

基于对偶回归和注意力机制的图像超分辨率重建网络^①



印珏泽, 周宁宁

(南京邮电大学 计算机学院, 南京 210023)
通信作者: 印珏泽, E-mail: 1336885928@qq.com

摘要: 针对单幅图像超分辨率 (single image super-resolution, SISR) 重建算法存在低分辨率图像 (LR) 到高分辨率图像 (HR) 的映射学习具有不适定性, 深层神经网络收敛慢且缺乏对高频信息的学习能力以及在深层神经网络传播过程中图像特征信息存在丢失的问题. 本文提出了基于对偶回归和残差注意力机制的图像超分辨率重建网络. 首先, 通过对偶回归约束映射空间. 其次, 融合通道和空间注意力机制构造了残差注意力模块 (RCSAB), 加快模型收敛速度的同时, 有效增强了对高频信息的学习. 最后, 融入密集特征融合模块, 增强了特征信息流动性. 在 Set5、Set14、BSD100、Urban100 四种基准数据集上与目前主流的单幅图像超分辨率算法进行对比, 实验结果表明该方法无论是在客观质量评价指标还是主观视觉效果均优于对比算法.

关键词: 单幅图像超分辨率; 通道注意力; 空间注意力; 对偶回归; 密集特征融合

引用格式: 印珏泽, 周宁宁. 基于对偶回归和注意力机制的图像超分辨率重建网络. 计算机系统应用, 2023, 32(2): 111-118. <http://www.c-s-a.org.cn/1003-3254/8939.html>

Image Super-resolution Reconstruction Network Based on Dual Regression and Attention Mechanism

YIN Jue-Ze, ZHOU Ning-Ning

(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: The single image super-resolution (SISR) reconstruction algorithm is ill-posed in the mapping learning from low-resolution (LR) image to high-resolution (HR) image, and the deep neural network has slow convergence and lacks the ability to learn high-frequency information. Moreover, image feature information tends to be lost during deep neural network propagation. In order to address these issues, this study proposes an image super-resolution reconstruction network based on dual regression and residual attention mechanism. Firstly, the mapping space is constrained by dual regression. Secondly, a residual attention module (RCSAB) is constructed by combining channel and spatial attention mechanisms, which not only accelerates the model convergence speed and effectively strengthens the learning of high-frequency information. Finally, a dense feature fusion module is introduced to enhance the fluidity of feature information. In addition, a comparison with the mainstream SISR algorithms is carried out on four benchmark datasets, namely, Set5, Set14, BSD100, and Urban100, and experimental results demonstrate that the proposed method is superior to other algorithms in terms of objective evaluation metrics and subjective visual effects.

Key words: single image super-resolution; channel attention; spatial attention; dual regression; dense feature fusion

1 引言

图像分辨率作为评估图像质量的重要指标, 其决

定了图像纹理细节的精细度. 图像分辨率越高, 意味着图像的纹理细节更清晰, 越有利于后续对图像的分析

① 收稿时间: 2022-06-20; 修改时间: 2022-07-18, 2022-07-29; 采用时间: 2022-08-15; csa 在线出版时间: 2022-10-28
CNKI 网络首发时间: 2022-11-15

与处理.然而,受限于硬件设备、外部环境、压缩技术等客观条件,无法获得分辨率非常高的图像,图像超分辨率技术能有效地解决上述问题.图像超分辨率技术旨在将低分辨率图像恢复为高分辨率图像,近年来,已被广泛应用于公安安防^[1]、医学成像^[2]和卫星遥感^[3]等现实领域,因此对图像超分辨率技术研究具有显著的意义.图像超分辨率技术可以分为单幅图像超分辨率重建、多帧图像超分辨率重建和视频超分辨率重建3类,本文主要研究的是单幅图像超分辨率重建技术.

传统单幅图像超分辨率算法可以分为3类.

(1) 基于插值的算法,是通过利用核函数根据已知邻近像素的值,计算出待插入像素的值.最常用插值算法为双三次(Bicubic)^[4]插值算法,当放大倍数较小时,利用Bicubic插值算法高效且稳定,但是放大倍数较大时,重建图像往往会因为细节的丢失存在明显的模糊失真问题.

(2) 基于重构的方法,通过利用图像的一些先验知识,例如图像的局部平滑性、非局部自相似性、稀疏性等特征来构成约束条件,然后迭代优化估计出最接近低分辨率图像成像模型的逆过程,从而重建高分辨率图像.常见的方法有最大后验概率法^[5]、迭代反向投影法^[6]、凸集投影法^[7]等,相比于插值法,重构法引入了更多额外信息获得了更好的重构效果,但是对先验知识利用的不充分或者无法获得足够的先验知识,所重建的图像存在平滑过度的问题.

(3) 基于学习的方法,通过利用提前在训练集上训练学习到的低分辨率图像与高分辨率图像间的映射关系来重建高分辨率图像.常见的有领域嵌入^[8]、流形学习^[9]、稀疏表示^[10]等方法,在对人脸、文本等图像重建,这些方法表现出良好的重构效果,但是其特征提取过程依赖于人工设计,十分复杂.

近年来,基于深度学习的方法通过卷积神经网络能够提取丰富的特征,取得了很好的重建效果,但目前大多数的超分模型注重加深网络而忽略了对特征信息的区别处理,使得模型无法有效地恢复图像的纹理细节等高频信息.同时,随着网络深度的增加,特征信息容易在向后传播过程中被削弱或丢失.此外,学习LR到HR的映射函数通常是一个不适定问题,存在无穷多的HR图像可以下采样到相同的LR图像.因此,映射函数的解空间可能非常大,学习到一个性能优异的映射函数非常困难.

本文提出的基于对偶回归和残差注意力机制的图像超分辨率重建网络.通过对偶回归学习方法,在原始回归任务基础上引入额外约束,缩小映射空间.融合通道和空间注意力机制构造的残差注意力模块能够对不同特征通道信息和相同特征通道不同空间位置信息进行自适应调节,使网络专注于高频信息的学习.此外,融入密集特征融合模块,增强了特征信息的流动性,使特征信息在网络传播过程中更不易丢失.

2 相关工作

2.1 基于深度学习的图像超分辨率技术

近年来,随着深度学习技术的快速发展,利用深度学习技术的图像超分辨率研究取得了显著进展.2015年Dong等人^[11]提出SRCNN(super resolution convolutional neural network),首次将卷积神经网络(convolutional neural network, CNN)应用到图像超分领域,结合稀疏编码的方法让CNN网络直接学习HR与LR间的端到端映射,取得了不错的重构效果.随后,Dong等人^[12]对SRCNN做出改进,提出FSRCNN(Fast-SRCNN)模型.FSRCNN引入反卷积层,不需要对原始低分辨率输入图像进行双三次插值,减少了计算量和输入误差且能够通过改变反卷积层实现不同尺度放大,同时改变了卷积核大小和网络深度,使FSRCNN模型获得了更快的收敛速度和更好的重构效果.Lim等人^[13]证明了批处理归一化层(batch normalization, BN)丢失了图像的尺度信息,于是去除了残差块中的BN层提出了EDSR(enhanced deep super-resolution)模型,节省内存资源的同时在相同的计算资源下能堆叠更多的网络层,并在每一层提取更丰富的特征信息,从而提高重建图像质量.Ledig等人^[14]首次将生成对抗网络应用到超分辨率重建领域,提出超分辨率生成对抗网络(super-resolution using a generative adversarial network),同时提出一种感知损失函数代替均方误差(mean square error, MSE)损失函数,通过对抗训练,重构得到的图像更符合人类视觉感知.Haris等人^[15]基于迭代反投影方法提出了DBPN(deep back-projection networks)模型,利用迭代升降采样,为每个阶段提供误差反馈机制.Zhang等人^[16]提出了RCAN(residual channel attention network)模型,提出了一种通道注意力机制来自适应调节通道特征.Dai等人^[17]提出了SAN(secondary-order attention network),通过使用二阶特征统计量对图像特征信息进行自适应

提取。

2.2 通道注意力机制

大多数基于 CNN 的 SR 模型对 LR 图像的特征提取过程存在以下两个问题。

(1) 不同特征通道所包含的信息存在一定差异, 采用平等处理的方式不利于重建图像高频细节。

(2) 卷积层中每个滤波器的卷积核仅能使用一个特征通道的局部感受野, 无法利用其他通道的感受野信息。

针对这两个问题, 通道注意力机制利用特征通道之间的相互依赖关系来自适应地对特征通道的权重进行调整, 其结构如图 1 所示。

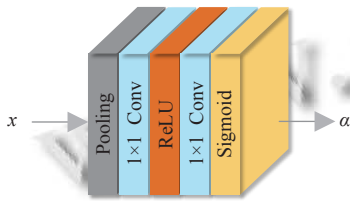


图 1 通道注意力模块

$x = [x_1, x_2, \dots, x_c] \in R^{H \times W \times C}$ 为通道注意力模块的输入, 首先, 利用全局平均池化将通道的全局空间信息压缩为一个通道向量 z , 计算方式如式 (1) 所示:

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

其中, z_c 为通道向量 z 中第 c 个元素, $x_c(i, j)$ 为第 c 个特征图在 (i, j) 处的像素值, H_{GP} 表示全局池化函数。

经过池化层后, 利用 Sigmoid 激活函数来获取不同特征通道间的依赖关系, 计算方式如式 (2) 所示:

$$\alpha = f(W_U \delta(W_D z)) \quad (2)$$

其中, f 和 δ 分别代表 Sigmoid 和 ReLU 激活函数, W_D 是起通道降维作用的卷积层权重, W_U 是经过 ReLU 激活后起通道升维的卷积层权重. α 为表示特征通道相互关系的统计量。

最后, 利用 α 对 x 进行调整, 输出结果如式 (3) 所示:

$$U_{CA}(x) = \alpha \cdot x \quad (3)$$

其中, $U_{CA}(x)$ 为通道注意力模块的输出。

2.3 空间注意力机制

同一特征通道所包含的信息因空间位置的不同而不同. 例如, 边缘或纹理区域通常含有更多的高频信息,

而平滑区域则含有更多的低频信息. 为了恢复图像的高频细节, 应该对高频区域给予更多的关注. 受此启发, 空间注意力模块将同一特征通道的低频信息与高频信息分离至不同的滤波器, 使得网络能更专注的学习高频信息, 重建图像具有丰富的高频细节. 其结构如图 2 所示。

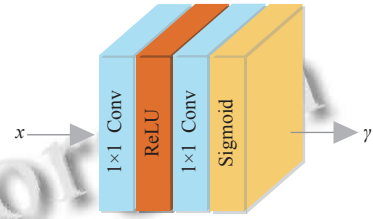


图 2 空间注意力模块

同样使用 $x = [x_1, x_2, \dots, x_c] \in R^{H \times W \times C}$ 作为空间注意力模块的输入. 首先, 通过 1×1 卷积层融合特征信息, 接着, 通过 ReLU 激活函数和 1×1 卷积层学习空间权重, 再利用 Sigmoid 激活函数获得空间向量 γ , 计算方式如式 (4) 所示:

$$\gamma = \sigma(W_{SA}^2 \delta(W_{SA}^1(x))) \quad (4)$$

其中, W_{SA}^1 和 W_{SA}^2 分别代表两个 1×1 的卷积层权重, δ 为 ReLU 激活函数, σ 代表 Sigmoid 激活函数。

最后, 利用 γ 对 x 进行调整, 输出结果如式 (5) 所示:

$$U_{SA}(x) = x \cdot \gamma \quad (5)$$

其中, $U_{SA}(x)$ 为空间注意力模块的输出。

3 本文算法

3.1 网络框架

本文所设计的网络总体结构基于 U-Net 网络^[18], 网络结构图如图 3 所示, 整体网络结构分为两个部分: 原始回归网络和对偶回归网络, 原始回归网络是一种渐进上采样结构, 首先经过 $\log_2(s)$ 次下采样, 进行初始特征提取, s 表示超分辨率的比例因子, 接着经过 $\log_2(s)$ 次上采样, 进行深层特征提取, 上采样过程中通过融合下采样过程生成的同尺寸特征图, 再次进行特征提取, 有效增强了网络的学习能力. 最后将重建的超分辨率图像输入至对偶回归网络, 对偶回归网络将重建的超分辨率图像进行下采样操作后与原始 LR 图像进行对比. 该过程为 LR-HR 的映射学习引入了额外约束, 有效约束了映射空间。

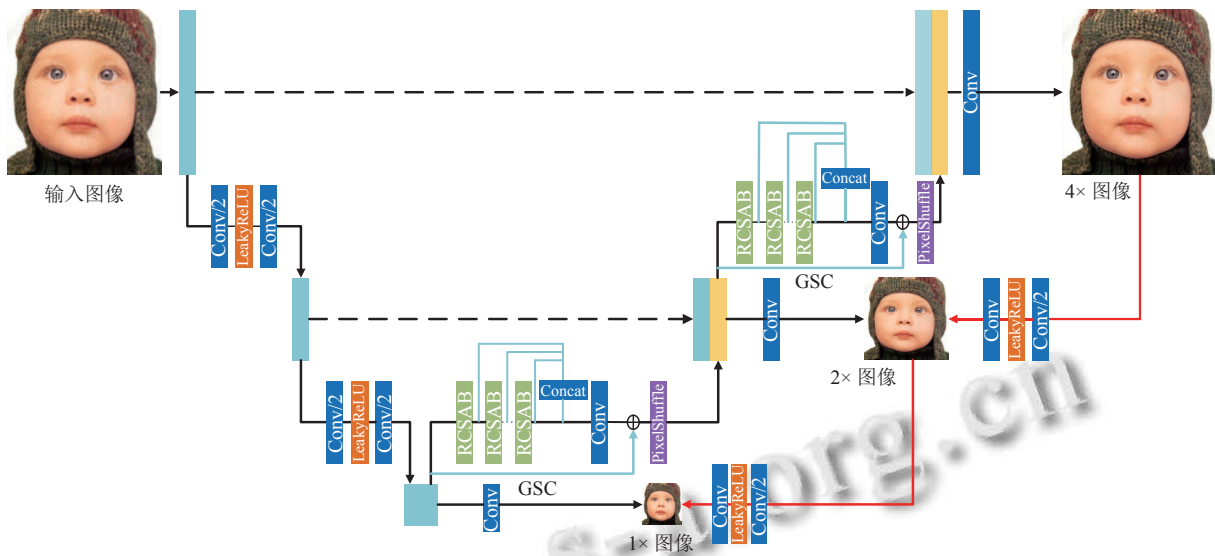


图3 整体网络结构

3.2 原始回归网络

原始回归网络(图3中的黑线部分)由上采样和下采样两个模块组成,下采样基本块由步长为2的3×3卷积层、LeakyReLU激活层和步长为1的3×3卷积层构成,用于初始特征提取,上采样基本块由B个残差注意力模块、密集特征融合模块和亚像素卷积层组成,其中残差注意力模块对图像进行深层特征提取,密集特征融合模块有效解决了随着网络深度的增加导致原始低分辨率图像的特征信息丢失问题.亚像素卷积层通过将特征图通道数扩充为原来的4倍,再通过像素重洗将分辨率扩大为原来的2倍.

3.3 对偶回归网络

对偶回归^[19]学习在语言翻译任务中首次被提出,简单有效地提高了翻译效果.随后应用于图像翻译^[20]、图像处理^[21]等监督学习任务中,学习模型的性能均得到一定程度的提升.

对偶回归学习通过对偶模型学习与原始模型的相反映射,使得原始模型和对偶模型形成一个闭环,二者将生成的信息互相反馈来进行训练.模型框架如图4所示.

在图像超分辨率任务中,原始模型用于预测生成SR图像,对偶模型用于对生成的SR图像进行下采样,利用对偶模型生成的图像X'与原始输入的低分辨率图像X的偏差对原始模型进行参数调节,通过引入这一额外约束,来减少可能的LR到HR的映射空间,从而学习到更优的映射函数.

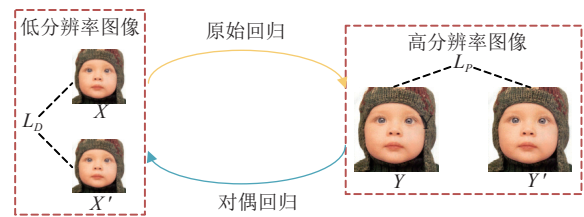


图4 对偶回归模型

对偶回归网络(图3中的红线部分),主要任务是学习下采样操作,原始模型利用原始任务训练过程中学习到的信息和对偶模型提供的反馈信息迭代优化.对偶模型的下采样基本块和原始模型下采样基本块结构相同,由步长为2的3×3卷积层、LeakyReLU激活层和步长为1的3×3卷积层构成.

3.4 融合通道空间注意力的残差模块

随着网络层数的加深,网络可以进行更加复杂的特征提取,但同时伴随着模型容易过拟合和梯度消失与爆炸等问题的出现,针对这些问题,He等人^[22]提出残差网络结构,由于残差网络结构所需要学习的特征映射只是输入与输出的残差,与不引入残差学习的网络结构所要学习特征映射相比更加简单,大大提升了深层次网络模型的性能.

残差学习最初是针对分类、检测等高级计算机视觉任务提出,直接应用超分辨率等低级计算机视觉任务效果不佳.本文基于EDSR模型中使用的残差模块并融合了通道和空间注意力机制,构建了RCSAB,有效提取特征图像中的高频信息,其结构如图5所示.

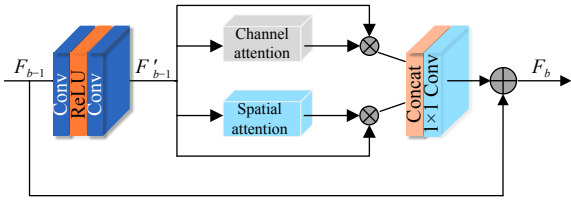


图5 RCSAB 模块

第 b 个 RCSAB 的输出结果如式 (6) 表示:

$$F_b = H_b(F_{b-1}) \quad (6)$$

其中, F_{b-1} 和 F_b 分别第 b 个 RCSAB 的输入和输出, H_b 表示第 b 个 RCSAB 函数, 具体操作细节如下。

首先, 将 F_{b-1} 依次通过 1×1 卷积层, ReLU 激活层和 1×1 卷积层来进行特征提取, 得到 F'_{b-1} , 计算方式如式 (7) 所示:

$$F'_{b-1} = W_R^2 \delta(W_R^1(F_{b-1})) \quad (7)$$

其中, W_R^1 和 W_R^2 分别代表两个 1×1 的卷积层权重, δ 为 ReLU 激活层。

其次, 将 F'_{b-1} 分别通过空间注意力模块和通道注意力模块得到两组特征图, 将两组特征图融合再输入至 1×1 卷积层, 该卷积层用于自适应融合两种注意力机制所生成的特征图。最后, 与原始输入相加得到 F_b , 计算方式如式 (8) 所示:

$$F_b = W_c [U_{CA}(F'_{b-1}), U_{SA}(F'_{b-1})] + F_{b-1} \quad (8)$$

其中, W_c 代表自适应融合两种注意力机制所输出特征的 1×1 卷积层权重, U_{CA} 和 U_{SA} 分别代表通道注意力模块和空间注意力模块。

3.5 密集特征融合模块

经过观察, 通过简单地堆叠更多地残差注意力模块不能获得更好的重建性能, 为了解决这一问题, 设计了一种密集特征融合模块。其结构如图 6 所示, 该模块通过融合所有特征信息来提取全局特征, 有效地解决了特征信息在传播过程中的丢失问题, 提高了网络学习效率。

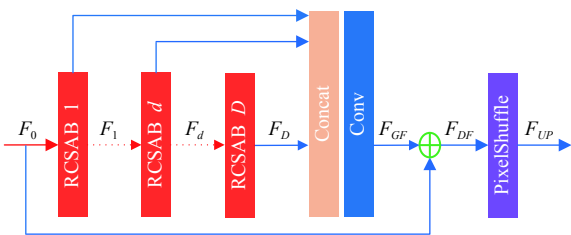


图6 密集特征融合模块

首先, 将所有 RCSAB 的输出进行拼接融合, 对融合后的特征进行特征提取获得全局特征 F_{GF} , 计算方式如式 (9) 所示:

$$F_{GF} = W_{GF}([F_1, \dots, F_d, \dots, F_D]) \quad (9)$$

其中, $[F_1, \dots, F_d, \dots, F_D]$ 表示对 D 个 RCSAB 生成的特征图的融合。 W_{GF} 代表一个 3×3 卷积层权重。

接着, 经过全局残差连接获得最终输入至亚像素卷积层的特征 F_{DF} , 计算方式如式 (10) 所示:

$$F_{DF} = F_0 + F_{GF} \quad (10)$$

其中, F_0 表示初始输入特征图。

3.6 损失函数

损失函数主要由原始重建网络的损失及对偶回归网络的损失两部分组成。给定一组 N 对样本 $S_p = \{(x_i, y_i)\}_{i=1}^N$, 其中 (x_i, y_i) 表示这组配对数据中的第 i 对低分辨率和高分辨率图像。训练损失 L 如式 (11) 所示:

$$L = \sum_{i=1}^N L_P\{P(x_i), y_i\} + \lambda L_D\{D(P(x_i)), x_i\} \quad (11)$$

其中, $P(x_i)$ 为初始模型预测的 SR 图像, $D(P(x_i))$ 为对偶模型经过下采样得到的低分辨率图像。 L_P 和 L_D 分别为原始重建网络损失和对偶回归网络损失。 λ 为控制对偶回归损失占比的权重。经过多次对比实验, 训练过程中将 λ 值设为 0.1, 具体对比实验结果将在实验部分给出。

4 实验分析

4.1 数据集和评价指标

实验采用 DIV2K 和 Flickr2K 数据集组成的 DF2K 数据集作为训练集。通过随机截取 RGB 输入低分辨率图像的 48×48 大小的图像块和对应的高分辨率图像块配对作为训练数据。测试集采用 Set5, Set14, BSD100, Urban100 此外, 本文使用数据增强技术^[23], 将训练数据随机旋转 90° , 180° , 270° 以及平移和翻转。

评价指标采用单幅图像超分辨率方法常用的两种客观评价指标, 分别是峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似性 (structure similarity, SSIM)。

PSNR 是有损变换中最常用的重构质量度量指标之一。对于图像的超分辨率, PSNR 是通过图像间的最大像素值和均方误差来定义, 计算方式如式 (12) 所示:

$$PSNR = 10 \cdot \lg \left(\frac{L^2}{MSE} \right) \quad (12)$$

其中, L 为最大像素值, MSE 是原始高分辨率图像与生成的高分辨率图像间的均方误差. $PSNR$ 数值越大, 表明生成图像与原始图像间的像素差值越小, 重建效果越好.

$SSIM$ 基于亮度、对比度和结构的比较, 计算方式如式 (16) 所示:

$$C_l(U, V) = \frac{2\mu_U\mu_V + C_1}{\mu_U^2 + \mu_V^2 + C_1} \quad (13)$$

$$C_c(U, V) = \frac{2\sigma_U\sigma_V + C_2}{\sigma_U^2 + \sigma_V^2 + C_2} \quad (14)$$

$$C_s(U, V) = \frac{\sigma_{UV} + C_3}{\sigma_U\sigma_V + C_3} \quad (15)$$

$$SSIM(U, V) = C_l(U, V) \times C_c(U, V) \times C_s(U, V) \quad (16)$$

其中, C_l 为亮度比较, C_c 为对比度比较, C_s 为结构比较, μ_U 和 μ_V 分别表示两幅对比图像的像素均值, σ_U 和 σ_V 分别表示两幅对比图像的标准差, σ_{UV} 表示两幅对比图像的协方差. $SSIM$ 反映两幅图像的结构相似性, 其数值越接近于 1, 表明重建效果越好.

4.2 模型细节

本文的实验基于 PyTorch 框架, 使用 2 块 NVIDIA GeForce 2080Ti 显卡进行网络训练, 每块显存容量为 10 GB. 优化器采用 Adam 优化器, $\beta_1 = 0.9$, $\beta_2 = 0.999$,

表 2 残差注意力模块与密集特征融合模块对照实验数据

对照项	Baseline	1	2	3	4	5	6	7
通道注意力 (CA)	×	√	×	×	×	√	×	×
空间注意力 (SA)	×	×	√	×	×	×	√	×
CA+SA	×	×	×	√	×	×	×	√
密集特征融合模块	×	×	×	×	√	√	√	√
参数量 (M)	0.51	4.80	2.26	5.55	2.72	7.01	4.47	7.76
$PSNR$ (dB)	32.39	32.48	32.45	32.63	32.52	32.60	32.58	32.75

从表 2 前 4 列或后 4 列可以看出, 融合了通道与空间注意力模块的残差模块测得的 $PSNR$ 数值最高, 组 3 相比于 Baseline 提升了 0.24 dB, 组 7 相比于组 4 提升了 0.23 dB, 证明了本文的残差注意力模块的有效性. 对比前 4 列和后 4 列可以看出, 融入密集特征融合模块后, 对于×4 模型虽然参数量提升了 2.21 M, 但是 $PSNR$ 值均得到了很大提升, 以组 7 和组 3 为例, 组 7 相较于组 3 的 $PSNR$ 值提升了 0.12 dB, 验证了密集特征融合模块的有效性.

4.4 与已有算法的对比

在相同实验设置下选取了 Bicubic、ESPCN、

batchsize 设为 32. 初始学习率被设定为 10^{-4} , 每 100 个 epochs 学习率下降一半. 在网络训练时, 将训练图像转换为 YCbCr 空间, 并只对 Y 通道做处理. 在对比实验中, 对于 4 倍尺度放大, 使用 30 个残差注意力模块, 特征通道数为 16, 对于 8 倍尺度放大, 使用 30 个残差注意力模块, 特征通道数为 8

4.3 消融实验分析

为了研究对偶回归损失权重对网络模型性能的影响, 在×4 模型中分别将 λ 设为 0.001, 0.01, 0.1, 1, 10 进行对照实验, 实验结果如表 1 所示.

表 1 λ 对×4 模型的性能影响

λ	0.001	0.01	0.1	1	10
Set5- $PSNR$ (dB)	32.62	32.65	32.73	32.59	32.48

从表 1 可以看出, λ 从 0.001 到增加到 0.1 时, 对偶回归损失有效地增强了模型的监督能力, 但是, 当 λ 进一步从 0.1 到增加到 10 时, 对偶回归网络损失对模型起负反馈作用, 抑制了模型地重建性能. 为了平衡原始回归损失与对偶回归损失, 本文将 λ 设置为 0.1.

为了研究残差注意力模块和密集特征融合模块的有效性, 将残差模块融合不同注意力模块和密集特征融合模块进行对照实验. 实验数据如表 2 所示. 其中 Baseline 为没有融合任何模块的残差模块.

SRCNN、FSRCNN、SRGAN、EDSR、DBPN、RCAN、SAN 与本文算法进行对比, 在 Set5、Set14、BSD100、Urban100 四种基准数据集上进行测试, 放大倍数分别为: ×4, ×8. 实验结果如表 3 所示.

从表 3 可以看出, 在 4 倍尺度放大时, 本文算法所重建图像的客观评价指标具有明显优势, 在 8 倍尺度放大时, 虽然由于尺度因子的增加各算法性能均有所下降, 但从客观结果表明, 本文算法性能仍优于其他算法.

同样从主观视觉角度进行对比, 本文分别从 Set5、Set14 以及 Urban100 数据集选取了 3 张图片并对特定区域进行放大, 结果如图 7-图 9 所示.

表3 不同算法的 PSNR 和 SSIM 数值比较

方法	尺寸	Set5		Set14		BSD100		Urban100	
		PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM
Bicubic	×4	28.42	0.810	26.10	0.702	25.96	0.667	23.15	0.657
ESPCN		29.21	0.851	26.40	0.744	25.50	0.696	24.02	0.726
SRCNN		30.48	0.862	27.50	0.751	26.90	0.710	24.52	0.722
FSRCNN		30.72	0.866	27.61	0.755	26.98	0.715	24.62	0.728
SRGAN		29.46	0.838	26.60	0.718	25.74	0.666	24.50	0.736
EDSR		32.48	0.898	28.81	0.787	27.72	0.742	26.64	0.803
DBPN		32.42	0.897	28.75	0.786	27.67	0.739	26.38	0.794
RCAN		32.63	0.900	28.85	0.788	27.74	0.743	26.74	0.806
SAN		32.64	0.900	28.92	0.788	27.79	0.743	26.79	0.806
本文算法		32.75	0.902	28.96	0.791	27.81	0.745	26.95	0.808
Bicubic	×8	24.39	0.657	23.19	0.568	23.67	0.547	20.74	0.515
ESPCN		25.02	0.697	23.45	0.598	23.92	0.574	21.20	0.554
SRCNN		25.33	0.690	23.76	0.591	24.13	0.566	21.29	0.544
FSRCNN		20.13	0.552	19.75	0.4820	24.21	0.568	21.32	0.538
SRGAN		23.04	0.626	21.57	0.495	21.78	0.442	19.64	0.468
EDSR		27.03	0.774	25.05	0.641	24.80	0.595	22.55	0.618
DBPN		27.25	0.786	25.14	0.649	24.90	0.602	22.72	0.631
RCAN		27.31	0.787	25.23	0.651	24.96	0.605	22.97	0.643
SAN		27.22	0.782	25.14	0.647	24.88	0.601	22.70	0.631
本文算法		27.52	0.798	25.26	0.655	25.01	0.608	22.98	0.646

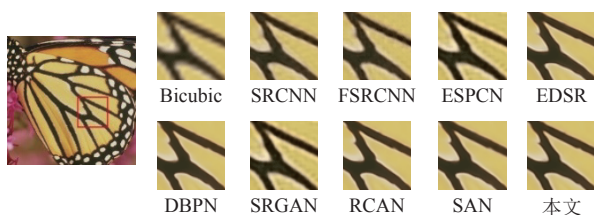


图7 “Set5-butterfly” ×4 主观视觉效果对比

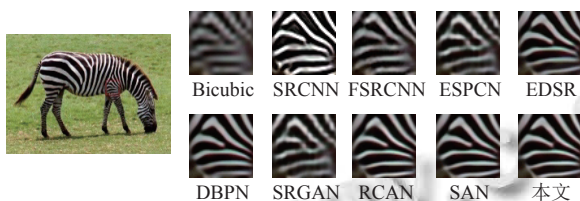


图8 “Set14-zebra” ×4 主观视觉效果对比

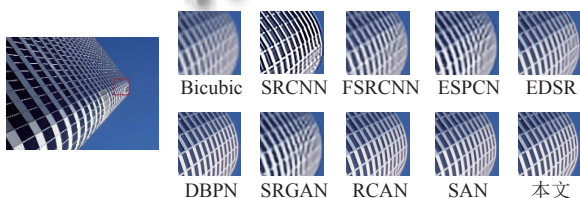


图9 “Urban100-Img005” ×4 主观视觉效果对比

从图7和图8看出, 本文方法构建的图像相较于其他方法更加清晰, 色彩与纹理细节更贴近真实图像. 此外, 观察图9可以看出, 相较于 Bicubic 插值算法,

SRCNN 算法准确地预测了窗沿方向, 但存在严重的伪影问题, FSRCNN 有效减轻了伪影现象, 但混淆了窗沿平行线方向, 其他对比算法也均存在不同程度的混淆现象与伪影问题, 相比较而言, 本文算法预测的窗沿方向最准确, 同时有效解决了伪影问题.

5 结论与展望

本文设计了一种基于对偶回归和残差注意力机制的单幅图像超分辨率网络结构, 对偶回归学习策略为 LR-HR 的映射学习引入了额外约束, 减少了可能的映射空间. 通过融合通道注意力模块和空间注意力模块至残差模块中对特征信息进行全局和局部调节, 有效提取了特征信息中的高频信息. 最后, 密集特征融合模块增强了特征信息流动性, 有效解决了网络传播过程中特征信息的丢失问题. 实验结果表明, 无论从客观指标评价还是主观视觉效果对比, 本文算法的重建效果均优于对比算法.

参考文献

1 Kouame D, Ploquin M. Super-resolution in medical imaging: An illustrative approach through ultrasound. Proceedings of 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. Boston: IEEE, 2009. 249-252.

- 2 Masutani EM, Bahrami N, Hsiao A. Deep learning single-frame and multiframe super-resolution for cardiac MRI. *Radiology*, 2020, 295(3): 552–561. [doi: [10.1148/radiol.2020192173](https://doi.org/10.1148/radiol.2020192173)]
- 3 Zhang L, Nie JT, Wei W, *et al.* Deep blind hyperspectral image super-resolution. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 32(6): 2388–2400.
- 4 Keys R. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1981, 29(6): 1153–1160. [doi: [10.1109/TASSP.1981.1163711](https://doi.org/10.1109/TASSP.1981.1163711)]
- 5 Irani M, Peleg S. Motion analysis for image enhancement: Resolution, occlusion, and transparency. *Journal of Visual Communication and Image Representation*, 1993, 4(4): 324–335. [doi: [10.1006/jvci.1993.1030](https://doi.org/10.1006/jvci.1993.1030)]
- 6 Yang X, Zhang Y, Zhou DK, *et al.* An improved iterative back projection algorithm based on ringing artifacts suppression. *Neurocomputing*, 2015, 162: 171–179. [doi: [10.1016/j.neucom.2015.03.055](https://doi.org/10.1016/j.neucom.2015.03.055)]
- 7 Stark H, Oskoui P. High-resolution image recovery from image-plane arrays, using convex projections. *Journal of the Optical Society of America A*, 1989, 6(11): 1715–1726. [doi: [10.1364/JOSAA.6.001715](https://doi.org/10.1364/JOSAA.6.001715)]
- 8 Rahiman VA, George SN. Single image super resolution using neighbor embedding and statistical prediction model. *Computers & Electrical Engineering*, 2017, 62: 281–292.
- 9 江俊君. 基于一致流形学习的人脸超分辨率算法研究 [博士学位论文]. 武汉: 武汉大学, 2014.
- 10 Yang JC, Wright J, Huang T, *et al.* Image super-resolution as sparse representation of raw image patches. *Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage: IEEE, 2008. 1–8.
- 11 Dong C, Loy CC, He KM, *et al.* Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295–307. [doi: [10.1109/TPAMI.2015.2439281](https://doi.org/10.1109/TPAMI.2015.2439281)]
- 12 Dong C, Loy CC, Tang XO. Accelerating the super-resolution convolutional neural network. *Proceedings of 14th European Conference on Computer Vision*. Cham: Springer, 2016. 391–407.
- 13 Lim B, Son S, Kim H, *et al.* Enhanced deep residual networks for single image super-resolution. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu: IEEE, 2017. 136–144.
- 14 Ledig C, Theis L, Huszár F, *et al.* Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 105–114.
- 15 Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 1664–1673.
- 16 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 (ECCV)*. Munich: Springer, 2018. 294–310.
- 17 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.
- 18 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. *Proceedings of 18th International Conference on Medical Image Computing and Computer-assisted Intervention*. Cham: Springer, 2015. 234–241.
- 19 Zhu S, Cao RS, Yu K. Dual learning for semi-supervised natural language understanding. *IEEE/ACM Transactions on Audio, Speech and Language Processing*, 2020, 28: 1936–1947.
- 20 Zhu JY, Park T, Isola P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks. *Proceedings of the 2017 IEEE International Conference on Computer Vision*. Venice: IEEE, 2017. 2242–2251.
- 21 Wang L, Li D, Zhu YS, *et al.* Dual super-resolution learning for semantic segmentation. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 3773–3782.
- 22 He K, Zhang X, Ren S, *et al.* Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016. 770–778.
- 23 Li Z, Yang JL, Liu Z, *et al.* Feedback network for image super-resolution. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 3862–3871.

(校对责编: 牛欣悦)