

# 非抽取小波边缘学习深度残差网络的单幅图像超分辨率重建

王相海<sup>1,2</sup>, 赵晓阳<sup>1</sup>, 王鑫莹<sup>2</sup>, 赵克云<sup>2</sup>, 宋传鸣<sup>2</sup>

(1. 辽宁师范大学地理科学学院, 辽宁大连 116029; 2. 辽宁师范大学计算机科学与技术学院, 辽宁大连 116081)

**摘 要:** 图像超分辨率重建作为一个典型的非适定问题一直受到重视, 尽管近年来出现了许多行之有效的卷积神经网络超分辨率重建模型, 但如何全面挖掘图像先验信息, 用以提高重建图像的细节清晰度仍有待深入研究. 本文提出一种基于非抽取 Wavelet 变换的边缘学习深度残差网络单幅图像超分辨重建模型 NDW-EDRN (Non-Decimated Wavelet Edge learning using Deep Residual Networks), 在图像经非抽取 Wavelet 变换后获得多冗余信息、平滑及梯度值较小的低频区域和边缘及梯度值较大的高频区域的基础上, 将整体网络框架设计为采用不同结构的 CNN (Convolutional Neural Networks) 模型来对低频子带与高频子带分别进行学习的策略: 对低频子带采用稠密跳跃连接的方式整体性学习低频子带间的映射关系; 对高频子带采用一种新型的 U-net 模型, 将图像退化过程中所丢失的边缘作为网络的期望输出, 通过基于块的跳跃连接来使网络更精细地学习缺失性边缘, 从而更加充分、有效地获取图像在退化过程中所丢失的边缘细节信息. 大量实验结果表明, 该网络模型能够有效提高重建图像的质量, 特别在恢复低分辨率图像的边缘信息方面具有一定的优势, 在一定程度上弥补了传统 CNN 网络模型捕捉图像细节信息的不足.

**关键词:** 卷积神经网络; 残差学习; 非抽取小波变换; 图像超分辨率重建; 纹理边缘信息

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 0372-2112(2022)07-1753-13

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

**DOI:** 10.12263/DZXB.20210854

## Single Image Super-Resolution Reconstruction Using Deep Residual Networks with Non-decimated Wavelet Edge Learning

WANG Xiang-hai<sup>1,2</sup>, ZHAO Xiao-yang<sup>1</sup>, WANG Xin-ying<sup>2</sup>, ZHAO Ke-yun<sup>2</sup>, SONG Chuan-ming<sup>2</sup>

(1. School of Geographical Science, Liaoning Normal University, Dalian, Liaoning 116029, China;

2. School of Computer and Information Technology, Liaoning Normal University, Dalian, Liaoning 116081, China)

**Abstract:** As a typical ill-posed problem, image super-resolution reconstruction has always been paid attention to. Although many effective super-resolution reconstruction models using convolutional neural networks have been developed in recent years, how to mine the prior information in the image entirely to improve the details in reconstructed image needs to be further studied. In this paper, a single image super-resolution reconstruction model using deep residual networks with non-decimated wavelet transform edge learning NDW-EDRN (Non-Decimated Wavelet Edge learning using Deep Residual Networks) is proposed. On the basis of low-frequency regions with multiple redundant information, smoothness and low gradient values, as well as high-frequency regions with edges and high gradient values that are obtained from an image after a non-decimated wavelet transform, CNN (Convolutional Neural Networks) model with different structures are designed to learn the low and high-frequency subbands separately in the overall network framework: dense skip connection is introduced for integrating learning the mapping between low-frequency subbands; a novel U-net model is designed, which makes the edges that lost in the process of image degradation as the expected output of the network, and block-based skip connection is designed at the same time for making the network learn the lost edges more robust, to obtain the lost edge details during image degradation more sufficiently and comprehensively. A large number of experimental results show that the network can improve the quality of the reconstructed image effectively and especially has a certain advantage in recovering low-resolution images lost, and makes up for the deficiencies of traditional CNN models in image learning detail information to some extent.

**Key words:** convolutional neural networks; residual learning; non-decimated wavelet transform; image super-resolution reconstruction; texture edge information

## 1 引言

卷积神经网络(Convolutional Neural Networks, CNN)<sup>[1,2]</sup>作为一种具有深层结构的前馈神经网络(Feedforward Neural Networks, FNN),其强劲的特征学习能力被计算机视觉领域所关注<sup>[3,4]</sup>,其中CNN所具有的网络权值共享特性以及对平移、缩放的高度不变性被应用于图像的超分辨率重建中,比如最早出现的超分辨率重建CNN模型SRCNN<sup>[5]</sup>及其改进模型FSRCNN(Fast SRCNN, FSRCNN)模型<sup>[6]</sup>,前者采用3层卷积层用以学习LR图像和相应HR图像之间端对端的直接映射;后者则通过将SRCNN预处理中的双三次插值替换为后处理的反卷积.这两种模型为基于CNN图像超分辨率重建研究奠定了基础.

随着研究的不断深入,如何通过加深CNN层数来提高重建图像的质量,以及如何减轻模型的计算负担一直受到重视.对于提高CNN层数的问题,有两个比较经典的模型:VDSR(Very Deep Super-Resolution, VDSR)<sup>[7]</sup>和DRCN(Deeply-Recursive Convolutional Network, DRCN)<sup>[8]</sup>,前者设计了20层的卷积层,后者则通过将同一卷积层进行16次递归提高CNN的层数.进一步为了解决更深层次网络所带来的梯度“爆炸”和参数训练增量问题,VDSR采用了残差学习与增大学习率的方法,DRCN模型则采用了递归监督和跳跃连接的方案;对于模型的计算负担问题,文献[9]提出一种子像素CNN模型ESPCN(Efficient Sub-Pixel Convolutional Neural network, ESPCN),该模型通过把各低分辨率通道的同一位置像素视为高分辨率通道的对应子像素,将低分辨率图像特征按照对应位置周期性的插入到高分辨率图像中,实现图像分辨率的提升;文献[10]提出一种基于拉普拉斯金字塔超分辨率模型LapSRN(Laplacian pyramid Super-Resolution Network, LapSRN),其通过拉普拉斯金字塔框架渐进重建多等级高分辨率残差,并采用Charbonnier损失函数与跳跃连接实现了模型的有效训练.此外,近年来生成对抗网络模型(Generative Adversarial Network, GAN)也为超分问题的求解提供了新思路.文献[11]提出了SRGAN网络模型用于图像的超分重建,取得了更加逼真的重建图像;文献[12]提出了一种感知生成对抗网络(Super Resolution Perceptual Generative Adversarial Network, SRPGAN),用实例标准化取代原始SRGAN模型中的批量标准化,进一步提升了重建图像的质量;文献[13]针对SRGAN模型重建图像的伪影问题,提出了一种增强SRGAN模型(Enhanced SRGAN, ESRGAN),利用稠密残差块

(Residual-in-Residual Dense Block, RRDB)作为基本单元,借用相对GAN网络<sup>[14]</sup>思想,通过预测两幅图像的相对真实性对感知损失做出改进,在有效去除伪影的同时进一步提升了图像的超分重建质量.

尽管目前基于深度神经网络的图像超分辨率重建得到了较好的发展,但仍有许多工作尚需研究和完善,比如网络的结构有待创新、效率有待提高,对有效特征的挖掘还有待加强,重建超分辨率图像的质量还被期望进一步提升等等.近年来基于多尺度几何分析(Multiscale Geometrical Analysis, MGA)的图像稀疏表示取得了很好的发展<sup>[15,16]</sup>,文献[17]构建了一种Wavelet域深度残差学习算法用于图像去噪和超分辨率重建;文献[18]构建了一个基于图像Wavelet变换子带的深度学习网络,将输入的低分辨率图像作为图像Wavelet变换的低频子带,通过预测图像Wavelet变换的高频子带系数作为构建高分辨率子带“丢失”的细节来实现超分辨率图像的重建.

上述两种方案对图像Wavelet变换后的各子带采用了相同的学习策略.事实上低频与高频子带表现出不同的统计特性,低频子带通常会显著影响着两幅图像的主体内容相似程度;而高频子带则在图像边缘等梯度值较大的区域发挥着重要作用.受此启发,本文首先对LR图像与HR图像非抽取Wavelet(Non-Decimated Wavelet Transform, NDWT)的低频和高频子带的统计特性进行分析,进而提出一种基于NDWT边缘学习的深度残差网络模型NDW-EDRN用于单幅图像超分辨率重建,对输入图像NDWT低频和高频子带采取不同的学习策略来学习LR子带与HR子带间的映射关系,对低频子带采用基于稠密跳跃连接的“整体式”学习方式,而对高频子带则采用了一种基于块U-net级联的“缺失式”学习方式,即放弃原始将低、高分辨率图像的高频子带直接进行学习,利用网络模型学习它们间更为稀疏的差值,使网络学习更具针对性.

## 2 单幅图像超分重建及图像NDWT子带分析

### 2.1 单幅图像超分辨率重建

单幅图像超分辨率重建是指从观察到的包含模糊、下采样以及噪声等退化情况的单幅低分辨率图像中估计出高分辨率图像的过程<sup>[19]</sup>,其成像模型的一般过程可描述为

$$\mathbf{I}^{\text{LR}} = (\mathbf{B} * \mathbf{I}^{\text{HR}}) \downarrow_{(m,n)} + \mathbf{N}_{\sigma} \quad (1)$$

其中, $\mathbf{I}^{\text{LR}}$ 是观测到的低分辨率图像, $\mathbf{B}$ 是模糊矩阵, $\mathbf{I}^{\text{HR}}$ 是

原始高分辨率图像,  $(m, n)$  是水平与垂直方向的降采样因子, 实际应用中一般假设二者相同,  $N_\sigma$  是方差为  $\sigma$  的噪声.

单幅图像超分辨率重建一般可以看成是图像经由  $I^{HR}$  退化为  $I^{LR}$  的反问题, 参见图 1. 由于图像  $I^{HR}$  在退化过程中损失了不可逆的高频细节信息, 因此该反问题的解并不唯一, 从而导致图像超分辨率重建问题具有高度的不稳定性.

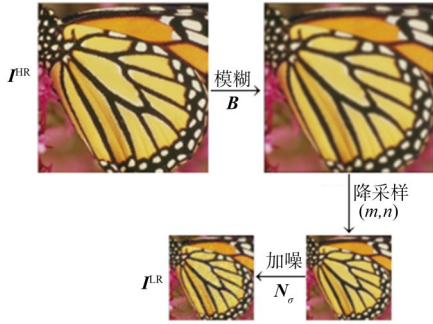


图 1 图像退化过程示意图

### 2.2 非抽取 Wavelet 变换

作为标准离散 Wavelet 变换的扩展, 文献 [20~22] 等基于不同的应用领域提出了非抽取 Wavelet 变换 NDWT. 假设低通滤波器  $\mathcal{H}$ 、高通滤波器  $\mathcal{G}$  为正交镜滤波器 (Quadrature Mirror Filters, QMF) [23], 即对于由有限个非零值序列  $\{h_n\}$  定义的  $\mathcal{H}$ , 满足如下正交性:

$$\sum_n h_n h_{n+2j} = 0 \quad (2)$$

其中  $j \neq 0, \sum_n h_n = 1$ . 同时定义  $\mathcal{G}$  为  $\{g_n | g_n = (-1)^n h_{1-n}\}$ , 容易看出,  $\mathcal{G}$  也满足正交性, 并与  $\mathcal{H}$  之间具有“互相交”关系, 即有

$$\sum_n h_n g_{n+2j} = 0 \quad (3)$$

对于信号  $\{\dots, x_{-2}, x_{-1}, x_0, x_1, \dots, x_2, \dots\}$ ,  $\mathcal{H}$  滤波、 $\mathcal{G}$  滤波过程为

$$\begin{cases} (\mathcal{H}x)_n = \sum_k h_{k-n} x_k \\ (\mathcal{G}x)_n = \sum_k g_{k-n} x_k \end{cases} \quad (4)$$

进一步定义  $D_0$  和  $D_1$  分别为基于偶数和奇数的二抽样算子:

$$\begin{cases} (D_0x)_j = x_{2j} \\ (D_1x)_j = x_{2j+1} \end{cases} \quad (5)$$

标准离散 Wavelet 变换是基于  $\mathcal{H}$ 、 $\mathcal{G}$  和  $D_0$  实现的 [21]: 对于原信号  $c: \{c_m | m=0, 1, \dots, n-1\}$ , 假设其具有周期性的边界条件, 定义第  $J$  层的平滑逼近信息  $c^j$  为

$$c_m^j = c_m, m=0, 1, \dots, n-1. \quad (6)$$

对于  $j=J-1, J-2, \dots, 1, 0$ , 递归的定义信号第  $j$  层的

平滑逼近信息  $c^j$  和细节信息  $d^j$ :

$$\begin{cases} c^j = D_0 \mathcal{H} c^{j+1} \\ d^j = D_0 \mathcal{G} c^{j+1} \end{cases} \quad (7)$$

其中  $c^j$  和  $d^j$  的长度分别为  $2^j$ . 可以看出在式 (7) 的递归分解中, 仅有当前层的平滑逼近信号参与下一层的平滑逼近信息与细节信息的分解.

NDWT 对上述标准 Wavelet 过程进行了扩展, 通过  $D_1$  和  $D_0$  算子在各个尺度上提取奇、偶索引的平滑逼近信息和细节信息: 对于初始信号  $c^j$ , 通过  $D_1 \mathcal{G} c^j$  和  $D_0 \mathcal{G} c^j$  获得奇、偶索引的细节信息, 通过  $D_1 \mathcal{H} c^j$  和  $D_0 \mathcal{H} c^j$  获得奇、偶索引的平滑逼近信息, 每一个奇、偶信息的长度分别为  $n/2$ , 这样所获得的细节信息 (即小波系数) 的总长度为  $2 \times n/2 = n$ .

进一步对所获得的奇、偶索引的细节信息分别进行  $D_1 \mathcal{G}$  和  $D_0 \mathcal{G}$  操作, 获得当前层的细节信息; 对所获得的奇、偶索引的平滑逼近信息进行  $D_1 \mathcal{H}$  和  $D_0 \mathcal{H}$  操作, 获得当前层的平滑逼近信息, 这样始终能够保持每层分解所获得的细节信息的总长度为  $n$ . 该过程重复进行, 直到完成整个变换. 具体过程参见图 2, 其中  $d_0$  和  $d_1$  为第一层的细节系数, 长度分别为  $n/2$ ;  $d_{00}, d_{10}, d_{01}$  和  $d_{11}$  为第二层的细节系数, 长度分别为  $n/4$ .

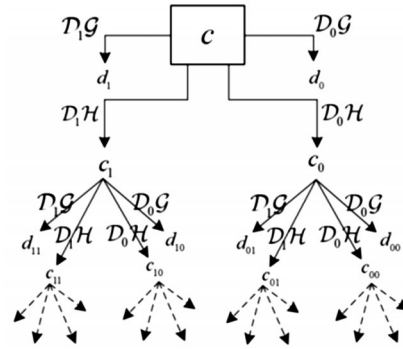


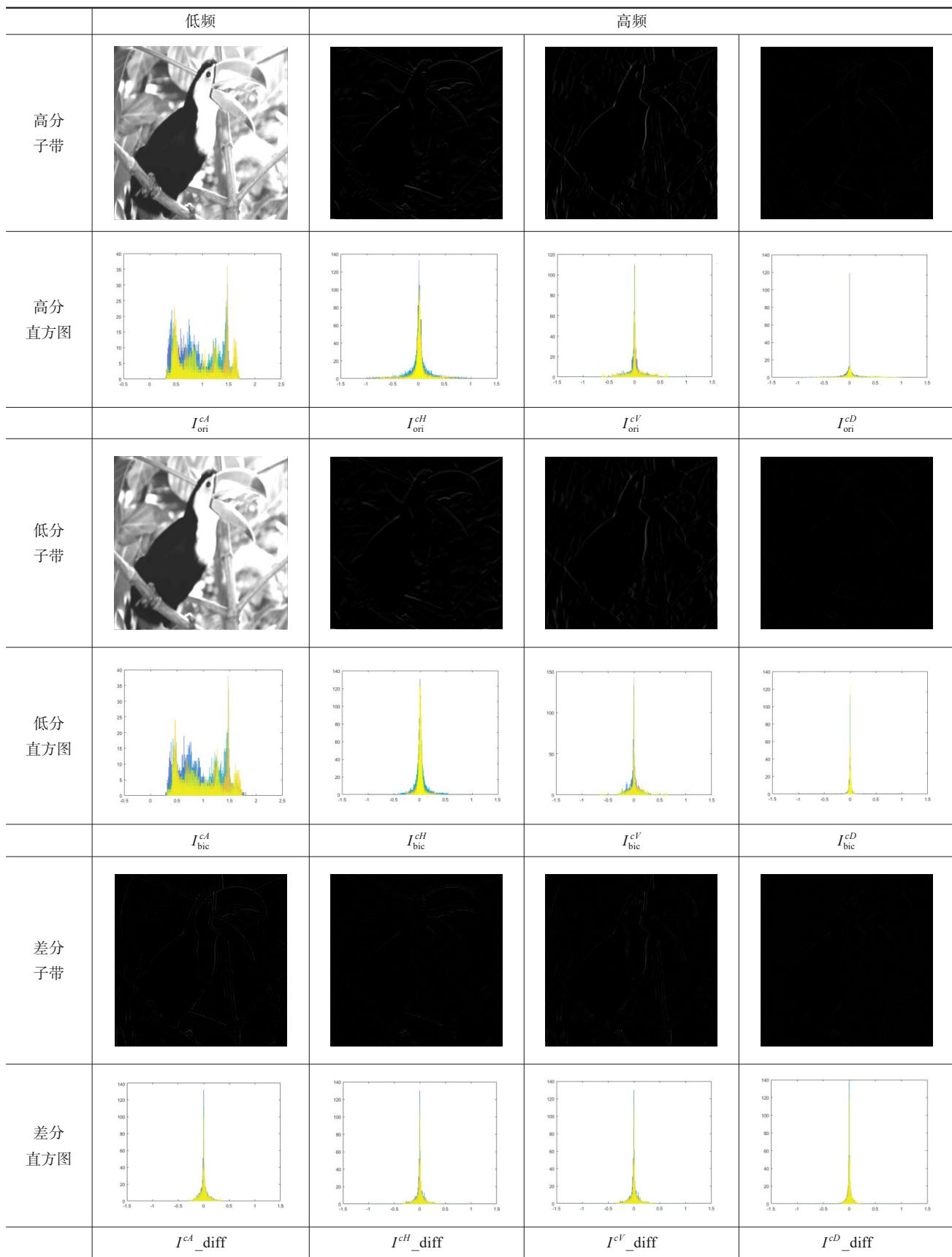
图 2 NDWT 过程流程图 [22]

NDWT 不仅保留了离散 Wavelet 变换的多尺度特性和方向特性, 同时还具有平移不变特性, 更多的获得了信号的细节信息, 从而为图像融合、图像超分辨率重建等应用奠定了基础.

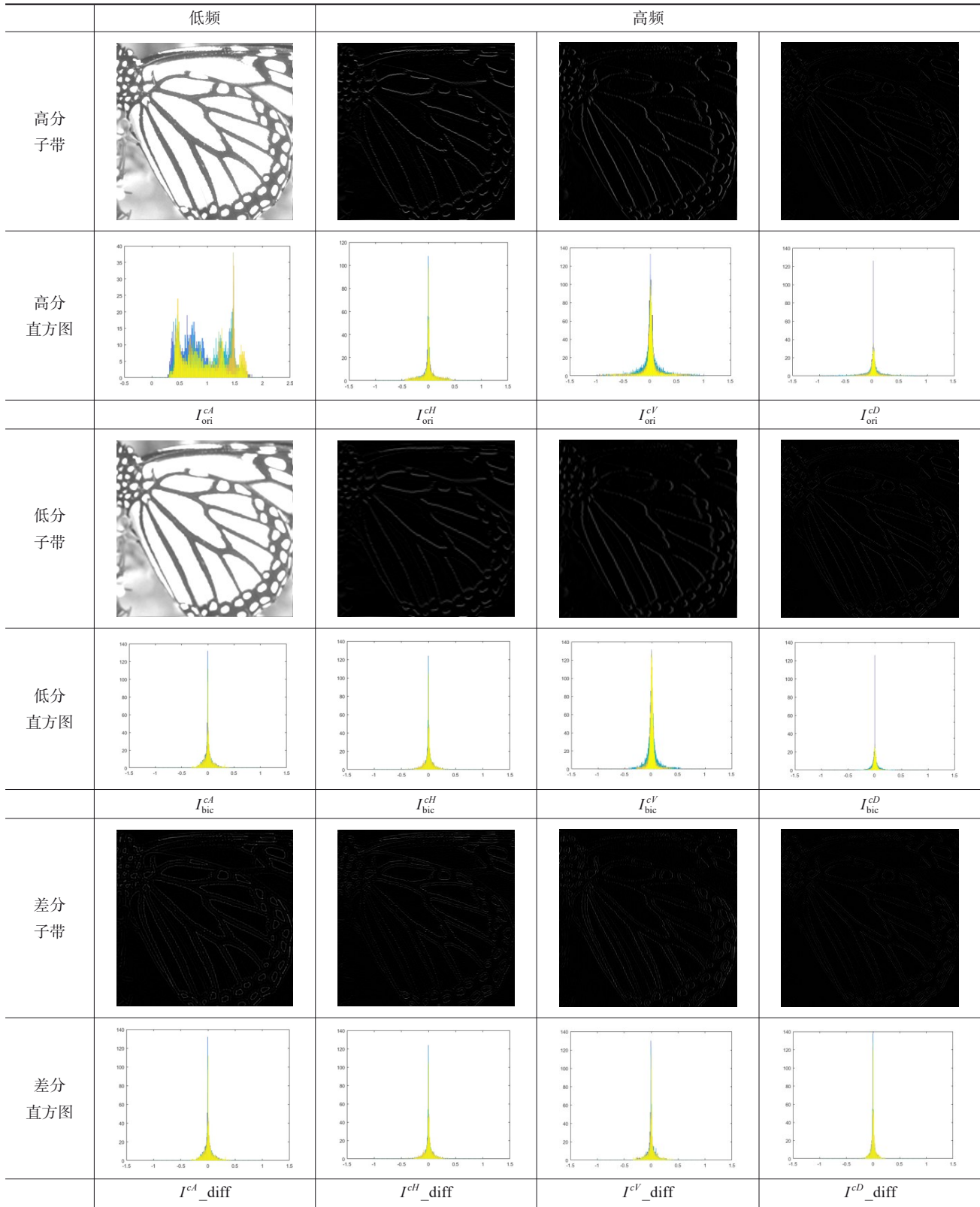
### 2.3 图像 NDWT 子带特性分析

选取文献 [24] 提供的 Set5 中具有丰富细节的 bird 图像和边缘明显的 butterfly 图像作为测试对象.

首先将 RGB 测试图像转换为 YCbCr, 并提取其 Y 通道, 记为  $I_{ori}$ ; 利用双三次插值函数得到初始重建图像, 记为  $I_{bic}$ , 将其作为模拟  $I_{ori}$  的退化图像 (即低分辨率图像); 对  $I_{ori}$  和  $I_{bic}$  分别进行一级 NDWT 分解, 分别获得两个低频子带  $I_{ori}^{cA}$  和  $I_{bic}^{cA}$ , 以及六个高频子带  $\{I_{ori}^{cH}, I_{ori}^{cV}, I_{ori}^{cD}\}$  与  $\{I_{bic}^{cH}, I_{bic}^{cV}, I_{bic}^{cD}\}$ . 进一步, 对各子带的系数直方图以及对应的差分直方图进行统计, 结果参见图 3.



(a) bird 图像



(b) butterfly 图像

图3 图像NDWT子带、差分子子带及其直方图统计

对图3的各子带及其统计结果进行观察与分析后,可分别得到以下结论:

(1)  $I_{\text{ori}}^{cA}$  与  $I_{\text{bic}}^{cA}$  子带在主体内容上极为相似,其差分直方图亦显示两者间差值均集中在小数值范围内,并且绝大多数为0,具有聚集性;

(2) 不论是  $I^{cH}$  还是  $I^{cV}$ ,  $I^{cD}$ , 其统计直方图均显示图像的任一高频子带绝大多数系数值分布在零点附近,只有少量系数值较大,具有“高峰拖尾”的稀疏性;

(3) 从高频差分子带可看出同一方向高频子带差值较大的系数大多位于图像退化过程中所丢失的边缘、拐点等梯度值较大的位置。

基于以上对各子带特性的分析与相关文献的支撑<sup>[25]</sup>可知:低频子带的恢复决定着图像的主体内容质量,它对应着较高的峰值信噪比值;高频子带的恢复则显著影响着图像的视觉感知质量,它呈现出较高的结构相似度.因此,本文设计两种不同的学习策略与网络结构来更有针对性地对图像的低、高频子带进行区分学习.对于低频子带,采用稠密跳跃连接的“整体式”学

习方式,将低分辨率图像的低频子带作为网络的输入,将高分辨率图像的低频子带作为网络的输出,网络采用稠密跳跃的连接方式,使深层次的网络学习更加充分,避免梯度弥散现象;对于高频子带,则采用块U-net级联的“缺失式”学习方式,将低分辨率的高频子带作为网络的输入,将高、低分辨率图像的高频子带的差值作为网络的输出,网络采用编解码与块级级联的方式,增强稀疏性的同时,防止梯度弥散,使低分辨率图像所丢失的边缘等高频信息得到有效的补充。

### 3 NDWT 边缘学习的深度残差网络

由前面的分析可以看出,图像经双三次插值后的LR图像与原始HR图像之间相比,主要在于图像退化过程中高频信息的丢失,而低频信息较为相近.为此本文构建如图4所示的深度残差网络模型NDW-EDRN (Non-Decimated Wavelet Edge learning using Deep Residual Networks, NDW-EDRN)用于单幅图像超分辨率重建。

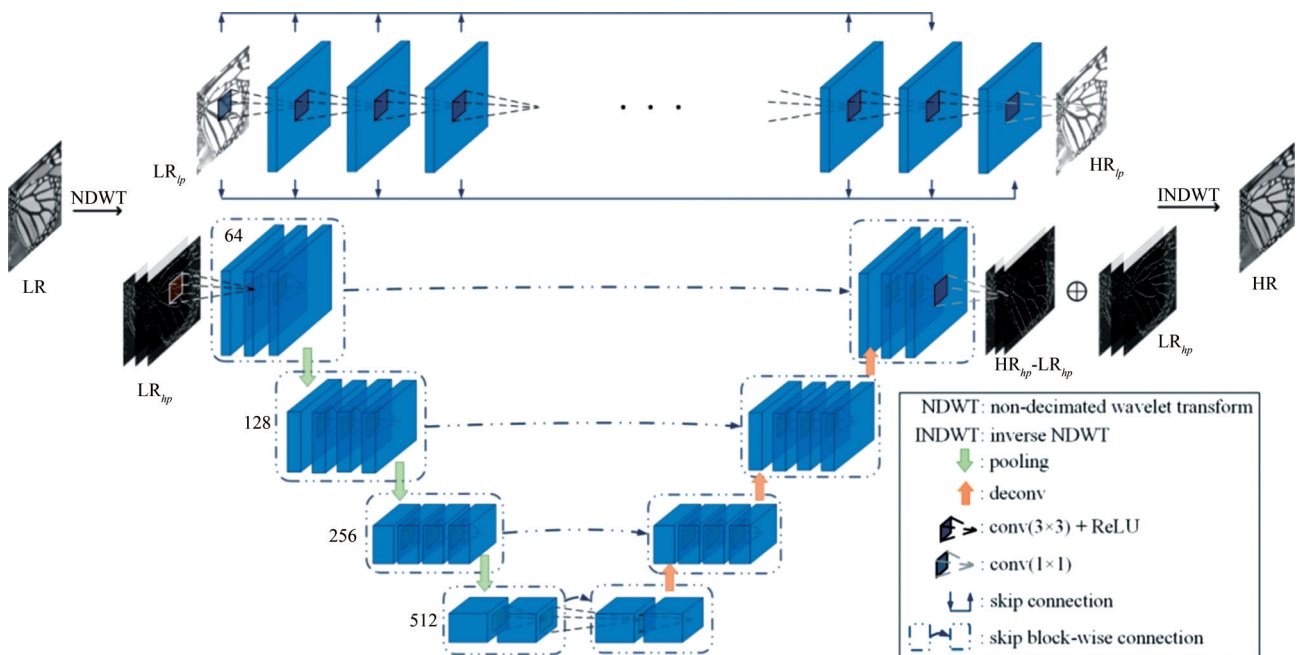


图4 深度残差网络模型NDW-EDRN结构图

模型NDW-EDRN对图像的NDWT子带采用了不同的学习方式,其中对低频子带间的映射关系学习采用了“整体式”的网络结构<sup>[17,18]</sup>,网络采用了20层的卷积操作,并通过稠密<sup>[26,27]</sup>的跳跃链接结构来充分学习低频子带内所包含的信息;而对高频子带间所丢失的边缘信息则采用“缺失式”的网络结构进行学习,表现为利用文献<sup>[28]</sup>所提出的U-net架构与文献<sup>[29,30]</sup>所提出的深度残差学习策略,构建一个学习丢失边缘的端到端网络,

提取出图像中更加深层且复杂的特征,弥补在图像退化过程中所丢失的重要信息.主要包含以下四种操作:

- (1) Conv(3×3)+ReLU: 每阶段获取特征图的主要方式;
- (2) Conv(1×1): 位于网络的最后一层,用于映射期望;
- (3) Pooling与Deconv: 扩大感受野并还原;
- (4) Skip block-wise connection: 本文在传统U-net

通过级联对应阶段子特征图增强网络强度的基础上,把单子特征图拓展为块子特征图,以像素级累加的方式将网络的编码、解码两端更紧密地联系在一起,进一步提升网络的有效性.

对于模型上线路的稠密链接部分旨在让当前第  $l$  层的输入接收来自前  $[l-1, l-2, \dots, 0]$  各层的输出. 若定义第  $l$  层的输入为  $I_l$ , 输出为  $O_l$ , 那么

$$\begin{cases} I_l = O_{l-1} + O_{l-2} + \dots + O_0 \\ O_l = f_l(I_l) \end{cases} \quad (8)$$

此处的  $f_l$  是对应当前  $l$  层的运算处理.

对于模型下线路的改进 U-net 部分旨在采用如下图 5 所示的以块为单位的跳跃级联, 达到信息紧密传递的目的.

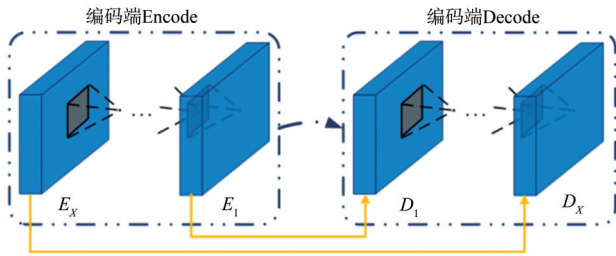


图5 块级级联示意图

若当前处于网络的第  $r$  级, 编解码端各有  $x$  个单子特征图, 则编码端的各单子特征图可用  $E_x^r, E_{x-1}^r, \dots, E_1^r$  表示, 定义其第  $t$  个的输入为  $IE_t^r$ , 输出为  $OE_t^r$ ; 解码端的各单子特征图可用  $D_1^r, D_2^r, \dots, D_x^r$  表示, 定义其第  $t$  个输入为  $ID_t^r$ , 输出为  $OD_t^r$ . 那么在解码端的各单子特征图的输入与输出可分别计算为

$$\begin{cases} \text{当 } t=1 \text{ 时, } ID_t^r = OE_t^r + \text{deconv}(OD_{t+1}^r) \\ \quad OD_t^r = f_2(ID_t^r) \\ \text{当 } t \neq 1 \text{ 时, } ID_t^r = OE_t^r + OD_{t-1}^r \\ \quad OD_t^r = f_2(ID_t^r) \end{cases} \quad (9)$$

NDW-EDRN 的总体损失函数如下所示:

$$L_k(\Theta_k) = \begin{cases} \frac{1}{2N} \sum_{i=1}^N \|F_k(y_i^k; \Theta_k) - x_i^k\|_F^2, & k \in \{cA\} \\ \frac{1}{2N} \sum_{i=1}^N \|F_k(y_i^k; \Theta_k) - (y_i^k - x_i^k)\|_F^2, & k \in \{cH, cV, cD\} \end{cases}$$

其中  $k$  表示子带类型,  $F_k(\cdot, \cdot)$  是当子带类型为  $k$  时所需训练的网络模型.  $x_i^k, y_i^k$  分别为第  $i$  个高分辨率与低分辨率的  $k$  类型子带块. 当网络走“整体式”学习策略时,  $k \in \{cA\}$ ,  $\Theta_k$  表示低频网络的参数集,  $\{(y_i^k, x_i^k)\}_{i=1}^N$  表示第  $i$  个 LR 低频-HR 低频的对应子块; 当网络走“缺失式”学习策略时,  $k \in \{cH, cV, cD\}$ ,  $\Theta_k$  表示高频网络的参数集,  $\{(y_i^k, y_i^k - x_i^k)\}_{i=1}^N$  表示第  $i$  个 LR 高频-差分高频的对应子

块,  $(y_i^k - x_i^k)$  表示 LR 缺少的高频信息.

模型的物理含义是对低频分量而言, 衡量 LR 与 HR 间对应子块的相似程度; 对高频分量而言, 训练出一组能够有效描述 LR 对比 HR 所缺失的边缘算子, 来有效补充高频信息.

## 4 实验与分析

### 4.1 实验设置

训练集: BSD500<sup>[31]</sup> (去除与测试集重复的部分)、General00<sup>[6]</sup> 和 91images<sup>[32]</sup>; 测试集: Set5<sup>[24]</sup>、Set14<sup>[33]</sup> 和 Urban100<sup>[34]</sup>; 采用双三次函数模拟下采样因子为 2 和 4 的 Y 通道退化; 设置“整体式”网络中子带块大小为  $64 \times 64$ , “缺失式”网络中子带块大小为  $192 \times 192$ ; 网络模型采用 Adam 优化算法求解<sup>[35]</sup>, 设置初始学习率为 0.001, 在第 40 个周期后衰减为 0.0001. 实验环境为 Intel(R) Xeon(R) CPU E5-2630 v4 @ 2.20 GHz 在 Ubuntu 14.04, Caffe7.5 平台上进行, NVIDIA Titan X GPU 用于网络训练.

### 4.2 定量与定性分析

表 1 给出与 SRCNN<sup>[5]</sup>、VDSR<sup>[7]</sup>、SRMD<sup>[36]</sup> 和 DSRN<sup>[37]</sup> 非多尺度且网络层数小于等于 20 的深度学习模型, 以及与 DWSR<sup>[18]</sup> 和 WaveResNet<sup>[17]</sup> 多尺度深度学习类模型分别放大 2 倍与放大 4 倍的 PSNR 与 SSIM 对比实验结果; 表 2 给出与 SRGAN<sup>[11]</sup>、SRPGAN<sup>[12]</sup> 生成对抗网络模型, 以及与 DRRN<sup>[38]</sup>、LapSRN<sup>[10]</sup> 和 MemNet<sup>[39]</sup> 非多尺度且网络层数大于 50 的深度学习模型别放大 2 倍与放大 4 倍的 PSNR 与 SSIM 对比实验结果. 两个表中分别对最优结果做出“加粗”处理, 对次优结果做出“下划线”处理. 同时图 6, 7, 8, 9 给出了 Set5、Set14 与 Urban100 数据集的部分重建彩色结果图及其细节处的放大对比图, 对各指标的最优及次优结果做出与表 1、2 中同样的处理.

表 1 的实验结果显示, 所提出的 NDW-EDRN 模型较近几年层数小于等于 20 的主流非多尺度深度学习模型和多尺度深度学习模型在客观评价指标 PSNR 与 SSIM 表现上均呈现着一定的优越性, 除了对 Set14 放大 2 倍的实验结果取得了次优结果, 其余均为最优结果. 对表 1 的统计结果进行分析后, 可以看出: PSNR 的次优结果大多分布于非多尺度深度学习模型实验结果中, 而 SSIM 的次优结果则绝大多数分布于多尺度深度学习模型实验结果中. 这说明, 非多尺度深度学习模型在恢复图像主体内容, 即与原始图像的像素级相似度呈现上表现良好; 而基于多尺度的深度学习模型则在恢复图像边缘结构时发挥着重要作用, 能获得比非多尺度深度学习模型更为优异的纹理结构级相似度.

表 2 展现了本文模型与近几年主流的生成对抗网络模型及层数大于 50 的主流非多尺度深度学习模型间的实验数据对比, 亦可看出, 本文模型相较于 SRGAN

表1 重建图像的定量比较 1

数据集	倍数	非多尺度深度学习模型(层数≤20)								多尺度深度学习模型				本文模型	
		SRCNN		VDSR		SRMD		DSRN		DWSR		WaveResNet		NDW-EDRN	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Set5	×2	36.66	0.954	37.53	<u>0.959</u>	37.53	<u>0.959</u>	<u>37.66</u>	<u>0.959</u>	37.43	0.957	37.57	<u>0.959</u>	<b>37.77</b>	<b>0.960</b>
	×4	30.04	0.861	31.35	0.884	31.59	<u>0.887</u>	31.40	0.883	31.39	0.883	31.52	0.886	<b>31.72</b>	<b>0.889</b>
Set14	×2	32.42	0.906	33.03	0.912	33.12	0.914	<u>33.15</u>	0.913	33.07	<b>0.917</b>	33.09	0.913	<b>33.26</b>	<u>0.914</u>
	×4	27.49	0.750	28.01	0.767	<u>28.15</u>	<b>0.772</b>	28.07	<u>0.770</u>	28.04	0.767	28.11	<u>0.770</u>	<b>28.27</b>	<b>0.772</b>
Urban100	×2	29.50	0.895	30.76	0.914	30.89	0.916	30.97	0.916	30.46	0.916	30.96	0.917	<b>31.30</b>	<b>0.926</b>
	×4	24.52	0.722	25.18	0.752	25.34	0.761	25.08	0.747	25.26	0.755	<u>25.36</u>	0.761	<b>25.51</b>	<b>0.763</b>
Average	×2	32.86	0.918	33.77	0.928	33.85	<u>0.930</u>	<u>33.93</u>	0.929	33.65	0.928	33.87	<u>0.930</u>	<b>34.11</b>	<b>0.933</b>
	×4	27.50	0.778	28.18	0.801	<u>28.36</u>	<u>0.807</u>	28.18	0.800	28.23	0.802	28.33	0.806	<b>28.50</b>	<b>0.808</b>

表2 重建图像的定量比较 2

数据集	倍数	生成对抗网络模型				非多尺度深度学习模型(层数>50)						本文模型	
		SRGAN		SRPGAN		DRRN_B1U25		LapSRNss-D5R8		MemNet		NDW-EDRN	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Set5	×2	—	—	29.67	0.950	37.74	<u>0.959</u>	37.72	<b>0.960</b>	<b>37.78</b>	<b>0.960</b>	<u>37.77</u>	<b>0.960</b>
	×4	29.40	0.847	22.68	0.880	31.68	<b>0.889</b>	<b>31.74</b>	<u>0.888</u>	<b>31.74</b>	<b>0.889</b>	<u>31.72</u>	<b>0.889</b>
Set14	×2	—	—	27.66	<u>0.911</u>	33.23	<b>0.914</b>	33.24	<b>0.914</b>	<b>33.28</b>	<b>0.914</b>	<u>33.26</u>	<b>0.914</b>
	×4	26.02	0.740	22.50	<b>0.786</b>	28.21	0.772	28.25	<u>0.773</u>	<u>28.26</u>	0.772	28.27	0.772
Urban100	×2	—	—	30.41	0.892	31.23	0.919	31.01	0.917	<b>31.31</b>	0.920	<u>31.30</u>	<b>0.926</b>
	×4	—	—	20.00	0.763	25.44	<u>0.764</u>	25.45	<b>0.765</b>	<u>25.50</u>	0.763	25.51	0.763
Average	×2	—	—	20.91	0.918	34.07	<u>0.931</u>	33.67	0.927	<b>34.12</b>	<u>0.931</u>	<u>34.11</u>	<b>0.933</b>
	×4	—	—	21.73	<b>0.810</b>	28.44	<u>0.808</u>	28.31	0.804	<b>28.50</b>	<u>0.808</u>	<b>28.50</b>	<u>0.808</u>

与 SRPGAN 在两个评价指标上都表现得更为稳健,更加具备实用性与稳定性;与层数大于 50 的主流非多尺度深度学习模型相比,本文模型在部分实验结果中可

与之持平甚至超过。

图 6、7 的生成结果同时在主观上也能够看出本文模型在图像边缘细节的重建方面,尤其重建倍数较大时,较

测试图像	GT	Bicubic	SRCNN	LapSRN	VDSR	NDW-EDRN
Set5 中 baby 图像						
	PSNR:	35.65	37.06	37.23	<u>37.25</u>	37.38
	SSIM:	0.979 8	0.985 1	<u>0.985 6</u>	<u>0.985 6</u>	<b>0.986 0</b>
Set14 中 baboon 图像						
	PSNR:	22.55	22.79	23.32	<u>23.42</u>	23.47
	SSIM:	0.794 2	0.827 9	0.839 1	<u>0.840 7</u>	<b>0.842 3</b>
Urban100 中 img_063 图像						
	PSNR:	21.14	24.19	23.80	<u>24.50</u>	24.61
	SSIM:	0.874 0	0.913 4	0.913 0	<u>0.916 8</u>	<b>0.917 8</b>

图6 对比其他非多尺度网络模型放大 2 倍的细节图

其他模型有着明显的优势,比如无论是butterfly图像中蝴蝶翅膀的纹路、lenna图像中的帽檐,亦或是img\_087图像

中的楼房边缘等都较其他模型重建后的图像边缘更清晰、稳定和有效,在重建质量上表现出一定的优势。

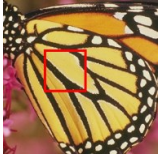
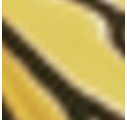
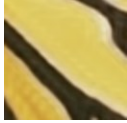
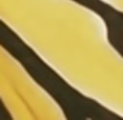
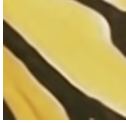
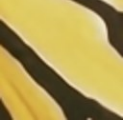
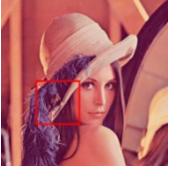
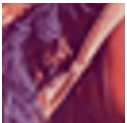
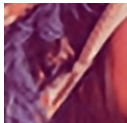
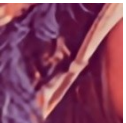
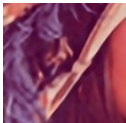
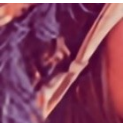






测试图像	GT	Bicubic	SRCNN	LapSRN	VDSR	NDW-EDRN
Set5中 butterfly 图像						
	PSNR:	20.90	23.95	25.52	<u>25.56</u>	25.89
	SSIM:	0.861 4	0.925 1	<u>0.947 0</u>	0.946 6	<b>0.950 0</b>
Set14中 lenna 图像						
	PSNR:	28.36	29.50	<u>29.88</u>	29.80	29.97
	SSIM:	0.963 2	0.970 0	<u>0.972 2</u>	0.971 8	<b>0.972 6</b>
Urban100中 img_087 图像						
	PSNR:	21.64	23.57	<u>24.64</u>	24.57	25.22
	SSIM:	0.838 0	0.884 5	<u>0.907 2</u>	0.905 7	<b>0.916 1</b>

图7 对比非多尺度其他网络模型放大4倍的细节

图8、9所示的多尺度网络模型的重建图像细节对比图显示了在恢复图像纹理、边缘时,这类的网络模型具备一定的优越性与普适性.说明在构建超分辨率重建网络时考虑多尺度变换,利用其提取图像结构信息的这一优势,使网络学习到更加具有针对性的边缘的

这一想法是成立且有效的.而本文在此基础上进一步归纳总结非抽取小波变换的低、高频两种子带系数分别呈现的规律后,提出了相应的网络学习模式,使低频子带中所蕴涵的信息被学习得更加完备,使高频子带中所缺失的信息得到指向性补充.


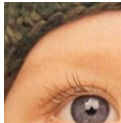

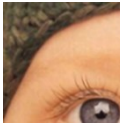








测试图像	GT	DWSR	WaveResNet	NDW-EDRN
Set5中 baby 图像				
	PSNR:	37.00	37.35	37.38
	SSIM:	0.984 0	<b>0.986 1</b>	0.986 0
Set14中 baboon 图像				
	PSNR:	23.38	23.44	23.47
	SSIM:	0.841 1	<b>0.842 3</b>	<b>0.842 3</b>
Urban100中 img_063 图像				
	PSNR:	24.57	24.60	24.61
	SSIM:	0.917 2	0.917 5	<b>0.917 8</b>

图8 对比其他多尺度网络模型放大2倍的细节图

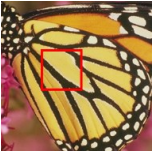
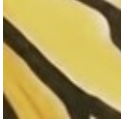
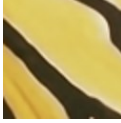
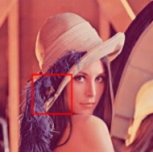
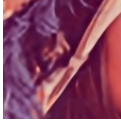
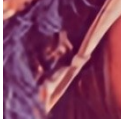



测试图像	GT	DWSR	WaveResNet	NDW-EDRN
Set5 中 butterfly 图像				
	PSNR:	25.36	—	25.89
	SSIM:	0.944 4	—	0.950 0
Set14 中 lenna 图像				
	PSNR:	29.89	—	29.97
	SSIM:	0.972 0	—	0.972 6
Urban100 中 img_087 图像				
	PSNR:	24.85	—	25.22
	SSIM:	0.907 6	—	0.916 1

图9 对比其他多尺度网络模型放大4倍的细节(注: WaveResNet 网络未公开四倍的预训练模型)

表3显示了本文模型与其他流行超分辨率重建模型的参数与计算量对比. 由于所提模型采用了“整体式”与“缺失式”两种联合学习模式来分别处理低频子带与高频子带, 这无疑增加了一定的参数量, 这也是需要在未来进一步优化地方.

表3 参数与计算量比较

	SR_	VD_	SR_	DR_	Lap_	Mem_	NSW
	CNN	SR	MD	RN	SRN	Net	-EDRN
Param (k)	57	665	1 483	297	827	2 910	5 736
Flops (G)	0.475	2.73	6.08	1.22	19.13	11.96	3.79

### 4.3 消融实验

为更全面地评估本模型, 表4给出了在非抽取小波变换域下, 各子带只采用同一种学习模式网络的消融实验重建结果的定量比较, 所有的对比使用与上文相同的超参数设置.

表4 消融实验重建数据集的定量比较

数据集	倍数	学习模式					
		整体式		缺失式		NDW-EDRN	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Set5	×2	37.69	0.952	37.74	0.957	<b>37.77</b>	<b>0.960</b>
Set14		33.14	0.909	33.23	0.912	<b>33.26</b>	<b>0.914</b>
Urban 100		31.22	0.921	31.26	0.925	<b>31.30</b>	<b>0.926</b>
Aver age	×2	34.02	0.927	34.08	0.931	<b>34.11</b>	<b>0.933</b>

与此同时, 由于缺失式网络层数大于整体式网络层数, 为避免层数这一嫌疑, 将缺失式网络设置为与整体式网络相同的层数. 即去除了解码端最后一子块

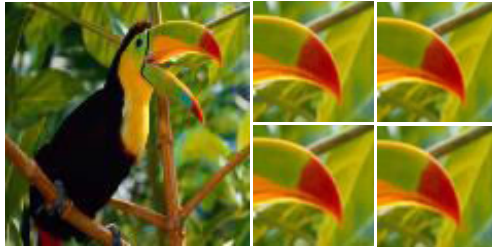
的最后一层以及该级的块级级联操作, 只保留对输入的残差设计, 将该网络记为“缺失式\_s”网络, 那么, 采用该“缺失式\_s”网络的 NDW-EDRN, 则记为“NDW-EDRN\_s”. 图10显示了2倍消融实验结果图, 表5给出相应的定量评价指标数值, 同样采用PSNR与SSIM来评估重建图像的质量.

由表4、5与图10可以看到: 双学习模式的网络比任意单一学习模式的网络在提升图像质量与图像结构相似度方面均有些许提升, 说明图像中低频信息在得到保证的同时, 还被有效地补充进了高频信息. 此外还注意到: (1) 相较于整体式学习, 缺失式学习在网络中更据主导地位, 说明重建边缘纹理信息是超分辨率任务中的重点与难点任务; (2) baby 图像的整体式比缺失式网络重建结果好, 这可能是因为 baby 图像中存在大量平滑信息, 整体式网络更完全地保留了原始低分辨率图像中的有效信息, 而高频信息较少, 导致缺失式网络所需补充的信息少. 因此, 双学习模式的网络可以融合上述两种网络的优点, 既可以对平滑图像学习全面的低频信息, 又可以

表5 消融实验重建图像的定量比较

测试图像	学习模式					
	整体式		缺失式_s		NDW-EDRN_s	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
bird	38.39	0.994 7	<u>38.42</u>	<u>0.995 1</u>	<b>38.46</b>	<b>0.995 3</b>
butter fly	31.95	0.987 4	<u>32.00</u>	<b>0.987 6</b>	<b>32.02</b>	<u>0.987 5</u>
img 068	35.79	0.974 1	<u>35.82</u>	<u>0.974 5</u>	<b>35.88</b>	<b>0.974 8</b>
baby	<u>36.61</u>	<u>0.984 5</u>	36.58	0.984 4	<b>36.64</b>	<b>0.984 6</b>

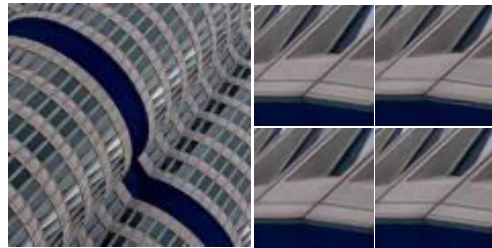
对复杂图像有效地补充高频信息,使网络的学习更健壮、鲁棒.



(a) bird



(b) butterfly



(c) img\_068



(d) baby

图10 消融实验放大2倍的细节(左上角:GT,右上角:“整体式”学习,左下角:“缺失式\_s”学习,右下角:NDW-EDRN\_s)

## 5 结论

本文从弥补低分辨率图像高频信息的角度为着眼点,针对传统深度卷积神经网络学习图像细节区域的不足,通过对NDWT子带间的对比与分析,获得高频子带内系数存在稀疏性、退化图像的低频子带与原始图像的低频子带间存在相似性、图像退化过程中大多丢失高频信息的结论,在此基础上提出一种对低、高频子带采用不同学习策略的深度残差网络模型—NDW-

EDRN,以解决超分辨率重建过程中难以有效补充边缘细节这一难点问题.该网络可主要分为两部分:一是采用稠密跳跃连接的方式整体性学习低频子带间的映射关系;二是对于高频子带采用一种新型的U-net模型,主要表现为将图像退化过程中所丢失的边缘作为网络的期望输出,并改进传统U-net的级联结构,通过采用基于块的跳跃连接来使网络更加有效地学习缺失性边缘.通过大量的实验可以看出,本文提出的NDW-EDRN模型与当前一些经典的单幅图像超分辨率重建网络模型相比,不论主观还是客观都更加有效地提高了重建图像的质量,特别在恢复低分辨率图像所缺失的边缘信息上得到良好效果.

## 参考文献

- [1] LECUN Y, BOSER B E, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. *Neural Computation*, 1989, 1(4): 541-551.
- [2] LECUN Y, BOTTOU L. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [3] 周飞燕,金林鹏,董军.卷积神经网络研究综述[J].*计算机学报*, 2017, 40(6): 1229-1251.  
Zhou Fei-yan, Jin Lin-peng, Dong Jun. Review of convolutional neural network[J]. *Chinese Journal of Computers*, 2017, 40(6): 1229-1251. (in Chinese)
- [4] KHAN S, RAHMANI H, SHAH S A A, et al. A guide to convolutional neural networks for computer vision[J]. *Synthesis Lectures on Computer Vision*, 2018, 8(1): 1-207.
- [5] DONG C, LOY C C, HE K, TANG X. Learning a deep convolutional network for image super-resolution[C]//*Proceedings of the 13th European Conference on Computer Vision*. Berlin: Springer, 2014: 184-199.
- [6] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[C]//*Proceedings of the European Conference on Computer Vision*. Berlin: Springer, 2016: 391-407.
- [7] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//*Proceedings of the Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 1646-1654.
- [8] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution[C]//*Proceedings of the Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2016: 1637-1645.
- [9] SHI W Z, CABALLERO J, HUSZAR F, et al. Real-time single image and video super-resolution using an efficient

- sub-pixel convolutional neural network[C]//Proceedings of the Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 1874-1883.
- [10] 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*, 2017, 41(11): 2599-2613.
- [11] CHRISTIAN L, LUCAS T, FERENC H, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the IEEE conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 105-114.
- [12] WU B, DUAN H, LIU Z, et al. SRPGAN: Perceptual generative adversarial network for single image super resolution[J/OL]. (2017-12-20) [2021-06-20]. <https://arxiv.org/abs/1712.05927>.
- [13] WANG X, YU K, WU S, et al. Esrgan: Enhanced super-resolution generative adversarial networks[C]//Proceedings of the European Conference on Computer Vision. Munich: Springer, 2018: 63-79.
- [14] JOLICOEUR-MARTINEAU A. The relativistic discriminator: A key element missing from standard GAN[J/OL]. (2018-09-10) [2021-06-25]. <https://arxiv.org/abs/1807.00734>.
- [15] 宋传鸣, 赵长伟, 刘丹, 王相海. 3D多尺度几何分析研究进展[J]. *软件学报*, 2015, 26(5): 1213-1236.  
Song Chuan-ming, Zhao Chang-wei, Liu Dan, Wang Xiang-hai. Advances in three-dimensional multiscale geometrical analysis[J]. *Journal of Software*, 2015, 26(5): 1213-1236. (in Chinese)
- [16] VYAS A, YU S, PAIK J. Multiscale Transforms with Application to Image Processing: Wavelets and Wavelet transform[M]. Singapore: Springer, 2018.
- [17] BAE W, YOO J, YE J C. Beyond deep residual learning for image restoration: Persistent homology-guided manifold simplification[C]//Proceedings of the Computer Vision and Pattern Recognition Workshops. Honolulu: IEEE, 2017:1141-1149.
- [18] GUO T, MOUSAVI H S, VU T H, et al. Deep wavelet prediction for image super-resolution[C]//Proceedings of the Computer Vision and Pattern Recognition Workshops. Honolulu: IEEE, 2017: 1100-1109.
- [19] YUE L, SHEN H, LI J, et al. Image super-resolution: The techniques, applications, and future[J]. *Signal processing*, 2016, 128(11):389-408.
- [20] SHENSA M J. The discrete wavelet transform: Wedding the atrous and mallat algorithms[J]. *IEEE Transactions on Signal Processing*, 1992, 40(10): 2464-2482.
- [21] NASON G P, SILVERMAN B W. The stationary wavelet transform and some statistical applications[C]//Wavelets and Statistics Lecture Notes in Statistics. Berlin: Springer, 1995: 281-300.
- [22] NASON G P. Wavelet methods in statistics with R[J]. *Journal of the Royal Statistical Society*, 2010, 173(1): 273-273.
- [23] DAUBECHIES I. Orthonormal bases of compactly supported wavelets[J]. *Comms Pure Appl Math*, 1988, 41(7): 909-996.
- [24] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single image super-resolution based on nonnegative neighbor embedding[C]//Proceedings of the British Machine Vision Conference. Guildford, UK: BMVA Press, 2012: 135-144.
- [25] XIN D, REN Y, MAI X, et al. Wavelet domain style transfer for an effective perception-distortion tradeoff in single image super-resolution[C]//Proceedings of the IEEE International Conference on Computer Vision. Seoul: IEEE, 2019: 3076-3085.
- [26] 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. Venice: IEEE, 2017: 4809-4817.
- [27] 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. Piscataway, NJ: IEEE, 2017: 4809-4817.
- [28] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation [C]//Proceedings of the International Conference on Medical Image Computing & Computer-assisted Intervention. Munich: Springer, 2015: 234-241.
- [29] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]//Proceedings of the Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer, 2015: 234-241.
- [30] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [31] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation[J]. *IEEE*

Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.

- [32] YAGN J, WEIGHT J, HUANG T S, MA Y. Image super-resolution via sparse representation[J]. IEEE Transactions Image Processing, 2010, 19(11): 2861-2873.
- [33] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[C]//Proceedings of the 7th International Conference Curves Surfaces. Berlin: Springer,2012: 11-730.
- [34] 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. Boston: IEEE, 2015: 5197-5206.
- [35] KINGMA D, BA J. Adam: A method for stochastic optimization[J/OL]. (2017-01-17) [2021-05-20]. <https://arxiv.org/abs/1412.6980>.
- [36] 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 & Pattern Recognition. Salt Lake City: IEEE, 2018: 3262-3271.
- [37] HAN W, CHANG S, LIU D, et al. Image super-resolution via dual-state recurrent networks[C]//Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition. Salt Lake City: IEEE, 2018: 1654-1663.
- [38] TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network[C]//Proceedings of the IEEE Conference on Computer Vision & Pattern Recognition. Honolulu: IEEE, 2017: 2790-2798.
- [39] TAI Y, YANG J, LIU X, XU C. Memnet: A persistent memory network for image restoration[C]//Proceedings of the IEEE International Conference on Computer Vision. Venice: IEEE, 2017:4549-4557.



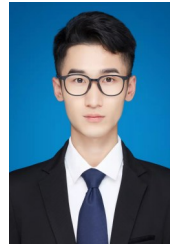
**赵晓阳** 女,1994年4月出生于辽宁省沈阳市.现为辽宁师范大学地理科学学院博士研究生.主要研究方向为深度学习与遥感影像智能信息处理.

E-mail: zxy\_lnnu@163.com



**王鑫莹** 女,1995年10月出生于黑龙江省齐齐哈尔市.现为辽宁师范大学计算机与信息技术学院硕士研究生.主要研究方向为遥感图像处理.

E-mail: 1603331385@qq.com



**赵克云** 男,1997年11月出生于辽宁省沈阳市.现为辽宁师范大学计算机与信息技术学院硕士研究生.主要研究方向为深度学习与遥感图像处理.

E-mail: 1184234775@qq.com



**宋传鸣** 男,1980年10月出生于辽宁省沈阳市.现为辽宁师范大学计算机与信息技术学院教授.主要研究方向为多媒体信息处理和智能图像处理.

E-mail: chmsong@163.com

## 作者简介



**王相海** 男,1965年11月出生于吉林省汪清县.现为辽宁师范大学地理科学学院和计算机与信息技术学院教授、博士生导师.主要研究方向为多媒体信息处理、遥感影像信息处理.

E-mail: xhwang@lnnu.edu.cn