 4 实验

### 4.1 数据集

在实验中, RANMC 模型的训练集选择了具有高质量图像的新 DIV2K<sup>[33]</sup> 数据集。在测试 RANMC 网络模型方面, RANMC 模型选取了 5 个标准基准数据集,如 Set5<sup>[34]</sup>, Set14<sup>[35]</sup>, B100<sup>[36]</sup>, Urban100<sup>[37]</sup> 和 Manga109<sup>[38]</sup>。在之前工作的基础上,使用了 PSNR 和 SSIM 对 YCbCr 颜色空间的 Y(亮度(luminance))通道上的训练和测试结果进行评估。在训练和测试 RANMC 模型的过程中,逐步增大因子(如  $\times 2, \times 3, \times 4, \times 8$ )。

### 4.2 实施细节

在 SCRAB 中,采用 3 种方法来增加训练数据集中的数据,其分别是缩放(scaling)、旋转(rotation)和翻转(flipping)。在 RANMC 模型的训练批量(raining batch)中,随机抽取 16 个大小为  $64 \times 64$  的低分辨率块(low-resolution patches),并将其反向传播的 1000 次迭代作为一轮(an epoch)。ADAM 优化器<sup>[39]</sup>(ADAM optimizer)被用于 RANMC 模型的训练中,其中把学习率设置为  $l_r = 10^{-4}$ 。同时,这些超参数也被用于 ADAM 优化器中,如  $\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 10^{-8}$ 。在最终的网络模型中,使用了 10 个 SCRAB,并且每个 SCRAB 的输出有 64 个特征映射,其中每个瓶颈层的输出也使用了 64 个特征映射,这里卷积层的大小被设置为  $1 \times 1$ 。在 RANMC 模型中,使用 Pytorch 框架进行训练,且 RANMC 模型的训练使用了 NVIDIA Geforce rtx2080 Ti。另外,最重要的是 RANMC 模型中不使用训练技巧,但是在相同的条件下,使用训练技巧的训练结果会与其他一流的模型进行比较。

在 RIMC 结构中,残差组数(Residual Groups, RGs)被设置为  $N = 10$ ,在每一个 RG 中, SCRAB 的数量被设置为 SCRAB=10。在 RANMC 的网络模型中,除了瓶颈层(见图 3(c))外,所有的卷积大小均被设为  $3 \times 3$ 。

### 4.3 与最先进的方法进行比较

为了展示 RANMC 模型的更好的性能,将 RANMC 模

型与更先进的基于卷积神经网络的超分辨率方法进行了比较,这些基于卷积神经网络的超分辨率方法包括FSR-CNN,SCN<sup>[40]</sup>,DRRN<sup>[41]</sup>,SRCNN,VDSR,SRMD<sup>[42]</sup>,DBPN<sup>[43]</sup>,MemNet<sup>[44]</sup>,MSLapSRN<sup>[45]</sup>,EDSR,LapSRN<sup>[46]</sup>,NLRN<sup>[47]</sup>,RDN<sup>[48]</sup>。此外,为了使RANMC的模型更有说服力,所有模型都被再次训练(除了EDSR直接使用了原论文的EDSR结果),并且两个常用的权威评价指标PSNR和SSIM被用来评估该模型的结果。在实验中的亮通道(luminance channel)上,所有记录的PSNR/SSIM测量值都

被计算。为了进一步改进RANMC模型,该模型还使用了自集成策略(self-ensemble strategy),并且把自集表示为RANMC+。

(1)采用PSNR/SSIM进行定量分析的结果。SR方法的最终评价结果如表1所列。这些SR方法包括RANMC模型和其他更先进的模型方法。结果表明,RANMC+在任何比例因子的所有测试数据集集中都表现得最好(见图1)。此外,即使RANMC模型不使用自集成策略,也能得到非常好的结果。

表1 通过在数据集上用平均数PSNR/SSIM的值对最先进的不同方法进行定量比较的结果

Table 1 Quantitative comparison of different most advanced methods by using average PSNR/SSIM value on dataset

Method	Scale	Set5	Set14	BSD100	Urban100	Manga109
		PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic	×2	33.66/0.9299	30.24/0.8688	29.56/0.8431	26.88/0.8403	30.80/0.9339
SRCNN	×2	36.66/0.9542	32.45/0.9067	31.36/0.8879	29.50/0.8946	35.60/0.9663
FSRCNN	×2	37.05/0.9560	32.66/0.9090	31.53/0.8920	29.88/0.9020	36.67/0.9710
VDSR	×2	37.53/0.9590	33.05/0.9130	31.90/0.8960	30.77/0.9140	37.22/0.9750
LapSRN	×2	37.52/0.9591	33.08/0.9130	31.08/0.8950	30.41/0.9101	37.27/0.9740
MemNet	×2	37.78/0.9597	33.28/0.9142	32.08/0.8978	31.31/0.9195	37.72/0.9740
EDSR	×2	38.11/0.9602	33.92/0.9195	32.32/0.9013	32.93/0.9351	39.10/0.9973
SRMD	×2	37.79/0.9601	33.32/0.9159	32.05/0.8985	31.33/0.9204	38.07/0.9761
NLRN	×2	38.00/0.9603	33.46/0.9159	32.19/0.8992	31.81/0.9246	—
DBPN	×2	38.09/0.9600	33.85/0.9190	32.27/0.9000	32.55/0.9324	38.89/0.9775
RDN	×2	38.24/0.9614	34.01/0.9212	32.34/0.9017	32.89/0.9353	39.18/0.9780
RANMC(ours)	×2	<u>38.28/0.9616</u>	<u>34.02/0.9224</u>	<u>32.34/0.9020</u>	<u>32.90/0.9353</u>	<u>39.20/0.9790</u>
RANMC+(ours)	×2	<b>38.30/0.9618</b>	<b>34.20/0.9230</b>	<b>32.40/0.9021</b>	<b>33.20/0.9373</b>	<b>39.22/0.9800</b>
Bicubic	×3	30.39/0.8682	27.55/0.7742	27.21/0.7385	24.46/0.7349	26.95/0.8556
SRCNN	×3	32.75/0.9090	29.30/0.8215	28.41/0.7863	26.24/0.7989	30.48/0.9117
FSRCNN	×3	33.18/0.9140	29.37/0.8420	28.53/0.7910	26.43/0.8080	31.10/0.9210
VDSR	×3	33.67/0.9210	29.78/0.8320	28.83/0.7990	27.14/0.8290	32.01/0.9340
LapSRN	×3	33.82/0.9227	29.87/0.8320	28.82/0.7980	27.07/0.8280	32.21/0.9350
MemNet	×3	34.09/0.9248	30.01/0.8350	28.96/0.8001	27.56/0.8376	32.51/0.9369
EDSR	×3	34.65/0.9280	30.52/0.8462	29.25/0.8093	28.80/0.8653	34.17/0.9476
SRMD	×3	34.12/0.9254	30.04/0.8382	28.97/0.8025	27.57/0.8398	33.00/0.9403
NLRG	×3	34.27/0.9266	30.16/0.8374	29.06/0.8026	27.93/0.8453	—
RDN	×3	34.71/0.9296	30.57/0.8468	29.26/0.8093	28.80/0.8653	34.13/0.9484
RANMC(ours)	×3	<u>34.74/0.9296</u>	<u>30.64/0.8468</u>	<u>29.28/0.8095</u>	<u>28.88/0.8662</u>	<u>34.15/0.9485</u>
RANMC+(ours)	×3	<b>34.84/0.9302</b>	<b>30.70/0.8480</b>	<b>29.35/0.8107</b>	<b>29.14/0.8708</b>	<b>34.35/0.9495</b>
Bicubic	×4	28.42/0.8104	26.00/0.7027	25.96/0.6675	23.14/0.6577	24.89/0.7866
SRCNN	×4	30.48/0.8628	27.50/0.7513	26.90/0.7101	24.52/0.7221	27.58/0.8555
FSRCNN	×4	30.72/0.8660	27.61/0.7750	26.98/0.7150	24.62/0.7280	27.90/0.8610
VDSR	×4	31.35/0.8830	28.02/0.7680	27.29/0.0726	25.18/0.7540	28.83/0.8870
LapSRN	×4	31.54/0.8850	28.19/0.7720	27.32/0.7270	25.21/0.7560	29.09/0.8900
MemNet	×4	31.74/0.8893	28.26/0.7723	27.40/0.7281	25.50/0.7630	29.42/0.8942
EDSR	×4	32.46/0.8968	28.80/0.7876	27.71/0.7420	26.64/0.8033	31.02/0.9148
SRMD	×4	31.96/0.8925	28.35/0.7787	27.49/0.7337	25.68/0.7731	30.09/0.9024
NLRG	×4	31.92/0.8916	28.36/0.7745	27.48/0.7346	25.79/0.7729	—
DBPN	×4	32.47/0.8980	28.82/0.7860	27.72/0.7400	26.38/0.7946	30.91/0.9137
RDN	×4	32.47/0.8990	28.81/0.7871	27.72/0.7419	26.61/0.8028	31.00/0.9151
RANMC(ours)	×4	<u>32.51/0.9000</u>	<u>28.83/0.7871</u>	<u>27.73/0.7420</u>	<u>26.61/0.8029</u>	<u>31.01/0.9153</u>
RANMC+(ours)	×4	<b>32.60/0.9003</b>	<b>28.92/0.7880</b>	<b>27.80/0.7430</b>	<b>26.85/0.8072</b>	<b>31.33/0.9170</b>
Bicubic	×8	24.40/0.6580	23.10/0.5660	23.67/0.5480	20.74/0.5160	21.47/0.6500
SRCNN	×8	25.33/0.6900	23.76/0.5910	24.13/0.5660	21.29/0.5440	22.46/0.6950
FSRCNN	×8	20.13/0.5520	19.75/0.4820	24.21/0.5680	21.32/0.5380	22.39/0.6730
SCN	×8	25.59/0.7071	24.02/0.6028	24.30/0.5698	21.52/0.5571	22.68/0.6963
VDSR	×8	25.93/0.7240	24.26/0.6140	24.49/0.5830	21.70/0.5710	23.16/0.7250
LapSRN	×8	26.15/0.7380	24.35/0.6200	24.54/0.5860	21.81/0.5810	23.39/0.7350
MemNet	×8	26.16/0.7414	24.38/0.6199	24.58/0.5842	21.89/0.5825	23.56/0.7387
MSLap	×8	26.34/0.7558	24.57/0.6273	24.65/0.5895	22.06/0.5963	23.90/0.7564
EDSR	×8	26.96/0.7762	24.91/0.6420	24.81/0.5985	22.51/0.6221	24.69/0.7841
DBPN	×8	<u>27.21/0.7840</u>	25.13/0.6480	24.88/0.6010	<u>22.73/0.6312</u>	25.14/0.7987
RANMC(ours)	×8	<u>27.11/0.7800</u>	<u>25.14/0.6485</u>	<u>24.90/0.6011</u>	<u>22.73/0.6313</u>	<u>25.15/0.7990</u>
RANMC+(ours)	×8	<b>27.22/0.7840</b>	<b>25.15/0.6485</b>	<b>24.92/0.6014</b>	<b>22.78/0.6315</b>	<b>25.16/0.7992</b>

注:加粗的数据为最好的结果,带下划线的数据为次好的结果

(2) 视觉结果。图 7 给出了可视化比较的效果。可以看出, RANMC 模型不仅可以恢复图像的许多细节, 而且可以减轻图像的模糊现象。从图像“img\_092”和“253027”中可以看到类似的结果。与此同时, 视觉上的比较也表

明, 具有表征能力的 RANMC 模型可以在 LR 空间中提取更复杂的图像特征。为了使以上分析更有说服力, 图 8 给出了 8 倍(8 $\times$ )的超分辨率视觉效果比较图, 其效果也非常好。

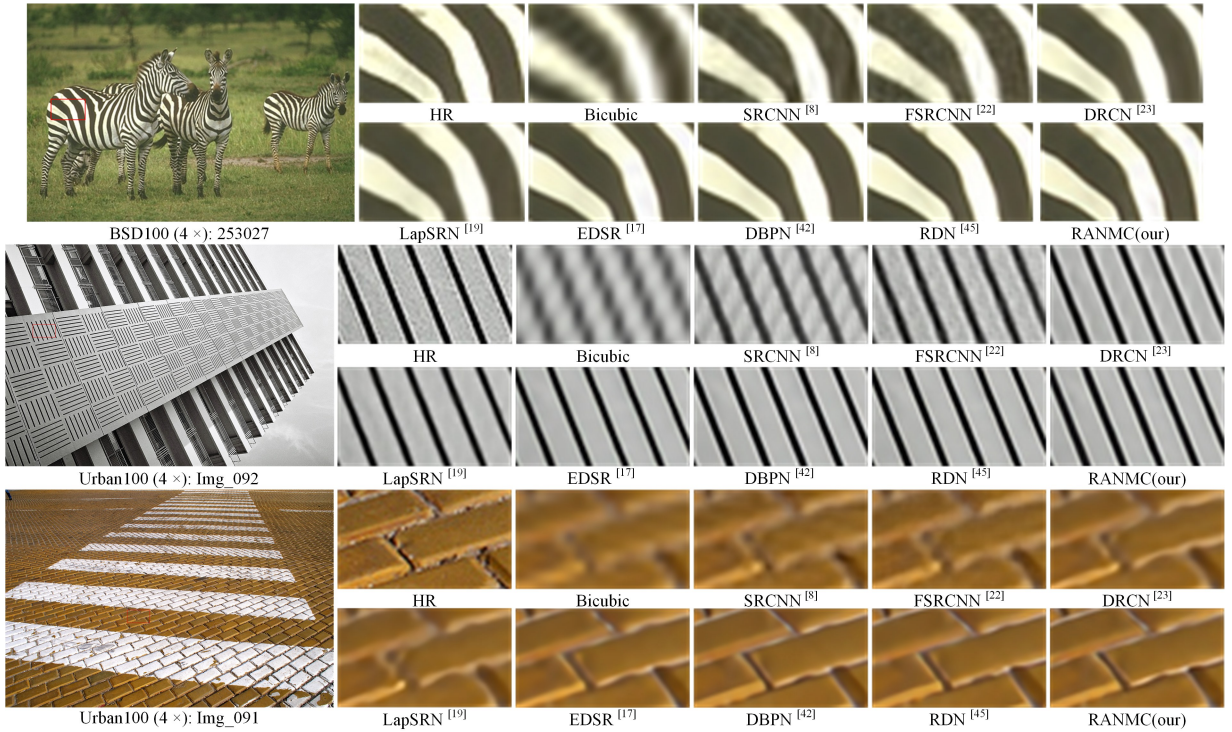


图 7 在 BSD 的“253027”和 Urban100 的“Img\_092”“Img\_091”中, 带有 BI 降解的 4 $x$  超分辨率视觉图像比较图

Fig. 7 Comparison of visual images for 4 $x$  super-resolution with bicubic degradation on “253027” of BSD and “Img\_092”“Img\_091” of Urban100

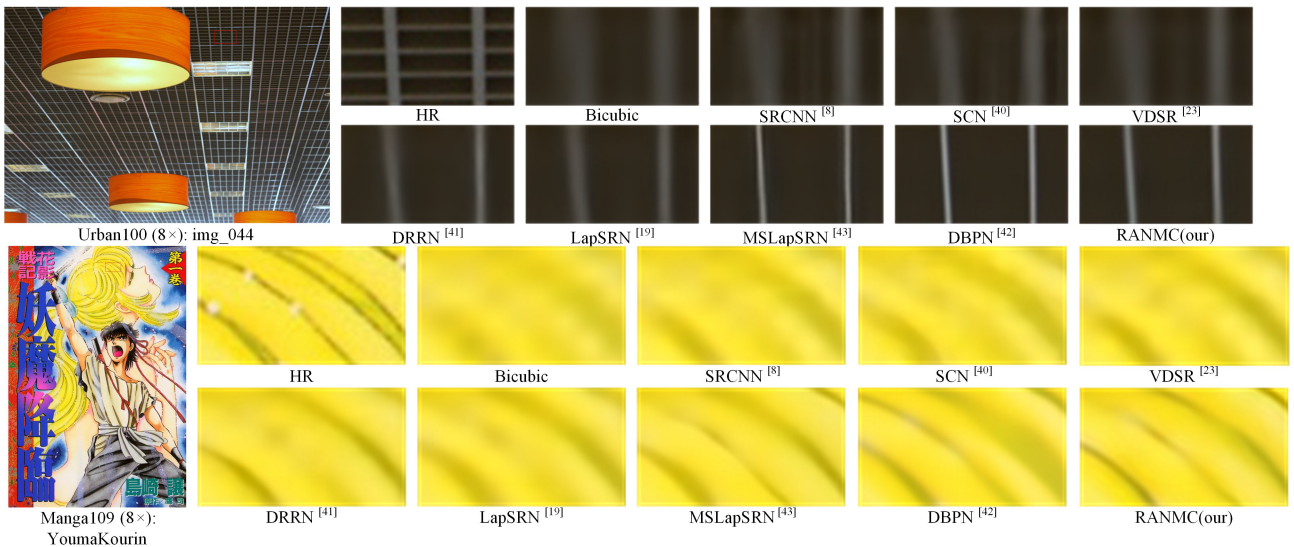


图 8 使用 Urban100 和 Manga109 数据集展示了带有双三次插值降解的 8 $x$  超分辨率的图像效果和模型比较

Fig. 8 Using Urban100 and Manga109 datasets to show the image effects and model comparison of 8 $x$  super-resolution with bicubic degradation

#### 4.4 定性分析

(1) RIMC 和 AMBlock 的益处。本文提出了一个极好的网络模型结构, 即跳连接残差注意块 (SCRAB), 其被应用在 RIMC 模型的结构中。为了更好地展示 SCRAB 的效果, 在 RANMC 模型中删除了跳连接 (Skip Connection, SC), 例如来自更深层次网络的短跳转连接 (Short Skip Connection, SSC) 和多跳转连接 (Multi-skip Connection, MC)。从表 2 中可以

看到, 只要删除 MC 和 SSC, PSNR 值在 Set5(5 $\times$ 2) 上就会很低, 且它与通道注意无关。表 2 是通过使用 Set5(2 $\times$ ) 数据集, 迭代 5 $\times$ 10<sup>4</sup> 次获得的最佳的 PSNR 结果。对 SISR 来说, PSNR 值较低说明简单的叠加残差块无法形成高性能更深层次的网络模型。基于此, 可以看出本实验中提出的 RIMC 结构在 RANMC 模型中起着重要的作用。同时, 对比表 2 中的 PSNR 可以看出, 有 AMBlock 网络的 PSNR 比没

有 AMBlock 网络的 PSNR 好。从表 2 中也可以观察到,即使没有 RIMC,AMBlock 也可以将这一性能从 37.10 dB 提高

到 37.19 dB。最终所有这些比较表明,在 RANMC 模型中使用 AMBlock 是明智的选择。

表 2 各个模块的峰值信噪比结果  
Table 2 PSNR results of various modules

Residual in multi-skip connection (RIMC)	MC	×	✓	×	✓	×	✓	×	✓
	SSC	×	×	✓	✓	×	×	✓	✓
Attention mechanism block (AMBlock)		×	×	×	×	✓	✓	✓	✓
PSNR on Set5 (2×)		37.10	37.62	37.64	37.70	37.19	37.39	37.52	37.58

(2)增加 SCRAB 的数量带来的好处。众所周知,深度网络可以极大地提高网络模型的效果和性能。在本文中,如果想要得到好的结果,增加 SCRAB 的数量应该是最简单的方法。为了验证增加 SCRAB 的数量对测试结果的影响,在模型中设计了一系列测试实验。从图 9 中就可以看到,随着

SCRAB 数量的增加,网络模型的性能不断提升。虽然增加 SCRAB 的数量可以使网络模型的性能更好,但它将大大增加计算量。为了解决计算量的问题,我们决定在 RANMC 模型中使用 10 个 SCRAB。虽然只使用了 10 个 SCRAB,但实验的结果比许多其他网络模型更好。

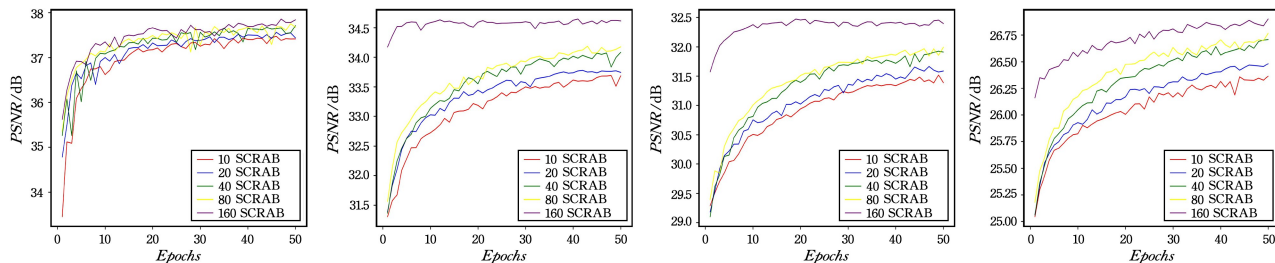


图 9 使用不同数量的跳连接残差注意块(SCRAB)来比较 RANMC 的性能

Fig. 9 Using different amount of skip connection residual attention block (SCRAB) to compare performance of RANMC

#### 4.5 模型参数和性能分析

从表 3(使用 Set(2X)的数据集)中可以看到 RANMC 模型的性能和大小比较的结果,并且还可以看到,通过降低网络模型的性能,NLRG 和 MemNet 获得了更少的参数。相比之下,RANMC 模型不仅减少了参数,而且提高了网络模型的性能。对于这些参数,虽然 RANMC 模型的参数不是最少的,但是其性能是最好的。这意味着 RANMC 模型可以在参数和性能之间实现更好的平衡。

表 3 各个模型不同参数的峰值信噪比结果

Table 3 PSNR results of different parameters of various modules

	EDSR	MemNet	NLRG	DBPN	RDN	RANMC
Parameter	43 M	677 K	330 K	10 M	22.3 M	10.4 M
PSNR	38.11	37.78	38.00	38.09	38.24	38.28

**结束语** 为了获得更真实的重建图像,在单幅图像超分辨中提出了许多训练操作的技巧。例如,文献[17]提出了多尺度混合训练。基于此,本文认为使用一些训练技巧也可以提高 RANMC 模型的性能,但是需要探索一种有效的模型来提高单幅图像超分辨率的性能和效果,而不是使用训练技巧。因此,本文提出了一种基于多跳连接的残差注意力(RANMC)网络。具体地说,多跳连接结构中的残差(RIMC)不仅可以得到具有 SSC 和 MC 的更深层的网络,而且可以充分利用低频信息和高频信息。与此同时,RANMC 模型的主网络通过使用 RIMC 结构和 SCRAB 来加深网络模型,并且为了进一步提高 RANMC 模型的代表能力,本文在 RANMC 模型中还提出了 AMBlock。考虑到特征在通道和空间维度上的相互依赖性,基于注意机制块,通道特征被自适应地调整,同时空间注意机制专注于特征的位置。在 AMBlock 的基础上,

可以有效地提高 RANMC 模型的代表能力,从而获得精确的图像分辨率。本文对带有双三次插值(bicubic, BI)退化的图像进行了大量实验,其结果表明,RANMC 模型在定量和可视化结果上都有较好的效果。虽然 RANMC 模型的性能比已有的许多先进的模型更好,但是该模型也有很多方面需要改进,如计算的复杂度、模型的结构等。在未来的工作中,将进一步提高模型的整体性能,从而实现更轻量化的模型结构。

#### 参考文献

- [1] IRANI M, PELEG S. Improving resolution by image registration [J]. GVGIP: Graphical Models and Image Processing, 1991, 53(3): 231-239.
- [2] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding[C]//BMVC. 2012.
- [3] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(2): 295-307.
- [4] WANG Y, LI A, WANG S Q. Research on the influence of training set automatic target recognition under super-resolution of remote sensing images[J]. Journal of Chongqing University of Technology(Natural Science), 2021, 35(2): 136-143.
- [5] WANG W, HU T, LI X W, et al. Study on Super-resolution Image Reconstruction of Leukocytes [J]. Computer Science, 2021, 48(4): 164-168.
- [6] ZOU W W W, YUEN P C. Very Low Resolution Face Recognition Problem [J]. IEEE Transactions on Image Processing A

- Publication of the IEEE Signal Processing Society, 2012, 21(1): 327-340.
- [7] KARRAS T, AILA T, LAINE S, et al. Progressive Growing of GANs for Improved Quality, Stability, and Variation[J]. arXiv: 1710.10196, 2018.
- [8] CABALLERO J. Cardiac image super-resolution with global correspondence using multi-atlas patchmatch[J]. *Medical Image Computing and Computer-assisted Intervention*, 2013, 16(Pt3): 9-16.
- [9] TIMOFTE R, DE V, GOOL L V. Anchored Neighborhood Regression for Fast Example-Based Super-Resolution[C]// *IEEE International Conference on Computer Vision*. IEEE, 2014.
- [10] TIMOFTE R, VINCENT D S, LUC V G. A+: Adjusted Anchored Neighborhood Regression for Fast Super-Resolution[C]// *Asian Conference on Computer Vision*. Cham: Springer, 2014.
- [11] PELEG T, ELAD M. A Statistical Prediction Model Based on Sparse Representations for Single Image Super-Resolution[J]. *IEEE Transactions on Image Processing*, 2014, 23(6): 2569-2582.
- [12] LAI W S, HUANG J B, AHUJA N, et al. Fast and Accurate Image Super-Resolution with Deep Laplacian Pyramid Networks[C]// *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2017.
- [13] HUANG J B, SINGH A, AHUJA N. Single Image Super-resolution from Transformed Self-Exemplars[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2015.
- [14] LEE C Y, XIE S, GALLAGHER P, et al. Deeply-supervised nets[C]// *Artificial Intelligence and Statistics*. 2015:562-570.
- [15] LI K, WU Z, PENG K C, et al. Tell me where to look: Guided attention inference network[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018: 9215-9223.
- [16] ZHANG K, GAO X, TAO D, et al. Single image super-resolution with non-local means and steering kernel regression[J]. *IEEE Transactions on Image Processing*, 2012, 21(11): 4544-4556.
- [17] ZHANG L, WU X. An edge-guided image interpolation algorithm via directional filtering and data fusion[J]. *IEEE transactions on Image Processing*, 2006, 15(8): 2226-2238.
- [18] CHAO D, CHEN C L, HE K, et al. Learning a Deep Convolutional Network for Image Super-Resolution[C]// *ECCV*. Springer International Publishing, 2014.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:770-778.
- [20] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017:136-144.
- [21] TONG T, LI G, LIU X, et al. Image super-resolution using dense skip connections[C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2017:4799-4807.
- [22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018:7132-7141.
- [23] WANG F, JIANG M, QIAN C, et al. Residual attention network for image classification[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017:3156-3164.
- [24] KIM J, KWON L J, MU L K. Accurate image super-resolution using very deep convolutional networks[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:1646-1654.
- [25] KIM J, KWON L J, MU L K. Deeply-recursive convolutional network for image super-resolution[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016:1637-1645.
- [26] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[C]// *European Conference on Computer Vision*. Cham: Springer, 2016:391-407.
- [27] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C]// *ICML*. 2010.
- [28] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016: 1874-1883.
- [29] YANG C Y, YANG M H. Fast direct super-resolution by simple functions[C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2013:561-568.
- [30] 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.
- [31] 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 (CVPRW)*. IEEE, 2017.
- [32] CHATTERJEE S, CHU W T. Thermal Face Recognition Based on Transformation by Residual U-Net and Pixel Shuffle Upsampling[M]// *Multimedia Modeling*, 2020:679-689.
- [33] TIMOFTE R, AGUSTSSON E, VAN G L, et al. Ntire 2017 challenge on single image super-resolution: Methods and results[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 2017:114-125.
- [34] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-Complexity Single Image Super-Resolution Based on Nonnegative Neighbor Embedding[C]// *BMVC*. 2012.
- [35] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[C]// *International Conference on Curves and Surfaces*. Berlin: Springer, 2010:711-730.
- [36] 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*. 2001:416-423.
- [37] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars[C]// *Proceedings of the*

- IEEE Conference on Computer Vision and Pattern Recognition. 2015:5197-5206.
- [38] 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.
- [39] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. *arXiv*:1412.6980, 2014.
- [40] WANG Z, LIU D, YANG J, et al. Deep networks for image super-resolution with sparse prior[C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2015:370-378.
- [41] TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017:3147-3155.
- [42] ZHANG K, ZUO W, ZHANG L. Learning a single convolutional super-resolution network for multiple degradations[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018:3262-3271.
- [43] HARIS M, SHAKHNAROVICH G, UKITA N. Deep back-projection networks for super-resolution[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018:1664-1673.
- [44] TAI Y, YANG J, LIU X, et al. Memnet: A persistent memory network for image restoration[C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2017:4539-4547.
- [45] LAI W S, HUANG J B, AHUJA N, et al. Fast and accurate image super-resolution with deep laplacian pyramid networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 41(11): 2599-2613.
- [46] LAI W S, HUANG J B, AHUJA N, et al. Deep laplacian pyramid networks for fast and accurate super-resolution[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017:624-632.
- [47] LIU D, WEN B, FAN Y, et al. Non-local recurrent network for image restoration[C]// *Advances in Neural Information Processing Systems*. 2018:1673-1682.
- [48] ZHANG Y, TIAN Y, KONG Y, et al. Residual dense network for image super-resolution[C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018: 2472-2481.



**LIU Zun-xiong**, born in 1967, Ph.D, professor. His main research interests include deep learning, computer vision, image processing, financial data analysis and machine learning theory algorithm and application.