

融合移位卷积与边缘检测的图像动态超分辨率重建^①



沈学利, 朱晓铭, 金海波

(辽宁工程技术大学 软件学院, 葫芦岛 125105)

通信作者: 朱晓铭, E-mail: 1945595956@qq.com

摘要: 针对固定网络架构和深度网络层导致的信息无法完全表达复杂场景预测高质量图像、高计算成本及部署困难等问题, 提出了一种具有宽网络结构的图像动态超分辨率网络 (wide dynamic super-resolution network, W-SDNet)。首先, 设计了一个由移位卷积残差增强结构组成的残差增强块, 以提高图像超分辨率的分层特征提取能力并减少计算成本。其次, 引入一个宽增强模块通过其双分支的 4 层并行结构, 在提取深度信息的同时利用动态网络的门机制选择性增强特征表达, 同时通过边缘检测算子融合的注意力机制增强边缘细节的表现力。紧接着, 采用组卷积和信道分割的细化块, 以防止在广泛增强块中组件间的干扰。最后, 通过一个构建块实现高质量图像的重建。实验结果表明, W-SDNet 在 5 个公开测试数据集上放大 4 倍时的峰值信噪比 (peak signal-to-noise ratio, PSNR) 指标均优于现有主流算法, 并且模型的参数量显著减少, 证明了 W-SDNet 在超分辨率重建的复杂度、性能及恢复时间方面的优势。

关键词: 图像超分辨率; 移位卷积; 动态网络; 边缘检测; 注意力机制

引用格式: 沈学利, 朱晓铭, 金海波. 融合移位卷积与边缘检测的图像动态超分辨率重建. 计算机系统应用, 2024, 33(9): 65-76. <http://www.c-s-a.org.cn/1003-3254/9618.html>

Dynamic Super-resolution Reconstruction of Images with Shift Convolution and Edge Detection

SHEN Xue-Li, ZHU Xiao-Ming, JIN Hai-Bo

(Software College, Liaoning Technical University, Huludao 125105, China)

Abstract: To address the challenges posed by fixed network architectures and deep network layers, such as incomplete expression of complex scene predictions, high computational costs, and deployment difficulties, this study proposes a new network called wide structure dynamic super-resolution network (W-SDNet). Initially, a residual enhancement block, consisting of shift convolution residual structures, is designed to enhance the capability of extracting hierarchical features for image super-resolution and to reduce computational costs. Next, a wide enhancement module is introduced, employing a dual-branch four-layer parallel structure to extract deep information while using a dynamic network's gating mechanism to selectively enhance feature expression. This module also utilizes an attention mechanism that integrates edge detection operators to improve the expressiveness of edge details. To prevent interference among components within the wide enhancement block, a refinement block utilizing group convolution and channel splitting is employed. Ultimately, high-quality image reconstruction is achieved through a construction block. Experimental results show that W-SDNet outperforms the existing mainstream algorithms in peak signal-to-noise ratio (PSNR) metrics when zoomed in 4 times on five publicly available test datasets, and the number of parameters in the model is significantly reduced. The results demonstrate the advantages of W-SDNet in terms of complexity, performance, and recovery time of super-resolution

① 基金项目: 国家自然科学基金面上项目 (62173171)

收稿时间: 2024-03-22; 修改时间: 2024-04-16; 采用时间: 2024-04-23; csa 在线出版时间: 2024-07-26

CNKI 网络首发时间: 2024-07-29

reconstruction.

Key words: image super-resolution; shift convolution; dynamic network; edge detection; attention mechanism

1 引言

图像超分辨率 (super-resolution, SR) 旨在从低分辨率图像 (low resolution, LR) 重建高分辨率图像 (high resolution, HR). 它是一种基础的低层次视觉任务, 其目标是利用先进的图像处理 and 机器学习技术重建缺失的细节, 提高图像质量. 在遥感成像、视频处理、医学图像分析等方面具有广泛的应用.

SR 方法可以概括为 3 种范式, 即基于插值的方法、基于优化的方法、基于学习的方法. 基于插值的方法主要依赖最近邻插值、双线性插值、双三次插值等操作来获得从 LR 图像到 HR 图像的映射. 尽管基于插值的方法简单易行, 但在高质量图像重建方面的性能较差. 为此, 部分研究者通过基于优化方法来恢复高质量的图像. 例如, 利用稀疏先验知识获得线性组合可有效预测 HR 图像. 迭代背投影 (iterative back-projection, IBP)^[1]通过迭代计算模拟低分辨率图像与观测的低分辨率图像之间距离, 不断更新预测的 HR 图像. 虽然以上基于优化的方法在图像超分辨率方面取得了很好的效果, 然而, 这种方法往往具有巨大的参数, 计算成本高, 不能在一些计算资源有限的设备上实际使用. 在单图像超分辨率 (single image super-resolution, SISR) 领域, 探索实用和实时的 SR 解决方案已经成为一个日益增长的趋势. 为了解决这一问题, 基于深度学习技术的图像 SR 方法应运而生. 与传统方法相比, 基于深度学习的方法能够更好地捕捉图像中的细节和结构信息, 提高图像重建的质量和准确性. 此外, 在深度学习中卷积神经网络 (convolutional neural network, CNN) 具有良好的自学习能力, 能够针对图像获取过程中存在的模糊和噪声等退化因素提供有效预防. 因此, CNN 在图像超分辨率领域得到了扩展.

Dong 等^[2]基于 CNN 提出了一种图像超分辨率方法 SRCNN (super-resolution CNN) 对图像超分辨率做出了重大贡献, 该方法只使用了 3 个 CNN 层进行特征提取. 虽然与传统的图像超分辨率方法相比, 它获得了更好的性能, 但可扩展性差. Kim 等^[3]以残差学习^[4]为基础, 提出了 VDSR (very deep convolutional networks super-resolution) 将网络深度增加到了 20 层, 并利用残

差块加速模型收敛, 提高性能, 但增加了计算复杂度, 并且可能产生梯度消失的问题. Lim 等^[5]提出的突破性的 EDSR (enhanced deep residual network for super-resolution) 将网络层数叠加到 60 层以上, 其卓越的性能充分体现了网络深度对图像重建质量的深刻影响. 另外, Tai 等^[6]利用门单元、跳跃连接和密集连接结构将获得的浅层特征继承到深层, 增强浅层对图像超分辨率的记忆能力. 虽然这些方法可以很好地实现图像的超分辨率, 但它们的输入较大, 可能会导致较大的计算成本.

随着注意力机制^[7-9]和 Transformer^[10,11]模型在图像处理领域的引入, 极大地推动了 SISR 的发展. Zhang 等^[8]设计的 RCAN (residual channel attention network) 通过残差结构来实现网络连接, 通过引入通道注意力来调整不同信道之间模型的权值, 增强网络对图像特征的建模能力. 肖子安等^[12]提出了一种多个级联的多尺度特征和坐标注意力模块, 实现对低分辨率遥感图像高频特征进行更充分的提取, 同时有效融合多尺度空间信息. Liang 等^[13]在图像恢复领域引入了 Swin Transformer, 并提出了 swinIR (image restoration using Swin Transformer) 方法. swinIR 是一种基于 Transformer 的图像超分辨率方法, 它利用了 Transformer 的强大表达能力, 实现了超分辨率领域先进的性能. 无论使用的是注意力机制还是 Transformer. 随着网络性能的提高, 模型参数的数量和计算工作量也显著增加, 使得超分辨率网络的实际应用变得不可行. 虽然这些方法是有效的, 并且极大地推动了图像超分辨率的发展, 但他们需要依赖更深的网络架构和强大的算力, 难以在资源受限设备实际应用. 此外, 以上方法获取的信息不能完全表达复杂场景下预测的高质量图像.

针对上述问题, 本文提出了一种宽结构图像动态超分辨率网络 (wide-structure dynamic image super-resolution network, W-SDNet), 旨在在不显著增加计算负担的前提下提升图像重建质量. W-SDNet 采用了一种递进动态网络架构, 通过融合浅层特征提取块、残差增强块、宽增强块、特征细化块和构造块, 实现了对复杂场景中丰富特征的有效提取和细化. 本网络的核

心在于动态调整网络结构以适应不同的图像内容,从而更精确地重建高质量图像.本文的主要贡献概括如下.

(1) 引入了移位卷积残差块,通过移位卷积层有效提取高频特征,同时降低模型的参数数量和计算成本.

(2) 提出了一个具有宽网络结构的动态宽增强块,结合蓝图可分离卷积和边缘检测算法的增强空间注意力模块,增强了模型对图像边缘特征表达能力.

(3) 通过采用组卷积和信道分割操作,进一步降低了计算复杂度,并利用特征细化块精确地学习和细化获得的特征.实验结果证明,相比现有的主流方法,W-SDNet在提升重建效果和降低模型复杂度方面展现了显著的优势.

2 相关工作

2.1 用于图像超分辨率的轻量级模型

为了提高图像处理的效果而增加网络深度,导致了巨大的计算成本,不利于超分辨率方法在现实生活中的应用.为了解决这个问题,研究人员提出了越来越多的精确和高效的模型. Kim等^[14]和 Tai等^[15]分别设计了DRCN (deeply-recursive convolutional network)和DRRN (deep recursive residual network)模型,利用递归结构和参数共享机制,大幅减少了模型参数的数量.然而,递归结构并没有减少计算工作量. Ahn等^[16]提出了一种具有群卷积的高效级联残差网络 (cascade residual network, CARN),提高了超分辨率的性能和效率. Hui等^[17]实现了一个轻量级信息多蒸馏网络 (iterative mean distillation network, IMDN),利用信息蒸馏机制 (information distillation mechanism, IDM) 逐级提取分层特征. RFDN (residual feature distillation network)^[18]改进了IMDN的体系结构,提出了残差特征精馏网络,用特征精馏连接代替IDM. 王一宁等^[19]利用密集块和反卷积提出一种拥有较少的层和参数的轻量密集神经网络模型,在训练速度和内存消耗方面均取得了良好的效果. 周登文等^[20]设计了轻量化且高效的区域互补注意力和多维注意力模块,在信道维度与空间维度对像素间的依赖关系建模,实现了轻量SR. Gao等^[21]将多路注意力模块、轻量级卷积块组合在一起,实现轻量级CNN,以便获得更准确的信息,提高图像超分辨率效果. 陈豪等^[22]针对网络复杂度较高的问题,提出了基于Transformer-CNN的轻量SR方法,使网络能够更好地捕获到图像中局部与全局的深度特征,降低了推

理时间.

根据上面提到的模型发现,优化网络架构对于实现图像超分辨率的轻量级CNN是有效的.虽然最近的轻量化SISR模型取得了良好的效果,但在高频恢复图像细节方面仍存在一些问题.所以考虑到获取的信息对复杂场景的鲁棒性,利用动态门、多视图思想和深度网络的设计原则,实现了一种基于CNN的轻量级的图像超分辨率.

2.2 图像超分辨率的动态神经网络

传统的神经网络结构一般是固定的,不能根据不同的场景灵活变化,这影响了他们在复杂场景中的表现.针对这一问题,研究人员开发了动态神经网络,动态网络能够根据输入数据特征及上下文信息,对网络结构和参数进行自适应调整,增加网络的自适应性和灵活性,并提高模型的鲁棒性和稳定性.由于动态网络模型优越性,已被应用于图像超分辨率领域.例如, Shi等^[23]使用了一种内容自适应技术,对齐现实世界中的图像以实现图像超分辨率.内容自适应技术可以根据输入数据的特征和上下文信息,动态地选择和调整网络的层次结构、连接方式和参数设置,使网络更好地适应不同的输入数据,并提高网络的性能和泛化能力. Wu等^[24]使用显著性感知模型来提取遥感图像中的显著性信息,然后通过动态路由策略将低分辨率图像的特征映射动态地分配给高分辨率图像的不同区域,以保留更多的细节信息进而实现超分辨率重建. Chen等^[25]通过边缘信息分配不同的网络来恢复图像超分辨率中的详细信息.为了提高图像超分辨率的性能,采用不同的注意机制分别提取局部和全局特征,从而为图像超分辨率提供更有用的详细信息.根据上述所提,可以看到动态网络对于图像的超分辨率是有效的.受此启发,本文应用了一个动态网络实现图像超分辨率.

3 本文方法

3.1 网络架构

基于CNN的单图像超分辨率方法通常倾向于增加网络深度以提高图像处理性能.然而,这种方法容易导致计算成本增加,并且不利于在低配置设备上实现.此外,单一的网络结构往往难以提取复杂场景中的丰富特征.针对这些问题,本文提出了一种新型的递进动态网络(W-SDNet).整体结构如图1所示,该网络由多个组件组成,包括浅层特征提取块、残差增强块、宽

增强块、特征细化块以及构造块。首先,低分辨率图像经过蓝图可分离卷积(blueprint separable convolution, BSConv)^[26]进行浅层特征的提取。接下来,通过由3个移位卷积残差块(shift-conv residual block, SC-ResBlock)组成的残差增强块,对图像的层次信息进行提取和增强,从而丰富特征的层次表达。为了应对复杂场景的挑战,引入了宽增强块。这个模块采用结合改进增强空间注意力机制的互补块(complementary block, CBlock)和一个动态门机制,以更有效地为复杂场景提取更丰富的信息。通过这种设计,网络能够更好地适应各种复杂场景,并提取更具表达力的特征。在宽增强块之后,使用了特征细化块。这个模块通过组卷积和信道分割

的策略,消除了宽增强块中可能产生的干扰,进一步优化了从宽增强块中学习到的特征。最后,构造块负责通过上采样和卷积层,将前述模块提取的特征和浅层特征进行整合,以重建高质量的图像。通过这种方式,W-SDNet不仅优化了特征提取的过程,同时也提升了最终图像的重建质量。实现过程如式(1)所示:

$$I_{HR} = W\text{-SDNet}(I_{LR}) = RB(FEB(WEB(REB(SF(I_{LR})))))) \quad (1)$$

其中, I_{LR} 表示低分辨率图像, $W\text{-SDNet}$ 表示 $W\text{-SDNet}$ 的函数, SF 、 REB 、 WEB 、 FEB 、 RB 分别表示浅层特征提取块、残差增强块、宽增强块、特征细化块、重建模块的函数。 I_{HR} 表示 $W\text{-SDNet}$ 输出的高分辨率图像。

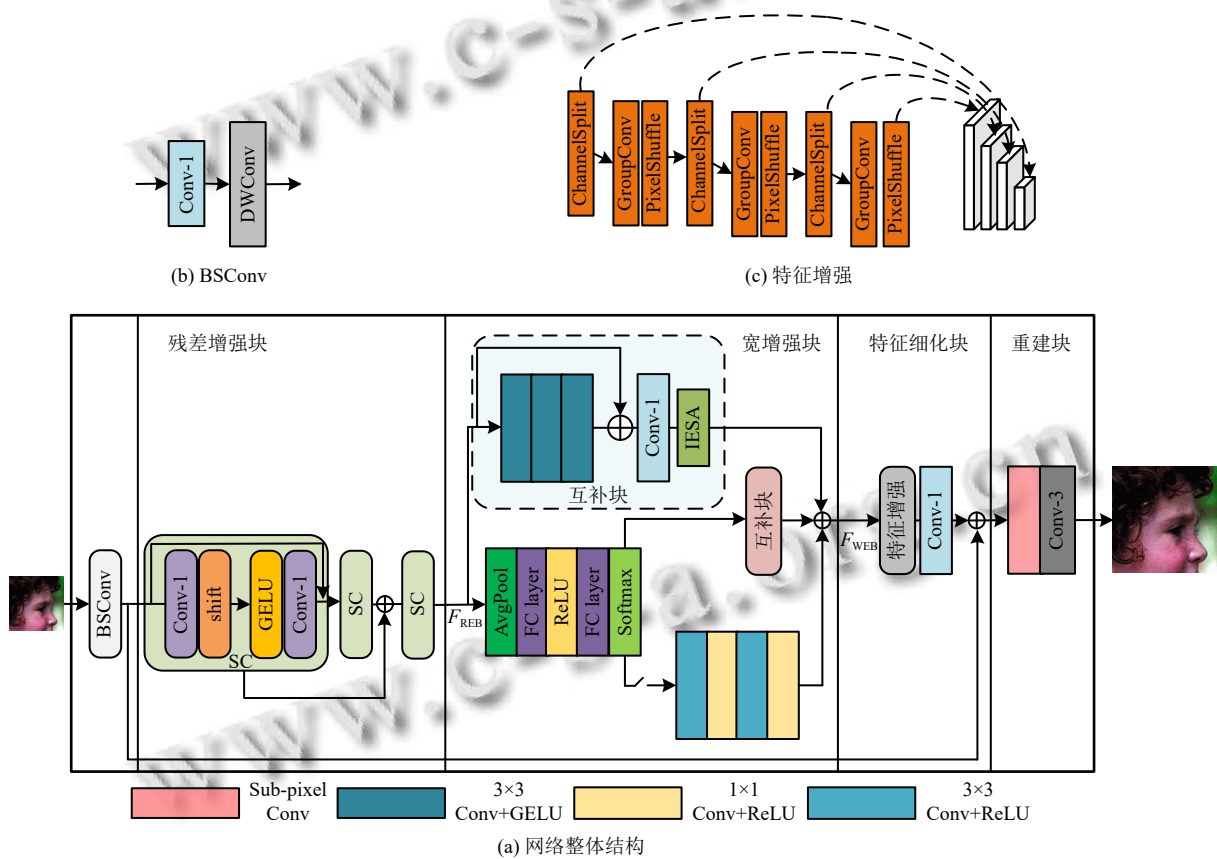


图1 网络结构图

3.2 浅层特征提取块

浅层特征提取部分采用蓝图可分离卷积对输入图像进行浅层特征提取,BSConv具体架构如图1(b)所示。BSConv是深度可分离卷积(depthwise separable convolution, DSConv)^[27]的改进变体,它由 1×1 卷积和深度卷积组成。更好地利用了核内相关性,相比于普通

卷积减少了冗余运算,文献[26]证明BSConv对于高效SR是有利的。 F_{SF} 表示提取得到的浅层特征,浅层特征提取过程可以表述为式(2):

$$F_{SF} = SF(I_{LR}) \quad (2)$$

3.3 残差增强块

残差网络已经广泛应用在图像超分辨率领域,大

量研究表明,通过增加网络层深度,所提取的特征就越丰富,但过深的网络层数会引起梯度爆炸或梯度消失的问题,而多个残差连接的残差块可以很好地解决这个问题,所以本文模型利用残差增强块对层次信息进行提取和增强.残差增强块采用3个移位卷积残差块(shift-conv residual block, SC-ResBlock)提取深度信息,SC-ResBlock是由移位残差卷积层, GELU^[28]激活函数, 1×1卷积层组成.最后使用残差学习操作收集浅层特征,以增强模型的记忆能力.

如图2(a)所示是文献[5]中提出的残差块,广泛应用于SR网络中.3×3卷积运算在平衡模型容量和计算成本方面具有一定优势,但会增加参数和计算开销,所以将3×3卷积层替换成计算成本更低的1×1卷积层,虽然1×1卷积很大程度上减少了参数数量和计算成本,但是由于1×1卷积操作只对单个像素的特征进行处理,这会影响建模,因此在这里通过空间移位操作^[29],如图3所示.通过重新排列特征图的空间位置,在不增加计算成本的情况下,增强了跨信道组的局部空间特征聚合. Shift-Conv层(简化为SC层)由1×1卷积层和空间移位操作堆叠,因此SC层扩展了正常的1×1卷积,具有局部特征聚合,且参数更少.与正常的3×3卷积相比,SC层在保持相当性能的同时显著减少了参数数量.在激活单元方面, GELU具有良好非线性表达能力,逐渐成为近期研究的首选.残差增强块具体操作如式(3):

$$\begin{aligned} F_{\text{REB}} &= \text{REB}(F_{\text{SF}}) \\ &= \text{SC}(F_{\text{SF}} + \text{SC}(\text{SC}(F_{\text{SF}}))) \end{aligned} \quad (3)$$

其中, REB表示残差增强函数, SC表示移位卷积残差操作, F_{REB} 表示经过残差增强操作的输出.

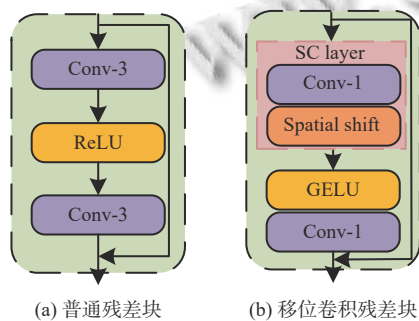


图2 两种残差块对比

3.4 宽增强块

现有图像超分辨率网络普遍通过增加网络深度来提升网络重建效果,但忽视了网络宽度对于网络影响,

根据GoogLeNet^[30],已知增大宽度可以提取更多的互补信息,以提高图像应用的性能.受此启发,构建了宽增强块,用以增加网络特征表达能力,提高网络性能,进一步证明增加网络宽度对图像重建具有积极影响.宽增强块紧跟在残差增强块之后,接收残差增强块的输出 F_{REB} 作为其输入,该模块由双分支的4层网络并行结构组成,每个分支各自执行不同的任务以增强特征表达.第1分支用于深度信息的提取,设计了一个互补块来实现.第2分支由一个动态网络组成,它使用平均池(AvgPool)、全连接层、ReLU激活函数、全连接层和Softmax组成的门机制来选择一个门,即选择的动态门.具体而言,动态网络通过门机制的Softmax层预测两个关于是否激活选定的动态门的概率值,这两个值的和为1,每个值代表概率,分别对应于是否激活选定的动态门.如果第2个值大于0.75,则该分支的输出为固定动态门和选定动态门输入,否则,它的输出只有一个固定门的输入.其中固定动态门结构与互补块结构相同,用以更好提取深层特征,选择动态门由核为3×3的Conv+ReLU和核为1×1的Conv+ReLU两种组合堆叠.最后通过残差学习收集互补特征,其中各输入输出通道均为64,过程可表示为式(4):

$$\begin{aligned} F_{\text{WEB}} &= \text{WEB}(F_{\text{REB}}) \\ &= (\text{CB}(F_{\text{REB}}) + \text{FDG}(\text{GM}(F_{\text{REB}}))) \Theta \text{SDG}(\text{GM}(F_{\text{REB}})) \end{aligned} \quad (4)$$

其中, F_{WEB} 表示宽增强块输出. CB、GM、FDG、SDG分别表示互补块、门机构、固定动态门、选定动态门函数, Θ 表示选择的门,如果第2个结果大于0.75, Θ 为残差学习操作.否则, Θ 表示被删除的部分.

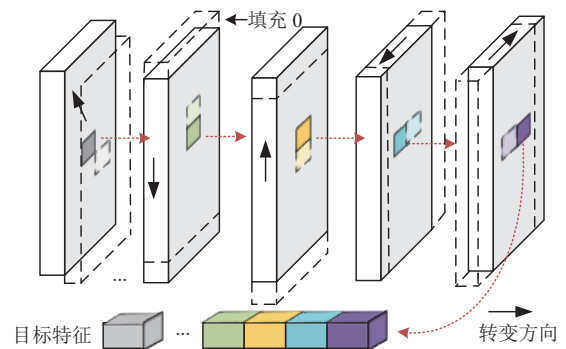


图3 空间位移操作说明

3.4.1 互补块

为了增强特征表达能力和模型的鲁棒性,在宽增强块中增加了互补块.受文献[31]启发,互补块采用剩

余局部特征块, 在保持模型容量的前提下, 显著减少推理时间, 互补块只用 3 个叠加的 Conv+GELU 进行特征提取, 过程如式 (5):

$$\begin{cases} F_{\text{refined1}} = \text{RM}_1(F_{\text{REB}}) \\ F_{\text{refined2}} = \text{RM}_2(F_{\text{refined1}}) \\ F_{\text{refined3}} = \text{RM}_3(F_{\text{refined2}}) \end{cases} \quad (5)$$

其中, RM_i 为第 i 个细化模块, F_{refined_i} 为第 i 个细化特征, 经过多个局部特征细化步骤后, 将最后细化的特征 F_{refined} 与跳过的特征 F_{REB} 相加, 得到 F_{refined} , 过程如式 (6):

$$F_{\text{refined}} = F_{\text{REB}} + F_{\text{refined3}} \quad (6)$$

最后通过一个 1×1 卷积进行特征融合和降维, 通过改进的增强空间注意力模块, 增强边缘信息和重要信息表达能力. 如式 (7):

$$F_{\text{WEB}} = \text{DSA}(C_1(F_{\text{refined}})) \quad (7)$$

其中, C_1 表示 1×1 卷积, DSA 表示注意力机制.

3.4.2 改进的增强空间注意力

注意力机制提出以来被广泛应用到图像领域, 并取得了令人满意的效果, ESA (enhanced spatial attention)^[17] 注意力是一种增强空间注意力的方法, 其在提升模型表达能力和改善长距离依赖方面具有有效性, 但随之而来的是参数的增加和计算复杂度的增加, 另一方面, 在对图像的高低频信息进行区分和整合的过程中需要额外关注边缘信息, 因为图像每个通道都包含大量高低频信息分量, 而高频信息分量往往含有许多边缘信息, 但现有方法大都不注重边缘的高频信息^[32]. 为解决以上问题, 设计了如图 4 所示的 IESA (improved enhanced spatial attention), 该注意力模块在 ESA 注意力机制的基础上, 融合了 Canny 边缘检测算子^[33], 对输入特征图进行边缘特征提取, 并与利用卷积操作得到的特征图进行融合运算. 得到边缘特征增强的特征图, 如式 (8) 所示:

$$\begin{cases} Y_c = \text{Canny}(Y_{\text{in}}) \\ Y_{\text{out}} = Y_e \otimes Y_c \end{cases} \quad (8)$$

其中, $\text{Canny}(\cdot)$ 为 Canny 边缘检测算子.

针对参数的增加和计算复杂度的增加的问题, 采用 BSConv 卷积代替原有的普通卷积, 因为与普通卷积相比, BSConv 采用分离卷积方式, 降低了计算复杂度和参数量, 提升计算速度, 同时 BSConv 采用不同

卷积核进行特征提取, 可以更好地捕捉到输入数据的局部和全局特征. 将经过边缘增强的特征图, 通过 BSConv 降低输入特征的通道维数. 然后使用大步卷积和大步最大池化层来减小空间大小. 通过一组卷积学习特征图空间维度之间的相关性, 采用基于插值的上采样来恢复空间大小. 为了提高效率结合跳跃连接, 通过 BSConv 对特征进行进一步处理, 恢复信道大小. 最后, 通过 Sigmoid 函数生成注意矩阵, 并乘以原始输入特征. 通过跳跃连接初始特征图与网络底层连接进行残差学习, 以降低网络训练难度.

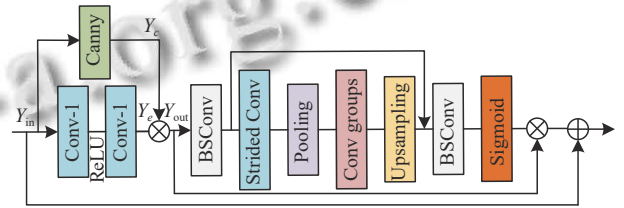


图 4 改进的增强空间注意力

3.5 特征细化块

为了防止宽增强块中某些成分的干扰, 对宽增强块中获得的特征进行特征细化. 通过组卷积逐渐添加高频细节, 并利用组卷积和信道分割操作来降低计算复杂度, 获得一组分辨率逐渐提高的特征图, 如图 1(c) 所示. 首先, 将输入的特征映射 $M_{\text{in}} \in \mathbb{R}^{C \times H \times W}$ 沿着通道维度分成两部分, 其中一部分包含 $1/4$ 的特征进行特征映射, 记为 $\tilde{M}_1 \in \mathbb{R}^{\frac{C}{4} \times H \times W}$, 并保持为当前分辨率. 使用组卷积和像素变换向剩余的 $3/4$ 特征映射 $\tilde{M} \in \mathbb{R}^{\frac{3C}{4} \times H \times W}$ 添加高频细节, 并将空间分辨率提高比例因子 2. 然后, 将二级分辨率的特征图分成两部分, 保留 $1/3$ 的特征图 $M_2 \in \mathbb{R}^{\frac{C}{4} \times 2H \times 2W}$ 作为当前分辨率. 其余 $2/3$ 特征映射 $\tilde{M}_2 \in \mathbb{R}^{\frac{2C}{4} \times 2H \times 2W}$ 的分辨率通过使用群卷积和像素变换再次加倍. 最后, 将第 3 级分辨率的特征图沿通道维度分成两半. 一半的特征映射 $M_3 \in \mathbb{R}^{\frac{C}{4} \times 4H \times 4W}$ 保持当前分辨率, 另一半 $\tilde{M}_3 \in \mathbb{R}^{\frac{C}{4} \times 4H \times 4W}$ 使用卷积层和像素 shuffle 操作进行两次升级. 最后一级的分辨率是原始输入的 8 倍, 最后一级的特征映射记为 $M_4 \in \mathbb{R}^{\frac{C}{4} \times 8H \times 8W}$. 因此, 得到了渐进分辨率的特征映射 $\{M_i\}_{i=1}^4$. 该过程具体过程如式 (9):

$$\begin{cases} M_1, \tilde{M}_1 = \text{Split}(M_{\text{in}}) \\ M_2, \tilde{M}_2 = \text{Split}(\text{PixelShuffle}(\text{GroupConv}_{3,3C}^{(3)}(\tilde{M}_1))) \\ M_3, \tilde{M}_3 = \text{Split}(\text{PixelShuffle}(\text{GroupConv}_{3,2C}^{(2)}(\tilde{M}_2))) \\ M_4 = \text{PixelShuffle}(\text{Conv}_{3,C}(\tilde{M}_3)) \end{cases} \quad (9)$$

其中, $\text{GroupConv}_{k,C}^{(n)}(\cdot)$ 表示一个卷积层, 其核大小为 $k \times k$, 信道数为 C , 组数为 n . $\text{PixelShuffle}(\cdot)$ 为像素重洗操作. 通过组卷积运算得到的特征图经过一个 1×1 卷积进行通道融合, 得到特征细化块输出的特征图 F_{FRB} , 具体操作如式 (10):

$$F_{\text{FRB}} = C_1(M_1, M_2, M_3, M_4) \quad (10)$$

3.6 重建块

重建块由亚像素卷积层和 3×3 卷积组成, 重建阶段首先由亚像素卷积对融合后特征图进行上采样, 再使用 3×3 卷积构建高质量图像. 得到最终高分辨率图像 F_{HR} 具体过程可表示为式 (11):

$$F_{\text{HR}} = \text{RB}(F_{\text{FRB}} + F_{\text{SF}}) = C_3(\text{Sub}(F_{\text{FRB}} + F_{\text{SF}})) \quad (11)$$

其中, $\{L_L^i, L_H^i\}_{i=1}^n$ 和 Sub 分别表示 3×3 卷积和亚像素卷积.

3.7 损失函数

在不增加计算成本的情况下, 增强了跨信道组的局部空间的均方误差 (mean squared error, MSE)^[34] 是 SR 任务的常用损失函数. 通过估计量与被估计量之间差异程度反应模型重建情况, 均方误差越小, 说明模型的预测能力越强. 通过实验对比 MSE 更适合模型, 所以为了更好地提升模型性能, 选择 MSE 作为损失函数来训练 W-SDNet 模型. 即将训练数据集的 $\{L_L^i, L_H^i\}_{i=1}^n$ 作用于 MSE 上训练 W-SDNet 模型, 其中 L_L^i 和 L_H^i 分别作为第 i 个低分辨率训练图像块和高分辨率训练图像块. 设 n 为训练图像 patch 的个数. 上面的描述可以表示为式 (12):

$$L(p) = \frac{1}{2n} \sum_{i=1}^n \|\text{W-SDNet}(L_L^i) - L_H^i\|^2 \quad (12)$$

其中, L 为 MSE 的损失函数, p 表示学习到的参数.

4 实验结果分析

4.1 数据集

选取使用广泛的 DIV2K^[35] 公共数据集训练模型. DIV2K 在 NTIRE 2017 挑战赛被提出, 并广泛应用于图像超分辨率领域. 该数据集由 1000 张高品质图像组成, 涵盖了包括人物、自然景观和生活场景在内的多种类型且分辨率均为 2K 图像. 选取其中 800 张用于训练, 100 张用于验证, 100 张用于测试. 测试时, 使用 Set5^[36]、Set14^[37]、BSD100^[38]、Urban100^[39]、Manga-

109^[40]、AIM2019^[35] 和 RealSR^[41] 测试数据集来评估不同算法的性能. 此外, 这些数据集中的每个图像都包含了 3 个版本 $\times 2$ 、 $\times 3$ 、 $\times 4$ 的 LR 和 HR 图像.

所提方法采用了 3 个关键指标来全面评估超分辨率重建模型的性能: Y 信道平均峰值信噪比 (PSNR)、结构相似度 (SSIM)、感知图像相似性 (learned perceptual image patch similarity, LPIPS). PSNR 和 SSIM 作为基础指标, 主要衡量图像质量的传统视觉特征, 其中高值代表更优的图像恢复质量. 为进一步评估图像的感知细节质量, 引入了 LPIPS 指标. LPIPS 通过分析图像深层特征的差异, 专注于图像内容的感知属性, 低 LPIPS 值表示重建图像与原始图像在视觉上更为相似.

4.2 实验设置

本文实验在 NVIDIA RTX 3090 GPU 上进行; 选用 Ubuntu 20.04 版本, Python 3.8.13 作为编程环境, 并选择深度学习框架 PyTorch 1.13.1, CUDA 版本 11.7, cuDNN 8.5.0; Batch 和 patch size 分别设置为 64, 步数是 $6E+5$. 在实验中, 初始学习率为 $1E-4$, 每走 $4E+5$ 步, 学习速度就会减半. 采用 Adam 优化器进行优化, 其中参数 β_1 为 0.9, β_2 为 0.999.

4.3 实验与结果分析

为了验证 W-SDNet 的先进性, 对 W-SDNet 算法与已有先进算法进行了广泛的实验对比. 实验结果如表 1-表 3 所示, 分别展示了所提算法与其他先进算法在 2 倍、3 倍、4 倍超分辨率任务上的客观评价指标结果 (其中加粗字体表示最优结果, 下划线字体表示次优结果, *表示重新复现的结果).

由表 1-表 3 可以看出, 本文方法在不同放大倍数上都取得了最优和次优的良好表现. 当放大倍数为 4 时, 相比于其他算法, 本文算法在 5 个测试数据集中均取得了最优的 PSNR. 在 Set5 数据集放大倍数为 3 时, 相比于 DSRNet^[42], 本文方法获得的 PSNR 和 SSIM 提高了 0.06 dB 和 0.0009, 同时参数量相比于 DSRNet 降低了 20.76%. 综上所述, 可以看出本文算法可以在减少参数量的同时获得更优秀的性能指标.

为了进一步评估图像的感知细节质量, 将提出的 W-SDNet 在质量更高的 AIM2019 和 RealSR 数据集上做了进一步的测试, 并与现有的主流方法 ZSSR、DASR 等方法进行了实验对比分析, 对比结果如表 4 所示. 实验结果表明, 相比于其他先进的超分辨率模型, 所提方

法在两个数据集上均展现出了优异的感知图像重建性能. 特别是在 RealSR 数据集上, 低 LPIPS 值表明本文

方法能够有效地减少图像间的感知差异, 重建出与原始图像在视觉上更为接近的高质量图像.

表 1 放大倍数为 2 时各算法在不同数据集上的平均 PSNR (dB) 和 SSIM 值对比

| 算法 | 参数量 (k) | Multi-Adds (G) | Set5 | Set14 | BSD100 | Urban100 | Manga109 |
|--------------------------|---------|----------------|--------------|--------------|--------------|--------------|--------------|
| Bicubic | — | — | 33.36/0.9299 | 30.24/0.8688 | 29.25/0.8431 | 26.88/0.8403 | 30.80/0.9339 |
| SRCNN ^[2] | 57 | 52.7 | 36.66/0.9542 | 32.45/0.9067 | 31.36/0.8879 | 29.50/0.8946 | 35.60/0.9663 |
| FSRCNN ^[43] | 12 | 6.0 | 37.00/0.9558 | 32.63/0.9088 | 31.53/0.8920 | 29.88/0.9020 | 36.67/0.9710 |
| VDSR ^[4] | 666 | 612.6 | 37.53/0.9587 | 33.05/0.9127 | 31.90/0.8960 | 30.77/0.9141 | 37.22/0.9750 |
| DRCN ^[15] | 1774 | 17974.3 | 37.63/0.9588 | 33.04/0.9118 | 31.85/0.8942 | 30.75/0.9133 | 37.55/0.9732 |
| LapSRN ^[44] | 251 | 29.9 | 37.52/0.9591 | 32.99/0.9124 | —/— | 30.41/0.9103 | 37.27/0.9740 |
| DSRNet ^[42] | 746 | 49.2 | 37.61/0.9584 | 33.30/0.9145 | 31.96/0.8965 | 31.41/0.9209 | 38.05/0.9753 |
| MSSR-Net ^[45] | 745 | — | 37.64/0.9590 | 33.21/0.9150 | 31.95/0.8970 | —/— | —/— |
| *ACNet-B ^[46] | 1283 | 8.09 | 37.60/0.9584 | 33.32/0.9141 | 31.97/0.8966 | 31.47/0.9221 | —/— |
| EBSR ^[47] | 82 | 2.27 | 37.64/0.9580 | 33.16/0.9130 | 31.88/0.8950 | 31.11/0.9170 | —/— |
| Our | 598 | 60.8 | 37.66/0.9588 | 33.36/0.9149 | 32.00/0.8967 | 31.47/0.9213 | 38.09/0.9759 |

表 2 放大倍数为 3 时各算法在不同数据集上的平均 PSNR (dB) 和 SSIM 值对比

| 算法 | 参数量 (k) | Multi-Adds (G) | Set5 | Set14 | BSD100 | Urban100 | Manga109 |
|--------------------------|---------|----------------|--------------|--------------|--------------|--------------|--------------|
| Bicubic | — | — | 30.39/0.8682 | 27.55/0.7742 | 27.21/0.7385 | 24.46/0.7349 | 26.95/0.8556 |
| SRCNN ^[2] | 57 | 52.7 | 32.75/0.9090 | 29.30/0.8215 | 28.41/0.7863 | 26.24/0.7989 | 30.48/0.9117 |
| FSRCNN ^[43] | 12 | 6.0 | 33.18/0.9140 | 29.37/0.8240 | 28.53/0.7910 | 26.43/0.8080 | 31.10/0.9210 |
| VDSR ^[4] | 666 | 612.6 | 33.66/0.9213 | 29.77/0.8314 | 28.82/0.7976 | 27.14/0.8279 | 32.01/0.9340 |
| DRCN ^[15] | 1774 | 17974.3 | 33.82/0.9226 | 29.76/0.8311 | 28.80/0.7963 | 27.15/0.8276 | 32.24/0.9343 |
| LapSRN ^[44] | 502 | 129.9 | 33.81/0.9220 | 29.79/0.8325 | 28.82/0.7980 | 27.07/0.8275 | 32.21/0.9350 |
| DSRNet ^[42] | 752 | 49.3 | 33.92/0.9227 | 30.10/0.8378 | 28.90/0.8003 | 27.63/0.8402 | 32.73/0.9391 |
| MSSR-Net ^[45] | 745 | — | 33.98/0.9240 | 30.04/0.8360 | 28.90/0.8060 | —/— | —/— |
| *ACNet-B ^[46] | 1283 | 8.09 | 33.97/0.9233 | 30.14/0.8379 | 28.93/0.8006 | 27.88/0.8447 | —/— |
| EBSR ^[47] | 82 | 2.27 | 33.89/0.9210 | 29.95/0.8340 | 28.82/0.7970 | 27.33/0.8330 | —/— |
| Our | 602 | 61.4 | 33.98/0.9236 | 30.15/0.8381 | 28.94/0.8007 | 27.64/0.8401 | 32.74/0.9393 |


表 3 放大倍数为 4 时各算法在不同数据集上的平均 PSNR (dB) 和 SSIM 值对比

| 算法 | 参数量 (k) | Multi-Adds (G) | Set5 | Set14 | BSD100 | Urban100 | Manga109 |
|--------------------------|---------|----------------|--------------|--------------|--------------|--------------|--------------|
| Bicubic | — | — | 28.42/0.8104 | 26.00/0.7027 | 25.96/0.6675 | 23.14/0.6577 | 24.89/0.7866 |
| SRCNN ^[2] | 57 | 52.7 | 30.48/0.8626 | 27.50/0.7513 | 26.90/0.7101 | 24.52/0.7221 | 27.58/0.8555 |
| FSRCNN ^[43] | 12 | 4.8 | 30.72/0.8660 | 27.61/0.7550 | 26.98/0.7150 | 24.62/0.7280 | 27.90/0.8610 |
| VDSR ^[4] | 666 | 612.6 | 31.35/0.8838 | 28.01/0.7674 | 27.29/0.7251 | 25.18/0.7524 | 28.83/0.8870 |
| DRCN ^[15] | 1774 | 17974.3 | 31.53/0.8854 | 28.02/0.7670 | 27.23/0.7233 | 25.14/0.7510 | 28.93/0.8854 |
| LapSRN ^[44] | 513 | 149.9 | 31.54/0.8852 | 28.09/0.7700 | 27.32/0.7275 | 25.21/0.7562 | 29.09/0.8900 |
| DSRNet ^[42] | 761 | 50.9 | 31.71/0.8874 | 28.38/0.7760 | 27.43/0.7303 | 25.65/0.7693 | 29.71/0.8975 |
| MSSR-Net ^[45] | 745 | — | 31.75/0.8910 | 28.35/0.7760 | 27.41/0.7310 | —/— | —/— |
| *ACNet-B ^[46] | 1283 | 8.09 | 31.72/0.8871 | 28.41/0.7763 | 27.42/0.7307 | 25.59/0.7760 | —/— |
| EBSR ^[47] | 82 | 2.27 | 31.64/0.8850 | 28.22/0.7720 | 27.30/0.7270 | 25.36/0.7610 | —/— |
| Our | 603 | 61.8 | 31.76/0.8879 | 28.42/0.7765 | 27.43/0.7309 | 25.66/0.7698 | 29.73/0.8977 |

表 4 不同方法在 AIM2019 和 RealSR 测试集放大 4 倍细节感知对比

| 方法 | AIM2019 | RealSR |
|----------------------|---------|--------|
| SRCNN ^[5] | 1.325 | 0.793 |
| ZSSR ^[48] | 0.634 | 0.386 |
| DASR ^[49] | 0.471 | 0.257 |
| Our | 0.382 | 0.201 |

将提出的 W-SDNet 与经典模型 LapSRN、VDSR, 先进轻量级模型 IMDN、RFDN, 基准模型 DSRNet 进行了视觉质量比较, 在 Urban100 测试数据集上放大 4 倍 SR 结果如图 5 和图 6 所示. 对于 img_078 可以发现 LapSRN 和 VDSR 生成的图像相对模糊, 条纹存在伪影, 线条变形, 而本文提出 W-SDNet 虽然存在形变,

但相比于其他方法纹理更清晰,更贴近高分辨率图像.对于发现 LapSRN、VDSR、IMDN、DSRNet 不仅产生的图像模糊,而且线条变形严重,不能准确恢复图像纹理细节,W-SDNet 恢复了图像大多细节,失真

更小具有更好的视觉效果.对于图片纹理细节恢复的优势,得益于本文提出的融合边缘检测算子的注意力机制,更加关注边缘纹理信息,使生成的 SR 图像可以更好表达边缘特征.

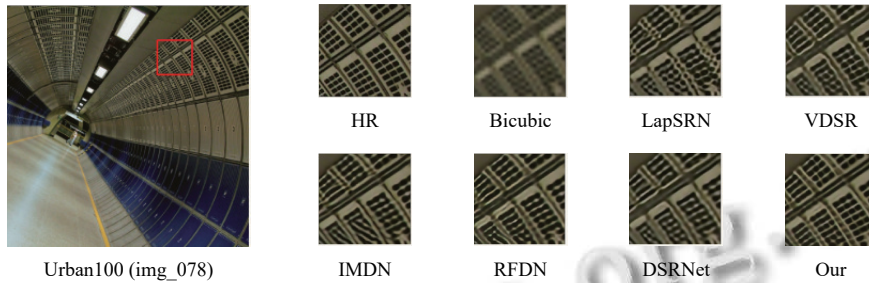


图5 放大倍数为4各算法在视觉对比

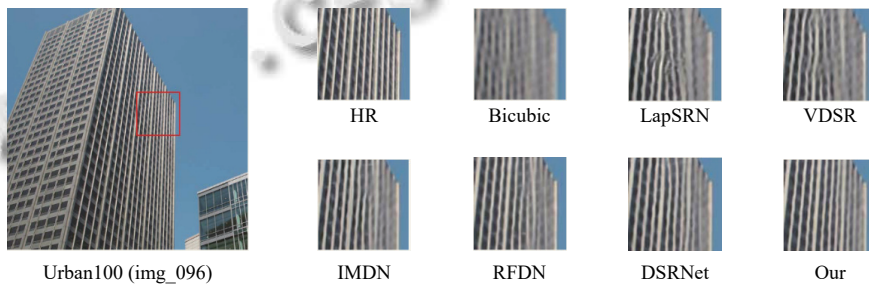



图6 放大倍数为4各算法在视觉对比

4.4 模型复杂度分析

在本节中,对算法的模型复杂度进行了研究,包括大小、运行时间、Flops,并与其他经典图像超分辨率算法进行了比较.如表5和图7所示,本文提出的方法平均运行时间在所提方法中排名第四,虽然SRCNN、FSRCNN、VDSR运行时间相比于本文方法更短,但本文算法的PSNR显著高于这3种模型,且参数数量和Flops少于VDSR.表5中IMDN与RFDN的PSNR略高于本文方法,但本文方法参数数量更少的同时单幅图像平均运行时间仅为这两种方法的1/4.通过以上实验对比分析,本文方法很好地平衡了模型性能和运行时间之间的矛盾,保持性能的同时降低了模型的参数数量和运行时间.

4.5 消融实验

为了验证算法各模块的有效性进行了消融实验,通过在Set14数据集上放大2倍的SSIM、PSNR的变化,验证移位残差块、互补块、注意力机制、特征增强块的有效性,以及门机制中阈值的选取和损失函数的选取.

表5 不同方法在Set14数据集放大4倍复杂度对比

| 算法 | PSNR (dB) | 参数量 (k) | Multi-Adds (G) | 运行时间 (ms) |
|------------------------|-----------|---------|----------------|-----------|
| SRCNN ^[5] | 27.50 | 57 | 52.7 | 1.90 |
| FSRCNN ^[43] | 27.61 | 12 | 4.8 | 2.22 |
| VDSR ^[7] | 28.01 | 666 | 612.6 | 8.95 |
| LapSRN ^[44] | 28.09 | 513 | 149.9 | 66.81 |
| IMDN ^[17] | 28.58 | 712 | 40.5 | 20.56 |
| RFDN ^[18] | 28.61 | 545 | 23.9 | 20.40 |
| DSRNet ^[42] | 28.38 | 746 | 50.9 | 5.971 |
| Our | 28.42 | 603 | 61.8 | 5.668 |

(1) 验证移位残差块,互补块,特征增强块有效性.在验证各模块有效性之前,将DSRNet作为基线网络,对DSRNet中对应模块进行替换,通过模型参数数量和PSNR变化来验证各模块对模型性能.通过表6所示,可以发现移位残差块和特征细化块可以在损失很少的PSNR情况下,极大地减少模型的参数.而互补块虽然增加了一些参数和运算量,但对模型的重建性能提高有帮助,通过各模块的组合,可以在不损失重建效果的情况下,减少模型参数,提升模型性能.这表明各模块对于提升图像超分辨率性能指标是有效的.

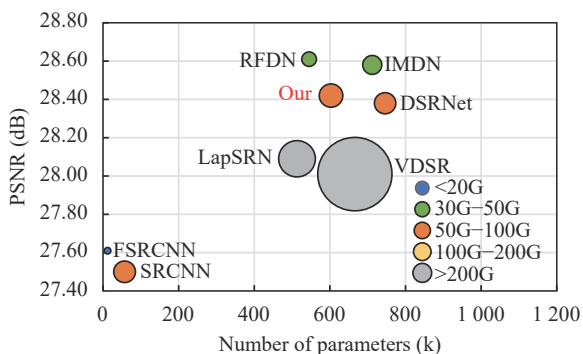


图7 在 Set14 数据集放大 4 倍参数量、Flops、PSNR 对比

表6 放大倍数为 4 时各模块的 PSNR 值和参数量对比

| SC-ResBlock | 互补块 | 特征细化块 | 参数量 (k) | PSNR (dB) |
|-------------|-----|-------|---------|-----------|
| × | × | × | 746 | 33.30 |
| √ | × | × | 613 | 33.28 |
| × | √ | × | 752 | 33.35 |
| × | × | √ | 718 | 33.25 |
| √ | √ | √ | 598 | 33.36 |

(2) 验证互补块 IESA 的有效性. 为了验证本文提出的 IESA 注意力模块的有效性, 在本文算法的基础上通过替换其他 3 种不同的注意力模型, 对比不同注意力机制对于网络参数量和 PSNR、SSIM 的变化情况, 验证增强注意力机制性能. 通过表 7 所示, 不同注意力对于网络参数量以及 PSNR、SSIM 具有不同的影响. 仅使用通道平均值的 SA 和 CCA 的模型虽然参数较少, 但性能明显比使用 ESA 和本文注意力模型的差. 而使用本文的模型保持与 ESA 相当参数的同时, 取得了比使用 ESA 的模型更好的性能. 所以尽管本文提出的注意力增加了少量参数, 但是获得了最好的 PSNR 和 SSIM. 证明该注意力模型具有有效性.

表7 使用不同注意力对模型参数量和性能影响

| 模块 | 参数量 (k) | PSNR (dB) | SSIM |
|-------|---------|-----------|--------|
| +SA | 621 | 33.18 | 0.9139 |
| +CCA | 632 | 33.14 | 0.9132 |
| +ESA | 595 | 33.27 | 0.9141 |
| +IESA | 598 | 33.36 | 0.9149 |

(3) 门机制中阈值的选取. 宽增强块中动态门阈值的选取对于动态神经网络十分重要, 阈值选取的好坏直接影响动态网络的性能, 所以选择一种二分法来设置他的阈值. 选取了 0.5、0.75、0.9 这 3 个值来作为动态门的备选阈值. 通过实验, 如表 8 所示, 发现当阈值为 0.75 时, 网络的 PSNR 高于其他阈值的 PSNR, 所以选择 0.75 作为动态门阈值.

(4) 损失函数的选取. 为了探索不同损失函数对于

模型的影响, 选取了在图像超分辨率领域使用最广泛的 L1 Loss (mean absolute error, MAE)、L2 Loss (mean squared error, MSE) 和平滑 L1 损失 (smooth L1 Loss, SLL) 3 种损失函数和文献[50]提出的 FFT 损失函数 (fast Fourier transform loss function, FFT). 在 Set14 数据集放大倍数为 4 倍下, 比较模型使用这 4 种不同损失函数的 PSNR.

表8 不同阈值动态门的 PSNR 比较

| 阈值 | PSNR (dB) |
|------|-----------|
| 0.5 | 33.24 |
| 0.75 | 33.36 |
| 0.9 | 33.27 |

通过表 9 所示, 发现本文方法使用 MSE 损失函数获得了最高的 PSNR, 并且在实验中发现 L1 Loss 收敛速度较慢, SLL Loss 的 PSNR 比较低, FFT Loss 开始收敛速度很慢, 后续甚至出现梯度消失的问题, 所以通过综合考量, 决定使用 MSE 作为本文模型的损失函数.

表9 不同损失函数对模型 PSNR 的影响

| 损失函数 | PSNR (dB) |
|------|-----------|
| L1 | 28.29 |
| MSE | 28.42 |
| SLL | 28.23 |
| FFT | — |

5 结语

本文提出的一种具有宽网络结构的图像动态超分辨率网络 (W-SDNet), 包含浅层特征提取块、残差增强块、宽增强块、特征细化块和构造块. 设计了引入移位卷积层的移位残差块 (SC-ResBlock), 减少模型参数量 and 计算成本. 构建了具有双分支 4 层并行网络结构的动态宽增强块, 并在其中嵌入了结合 BSCnv 和边缘检测算法的增强空间注意力模块 (IESA), 提取更丰富的信息, 增强所获得的超分辨率模型对不同应用场景的适用性. 最后选用组卷积和信道分割操作来降低计算复杂度, 提高模型重建速度. 通过大量实验结果表明, W-SDNet 不仅可以方便地获取丰富的结构信息, 而且具有轻量化的特点, 很好地平衡了图像重建性能和运行时间之间的矛盾. 未来希望在此模型的基础上实现任意尺度图像超分辨率.

参考文献

1 Yoo JS, Kim JO. Noise-robust iterative back-projection. IEEE Transactions on Image Processing, 2020, 29:

- 1219–1232. [doi: [10.1109/TIP.2019.2940414](https://doi.org/10.1109/TIP.2019.2940414)]
- 2 Dong C, Loy CC, He KM, *et al.* Learning a deep convolutional network for image super-resolution. Proceedings of the 13th European Conference on Computer Vision. Zurich: Springer, 2014. 184–199.
 - 3 Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 1646–1654.
 - 4 He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 770–778.
 - 5 Lim B, Son S, Kim H, *et al.* Enhanced deep residual networks for single image super-resolution. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu: IEEE, 2017. 1132–1140.
 - 6 Tai Y, Yang J, Liu XM, *et al.* MemNet: A persistent memory network for image restoration. Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017. 4549–4557.
 - 7 Hu J, Shen L, Sun G. Squeeze-and-excitation network. Proceeding of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.
 - 8 Zhang YL, Li KP, Li K, *et al.* Image super-resolution using very deep residual channel attention networks. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 294–310.
 - 9 Dai T, Cai JR, Zhang YB, *et al.* Second-order attention network for single image super-resolution. Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 11057–11066.
 - 10 Dosovitskiy A, Beyer L, Kolesnikov A, *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. Proceedings of the 9th International Conference on Learning Representations. OpenReview.net, 2021.
 - 11 Wang WH, Xie EZ, Li X, *et al.* PVT v2: Improved baselines with pyramid vision Transformer. Computational Visual Media, 2022, 8(3): 415–424. [doi: [10.1007/s41095-022-0274-8](https://doi.org/10.1007/s41095-022-0274-8)]
 - 12 肖子安, 张静, 苑馨方, 等. 基于多尺度特征提取和坐标注意力的光学遥感图像超分辨率重建. 计算机系统应用, 2023, 32(7): 105–112. [doi: [10.15888/j.cnki.csa.009168](https://doi.org/10.15888/j.cnki.csa.009168)]
 - 13 Liang JY, Cao JZ, Sun GL, *et al.* SwinIR: Image restoration using Swin Transformer. Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops. Montreal: IEEE, 2021. 1833–1844.
 - 14 Kim J, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 1637–1645.
 - 15 Tai Y, Yang J, Liu XM. Image super-resolution via deep recursive residual network. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 2790–2798.
 - 16 Ahn N, Kang B, Sohn KA. Fast, accurate, and lightweight super-resolution with cascading residual network. Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich: Springer, 2018. 256–272.
 - 17 Hui Z, Gao XB, Yang YC, *et al.* Lightweight image super-resolution with information multi-distillation network. Proceedings of the 27th ACM International Conference on Multimedia. Nice: ACM, 2019. 2024–2032.
 - 18 Liu J, Tang J, Wu GS. Residual feature distillation network for lightweight image super-resolution. Proceedings of the 2020 Workshops on Computer Vision. Glasgow: Springer, 2020. 41–55.
 - 19 王-一宁, 赵青杉, 秦品乐, 等. 基于轻量密集神经网络的医学图像超分辨率重建算法. 计算机应用, 2022, 42(8): 2586–2592.
 - 20 周登文, 王婉君, 马钰, 等. 基于区域互补注意力和多维注意力的轻量级图像超分辨率网络. 模式识别与人工智能, 2022, 35(7): 625–636. [doi: [10.16451/j.cnki.issn1003-6059.202207005](https://doi.org/10.16451/j.cnki.issn1003-6059.202207005)]
 - 21 Gao DD, Zhou DW. A very lightweight and efficient image super-resolution network. Expert Systems with Applications, 2023, 213: 118898. [doi: [10.1016/j.eswa.2022.118898](https://doi.org/10.1016/j.eswa.2022.118898)]
 - 22 陈豪, 夏振平, 程成, 等. 基于 Transformer-CNN 的轻量级图像超分辨率重建网络. 计算机应用, 2024, 44(1): 292–299.
 - 23 Shi YK, Zhong HY, Yang ZJ, *et al.* DDet: Dual-path dynamic enhancement network for real-world image super-resolution. IEEE Signal Processing Letters, 2020, 27: 481–485. [doi: [10.1109/LSP.2020.2978410](https://doi.org/10.1109/LSP.2020.2978410)]
 - 24 Wu HL, Ni N, Zhang LB. Lightweight stepless super-resolution of remote sensing images via saliency-aware dynamic routing strategy. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5601717.
 - 25 Chen BH, Lin MB, Sheng KK, *et al.* ARM: Any-time super-resolution method. Proceedings of the 17th European Conference on Computer Vision. Tel Aviv: Springer, 2022. 254–270.
 - 26 Haase D, Amthor M. Rethinking depthwise separable convolutions: How intra-kernel correlations lead to improved MobileNets. Proceedings of the 2020 IEEE/CVF Conference

- on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 14588–14597.
- 27 Chollet F. Xception: Deep learning with depthwise separable convolutions. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 1800–1807.
- 28 Hendrycks D, Gimpel K. Gaussian error linear units (GELUs). arXiv:1606.08415, 2016.
- 29 Wu G, Jiang JJ, Jiang K, *et al.* Fully 1×1 convolutional network for lightweight image super-resolution. arXiv: 2307.16140, 2023.
- 30 Szegedy C, Liu W, Jia YQ, *et al.* Going deeper with convolutions. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 1–9.
- 31 Kong FY, Li MX, Liu SW, *et al.* Residual local feature network for efficient super-resolution. Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New Orleans: IEEE, 2022. 765–775.
- 32 李众, 王雅婧, 马巧梅. 基于空洞卷积的医学图像超分辨率重建算法. 计算机应用, 2023, 43(9): 2940–2947.
- 33 Canny J. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, PAMI-8(6): 679–698. [doi: [10.1109/TPAMI.1986.4767851](https://doi.org/10.1109/TPAMI.1986.4767851)]
- 34 Wang Z, Bovik AC. Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE Signal Processing Magazine, 2009, 26(1): 98–117. [doi: [10.1109/MSP.2008.930649](https://doi.org/10.1109/MSP.2008.930649)]
- 35 Timofte R, Gu SH, Wu JQ, *et al.* NTIRE 2018 challenge on single image super-resolution: Methods and results. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018. 852–863.
- 36 Bevilacqua M, Roumy A, Guillemot C, *et al.* Low-complexity single-image super-resolution based on nonnegative neighbor embedding. Proceedings of the 2012 British Machine Vision Conference. Surrey: BMVA Press, 2012. 1–10.
- 37 Yang JC, Wright J, Huang TS, *et al.* Image super-resolution via sparse representation. IEEE Transactions on Image Processing, 2010, 19(11): 2861–2873. [doi: [10.1109/TIP.2010.2050625](https://doi.org/10.1109/TIP.2010.2050625)]
- 38 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. Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV). Vancouver: IEEE, 2001. 416–423.
- 39 Huang JB, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE, 2015. 5197–5206.
- 40 Aizawa K, Fujimoto A, Otsubo A, *et al.* Building a manga dataset “Manga109” with annotations for multimedia applications. IEEE MultiMedia, 2020, 27(2): 8–18. [doi: [10.1109/MMUL.2020.2987895](https://doi.org/10.1109/MMUL.2020.2987895)]
- 41 Cai JR, Zeng H, Yong HW, *et al.* Toward real-world single image super-resolution: A new benchmark and a new model. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 3086–3095.
- 42 Tian CW, Zhang XY, Zhang Q, *et al.* Image super-resolution via dynamic network. arXiv:2310.10413, 2023.
- 43 Dong C, Loy CC, Tang XO. Accelerating the super-resolution convolutional neural network. Proceedings of the 4th European Conference on Computer Vision. Amsterdam: Springer, 2016. 391–407.
- 44 Lai WS, Huang JB, Ahuja N, *et al.* Deep Laplacian pyramid networks for fast and accurate super-resolution. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 5835–5843.
- 45 欧阳宁, 韦羽, 林乐平. 结合感知边缘约束与多尺度融合网络的图像超分辨率重建方法. 计算机应用, 2020, 40(10): 3041–3047.
- 46 Tian CW, Xu Y, Zuo WM, *et al.* Asymmetric CNN for image superresolution. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2022, 52(6): 3718–3730. [doi: [10.1109/TSMC.2021.3069265](https://doi.org/10.1109/TSMC.2021.3069265)]
- 47 Wei RJ, Zhang SW, Liu ZC, *et al.* EBSR: Enhanced binary neural network for image super-resolution. arXiv:2303.12270, 2023.
- 48 Shocher A, Cohen N, Irani M. “Zero-shot” super-resolution using deep internal learning. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 3118–3126.
- 49 Wang LG, Wang YQ, Dong XY, *et al.* Unsupervised degradation representation learning for blind super-resolution. Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021. 10576–10585.
- 50 Zhang R, Isola P, Efros AA. Split-brain autoencoders: Unsupervised learning by cross-channel prediction. Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017. 645–654.

(校对责编: 孙君艳)