

## 基于混合深度卷积网络的图像超分辨率重建

胡雪影<sup>1,2</sup>, 郭海儒<sup>1</sup>, 朱 蓉<sup>2\*</sup>

(1. 河南理工大学 计算机科学与技术学院, 河南 焦作 454000; 2. 嘉兴学院 数理与信息工程学院, 浙江 嘉兴 314000)

(\* 通信作者电子邮箱 zr@mail.zjxu.edu.cn)

**摘 要:**针对传统图像超分辨率重建方法存在的重建图像模糊、噪声量大、视觉感差等问题,提出了一种基于混合深度卷积网络的图像超分辨率重建方法。首先,在上采样阶段将低分辨率图像放缩至指定大小;然后,在特征提取阶段提取低分辨率图像的初始特征;接着,将提取到的初始特征送入卷积编解码结构进行图像特征去噪;最后,在重建层用空洞卷积进行高维特征提取与运算,重建出高分辨率图像,并且使用残差学习快速优化网络,在降低噪声的同时,使重建图像的清晰度及视觉效果更优。在Set14数据集放大尺度 $\times 4$ 的基准下,将所提方法与双三次插值(Bicubic)、锚定邻域回归(A+)、超分辨卷积神经网络(SRCNN)、极深度超分辨网络(VDSR)、编解码网络(REDNet)等超分辨率重建方法进行对比。在超分辨实验中,所提方法与对比方法比较,峰值信噪比(PSNR)分别提升了2.73 dB、1.41 dB、1.24 dB、0.72 dB和1.15 dB,结构相似性(SSIM)分别提高了0.067 3、0.020 9、0.019 7、0.002 6和0.004 6。实验结果表明,混合深度卷积网络能够有效地对图像进行超分辨率重建。

**关键词:**图像超分辨率重建;图像特征去噪;混合深度卷积网络;反卷积;空洞卷积

**中图分类号:**TP391.4 **文献标志码:**A

### Image super-resolution reconstruction based on hybrid deep convolutional network

HU Xueying<sup>1,2</sup>, GUO Hairu<sup>1</sup>, ZHU Rong<sup>2\*</sup>

(1. College of Computer Science and Technology, Henan Polytechnic University, Jiaozuo Henan 454000, China;

2. College of Mathematics and Information Engineering, Jiaxing University, Jiaxing Zhejiang 314000, China)

**Abstract:** Aiming at the problems of blurred image, large noise, and poor visual perception in the traditional image super-resolution reconstruction methods, a method of image super-resolution reconstruction based on hybrid deep convolutional network was proposed. Firstly, the low-resolution image was scaled down to the specified size in the up-sampling phase. Secondly, the initial features of the low-resolution image were extracted in the feature extraction phase. Thirdly, the extracted initial features were sent to the convolutional coding and decoding structure for image feature denoising. Finally, high-dimensional feature extraction and computation were performed on the reconstruction layer using the dilated convolution in order to reconstruct the high-resolution image, and the residual learning was used to quickly optimize the network in order to reduce the noise and make the reconstructed image have better definition and visual effect. Based on the Set14 dataset and scale of 4x, the proposed method was compared with Bicubic interpolation (Bicubic), Anchored neighborhood regression (A+), Super-Resolution Convolutional Neural Network (SRCNN), Very Deep Super-Resolution network (VDSR), Restoration Encoder-Decoder Network (REDNet). In the super-resolution experiments, compared with the above methods, the proposed method has the Peak Signal-to-Noise Ratio (PSNR) increased by 2.73 dB, 1.41 dB, 1.24 dB, 0.72 dB and 1.15 dB respectively, and the Structural SIMilarity (SSIM) improved by 0.067 3, 0.020 9, 0.019 7, 0.002 6 and 0.004 6 respectively. The experimental results show that the hybrid deep convolutional network can effectively perform super-resolution reconstruction of the image.

**Key words:** image super-resolution reconstruction; image feature denoising; hybrid deep convolutional network; deconvolution; dilated convolution

## 0 引言

图像超分辨率(Super Resolution, SR)重建旨在通过输入一幅或多幅低分辨率(Low Resolution, LR)图像重建出包含丰富细节的高分辨率(High Resolution, HR)图像。重建技术作

为一个不适定的逆问题,其过程需要尽可能收集和分析更多的相邻像素以获取更多的线索,用于上采样过程中补充丢失的像素信息。而单图像超分辨率(Signal Image Super Resolution, SISR)<sup>[1-2]</sup>重建则是要利用一幅图像中包含的丰富信息以及从样本图像中得到的视觉先验,识别重要的视觉线

收稿日期:2019-12-24;修回日期:2020-03-09;录用日期:2020-03-11。

基金项目:浙江省重点研发计划项目(2019C03099);浙江省自然科学基金资助项目(LY19F020017)

作者简介:胡雪影(1993—),女,河南信阳人,硕士研究生,主要研究方向:智能信息处理、图像处理;郭海儒(1977—),男,山西朔州人,副教授,博士,CCF会员,主要研究方向:智能信息处理、图像处理;朱蓉(1973—),女,浙江嘉兴人,教授,博士,CCF会员,主要研究方向:智能信息处理、机器学习、数据挖掘。

索,填充细节,并尽可能忠实和美观地呈现。近年来,重建技术在医疗影像、卫星遥感、视频监控等领域都有着广泛的应用。

在计算机视觉任务中,为由低分辨率图像重建出高分辨率图像,研究者们已经提出多种 SISR 方法,主要可分为传统的 SISR 方法<sup>[3-4]</sup>和深度学习方法<sup>[5]</sup>。传统的 SISR 方法是基于经验算法实现,会遗忘大量的高频信息且需要一定量人为干预,方法的性能结果与实际应用要求有着较大的差距。

由于出色的性能,深度学习方法已成为模式识别与人工智能领域研究的热点,其中的卷积网络(Convolutional Network, ConvNet)<sup>[6]</sup>近年来更是在计算机视觉任务中取得巨大成功。Dong 等<sup>[7]</sup>提出的超分辨率卷积神经网络(Super-Resolution Convolutional Neural Network, SRCNN)使用了三层卷积网络首次将深度学习运用到图像超分辨率重建中,网络以端到端的方式学习从 LR 到 HR 的映射,不需要传统方法中的任何工程特性,并且获得比传统方法更先进的性能。Kim 等随后提出的深度递归卷积网络(Deeply-Recursive Convolutional Network, DRCN)<sup>[8]</sup>和极深度超分辨率网络(Very Deep Super-Resolution network, VDSR)<sup>[9]</sup>取得了当时最先进的性能。Wang 等<sup>[10]</sup>在生成对抗网络为主体的增强型超分辨率生成对抗网络(Enhanced Super-resolution Generative Adversarial Networks, ESRGAN)中,融合感知损失提高对低分辨率图像的特征重构能力,还原图像缺失的高频语义信息,取得了逼真视觉效果;Mao 等<sup>[11]</sup>提出具有对称结构的编解码网络(Restoration Encoder-Decoder Network, REDNet)对 SR 过程产生的噪声有抑制作用,能恢复出更加干净的图片。残差通道注意力网络(Residual Channel Attention Network, RCAN)<sup>[12]</sup>通过在卷积网络引入注意力机制,通过考虑卷积特征通道之间的相互依赖性来自适应地重新缩放通道方式的特征;Hui 等<sup>[13]</sup>提出的信息蒸馏网络(Information Distillation Network, IDN),在信息蒸馏模块里,局部特征可被有效提取且有着轻量级参数和计算复杂度。深度 SR 方法优秀的性能在特定的图像超分辨中得到应用<sup>[14-16]</sup>。

但上述方法在成功将深度学习技术引入超分辨率问题的同时,也具有一定的局限性。SRCNN 存在着依赖于小图像区域的上下文,训练收敛太慢,网络只适用于单尺度的缺陷;VDSR 虽然解决了 SRCNN 存在的问题,但 VDSR 较小的感受野使得其网络卷积层简单链式堆叠,鲁棒性与泛化性得不到保证;DRCN 则由于使用了循环卷积网络,导致训练时间长,存储量大,不便于实际应用;而 ESRGAN 采用对抗的方式生成 SR 图像,破坏了 SR 图像与 HR 图像像素间的对应关系,且网络规模庞大不利于训练与使用。

针对上述 SR 方法存在的问题,本文提出一种基于混合深度卷积网络的图像超分辨率重建方法。该方法中卷积与反卷积构建的编解码结构去除图像 SR 过程产生的噪声,空洞卷积扩大卷积感受野,通过不同卷积方式的混合使用构建端到端网络,重建出符合原高分辨率图像的超分辨率图像。

与前述 SISR 方法相比,本文所提方法具有如下特点:

- 1) 网络直接学习低分辨率和高分辨率图像之间的端到端映射,采用卷积与反卷积级联编解码方式,消解图像特征中含有的噪声。
- 2) 在图像重建部分中应用空洞卷积层,使小的卷积核获得大的感受野,提升重建效果的同时降低计算量。
- 3) 使用残差学习快速优化网络,通过在网络中添加多个跳跃连接,加速训练过程并提高重建结果性能。

## 1 相关工作

### 1.1 高分辨率图像成像模型

实验研究中,低分辨率图像成像模型通常表示为:

$$I^L = DI^H B + n \quad (1)$$

其中: $I^L$ 为低分辨率图像; $I^H$ 为高分辨率图像; $D$ 为下采样算子; $B$ 为模糊算子; $n$ 为加性噪声。

由式(1)推导由低分辨率图像得到高分辨图像的超分辨率重建过程模型:

$$I^H = D^{-1} I^L B^{-1} - D^{-1} n B^{-1} \quad (2)$$

其中: $B^{-1} D^{-1} n$ 可简化为 $S$ (生成噪声),得到式(3):

$$I^H = D^{-1} I^L B^{-1} - S \quad (3)$$

由式(3),经过上采样( $D^{-1}$ )、去模糊( $B^{-1}$ )、降噪( $-S$ )等一系列处理操作,可将低分辨率图像重建为高分辨率图像。

### 1.2 卷积编解码结构用于图像特征去噪

计算机视觉的图像处理非常依赖于图像的特征。图像的特征就是用较少的数据来描述捕获图像内容中包含的重要信息,因此,图像特征本质上可以理解图像的稀疏化表示<sup>[17]</sup>。由此,Zeiler 等<sup>[18]</sup>提出了反卷积网络(DeConvolutional Network, DeCNN)的概念,这是一种基于正则化图像稀疏表示方法,能够提取出图像特征并利用这些特征重构图像;陈扬钦等<sup>[19]</sup>将基于 $L_1$ 正则化反卷积网络模型应用于图像表述与复原,去除图像噪声;Long 等<sup>[20]</sup>在用于图像分割的全卷积网络(Fully Convolutional Network, FCN)模型中使用反卷积层(Deconvolution layer)获得优异效果,突出证明了反卷积网络在图像处理中的巨大作用;Xu 等<sup>[21]</sup>发现真实的模糊退化很少符合理想的线性卷积模型,采用反卷积网络捕获图像退化特征进而复原图像(式(4)),而不是单纯地从生成模型的角度对异常值进行完美建模<sup>[22]</sup>。

$$\sum_{k=1}^{K_1} (Z_k^i \otimes f_{k,c}) = x_c^i \quad (4)$$

其中: $x$ 为输入图片; $c$ 为图片的通道数量; $i$ 为图像像素; $k$ 为特征图数量; $\otimes$ 为卷积操作; $Z$ 为特征图,是局部隐变量,对于每个输入 $x$ 都不一样。 $f$ 是卷积核,为全局变量,对所有输入 $x$ 都一样。

借鉴上述思想及方法,将卷积与反卷积联结使用,并且增加卷积深度,使其在图像超分辨率重建去噪过程中发挥更大的作用,进一步证实较深的卷积与反卷积层联合使用可以显著提高超分辨率的性能。

### 1.3 空洞卷积用于扩大卷积感受野

研究者们对深度学习的不断探索,贡献了很多经典的卷积神经网络。LeCun 等的开拓性成果 LeNet<sup>[23]</sup>,是大量神经网络架构演进的起点,为后续相关深度学习研究领域提供灵感;和 LeNet 相似的 AlexNet<sup>[24]</sup>,不仅解决了网络较深时存在的梯度弥散问题,提高训练速度,而且还增强了模型的泛化能力;继承 LeNet 以及 AlexNet 部分框架结构的 VGG<sup>[25]</sup>更是直接利用了多层小卷积核叠加的优点(3个 3×3 的卷积核的链式叠加正则等效于 1 个 7×7 的卷积核),这种设计不仅可以大幅度地减少参数,其本身带有正则性质的卷积图能够更容易学习一个通用的、可表达的特征空间;残差网络(Residual Network, ResNet)<sup>[26]</sup>则是在网络中增加直连通道,允许原始输入信息直接传到后面的网络层中。这不仅可以极快地加速神经网络的训练,也能比较大地提升模型的结果准确率等。

然而深度卷积神经网络对于其他任务存在着一些致命性的缺陷。例如:下采样/池化过程导致内部数据结构丢失、空间层级化信息丢失、小物体信息无法重建等。而图像超分辨率重建过程需要更多的相邻像素(即依赖大感受野)。因此如何在卷积感受野大小与网络参数数量二者中找到一个平衡点,是一个至关重要的问题。而空洞卷积(扩张卷积)在不使用会导致信息损失的 pooling 层且计算量相当的情况下,提供更大的感受野,这恰恰与图像超分辨率重建的理念(SISR 网络不需要 pooling 层)相重合。因此,空洞卷积的设计恰好可解决上述如何找到最佳平衡点问题。

空洞卷积诞生于图像分割领域<sup>[27]</sup>,图像输入到网络中经

过CNN提取特征,再经过pooling降低图像尺度的同时增大感受野。在图像需要全局信息或者语音文本需要较长的序列信息依赖的问题中,都有很好的应用效果<sup>[20,28-29]</sup>。对比传统的卷积加池化操作,普通多层卷积联结使用,只能达到卷积层数与感受野大小成线性关系,而空洞卷积与卷积的联合使用,可以使感受野与层数呈指数级增长。如图1所示,圆点为卷积点,深色区域为感受野。

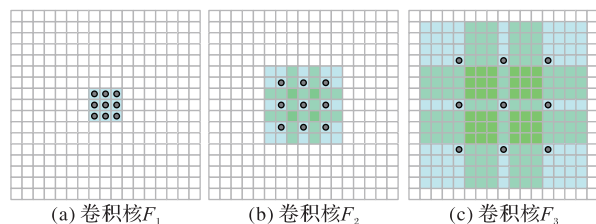


图1 不同比率空洞卷积感受野对比

Fig. 1 Comparison of receptive fields of dilated convolutions with different ratios

将图1中不同比率点距的空洞卷积层联结,级联卷积效果如图2所示。

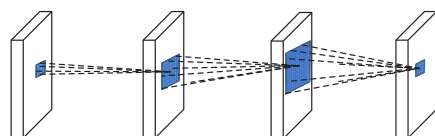


图2 不同比率空洞卷积多层级联

Fig. 2 Multi-cascade of dilated convolutions with different ratios

空洞卷积网络参数数量呈线性增长,而使卷积感受域指数扩展,保证网络参数数量限度的同时增大感受野。图1(a)是由 $3 \times 3$ 卷积核 $F_1$ 通过1倍的卷积产生的,其卷积点间距为1,其每个卷积具有 $3 \times 3$ 的感受野;(b) $F_2$ 是由 $F_1$ 通过2倍的卷积产生的,卷积点间距为2, $F_2$ 中的每个卷积核的感受野为 $7 \times 7$ ;(c) $F_3$ 是由 $F_2$ 通过4倍的卷积产生的,卷积点间距为2, $F_3$ 中的每个卷积核的感受野为 $15 \times 15$ 。与每一层关联的参数数量相同。

在图像超分辨重建网络中,采用卷积与空洞卷积的级联使用,扩大了卷积感受野,在不增加网络容量的前提下能获得

更多相邻像素信息。

#### 1.4 残差学习用于深度网络

传统的卷积层或全连接层在进行信息传递时,或多或少会存在信息损耗丢失问题。He等<sup>[26]</sup>为解决深层网络训练困难的问题提出的残差网络(ResNet)<sup>[26]</sup>在卷积网络中增加直通通道,允许原始输入信息直接传到后面的网络层。不仅可以极大地加速神经网络的训练,并且提升了模型的准确率。确切地说,残差网络通过直接将输入信息绕道传到输出的方式,在某种程度上解决了图像特征细节缺少的问题,保护图像信息的完整性。即整个网络只需要学习输入、输出的残差部分,简化了学习目标和难度。残差块的结构如图3所示。

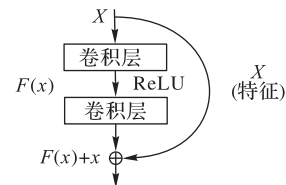


图3 残差结构

Fig. 3 Residual structure

如图3所示,假设原始网络学习函数为 $H(x)$ ,将其分解为 $H(x) = F(x) + x$ 。分解后原始网络(图3垂直向下的流程)拟合 $F(x)$ ,分支 $x$ (图3弯曲部分,为跳跃连接)。在ResNet中引入了残差函数 $F(x) = H(x) - x$ (即目标值与输入值的偏差),通过训练拟合 $F(x)$ ,进而由 $F(x) + x$ 得到 $H(x)$ 。如果 $F(x) = 0$ ,则相当于引入了一个恒等映射到图像超分辨率重建任务中,使用残差学习通过在卷积网络添加跳跃连接可将图像初始特征直接传输到网络中后层,协助梯度的反向传播,加快训练进程,提高结果性能。

## 2 本文方法

### 2.1 网络模型

本文所构建的网络模型主要由四个模块组成:上采样、特征提取、编解码结构去噪和图像重建。上采样模块将低分辨率图片通过简单上采样,得到与高分辨率图像具有相同像素的低分辨率图像;特征提取模块将图像表示为可供处理的特征映射;编解码结构完成图像特征去噪任务;任务完成后,将反卷积去噪中的最终特征映射输入重建模块,最终生成超分辨率重建图像。具体模型如图4所示。

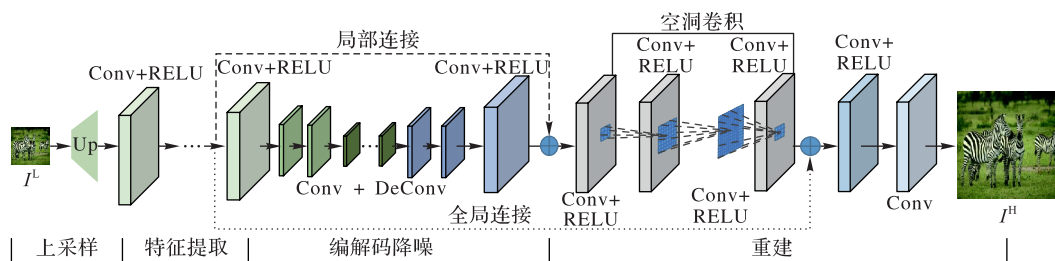


图4 网络架构

Fig. 4 Network architecture

#### 2.1.1 上采样

在基于深度学习的图像超分辨率重建任务中,由高分辨率图像的重建模型(式(3)),需要通过上采样将低分辨率图像放缩至与高分辨率图像具有相同像素数量的次低分辨率图像。而上采样中最简单的方式是重采样和插值:主要通过将输入图片重新缩放到一个想要的尺寸,同时计算每个点的像素点。文献[7-9]即是采用在图像输入网络前使用双三次线性插值对低分辨率图像进行上采样,但此方法形式上增加了人工干预,过多添加工程特征,影响重建效果。受文献[20,

22]的思想启发,本文采用反卷积网络结构将输入低分辨率图像上采样到目标图像大小。此部分使用卷积核为 $3 \times 3$ 的反卷积层将低分辨率图像上采样至目标图像同等尺度大小,作为特征提取层的输入。

#### 2.1.2 图像特征提取

传统图像恢复中,图像特征提取的一种策略是先将图像块进行密集提取<sup>[30]</sup>,之后用一组预先训练的基底(如主成分分析(Principal Component Analysis, PCA)<sup>[31]</sup>、离散余弦变换(Discrete Cosine Transform, DCT)<sup>[32]</sup>等)来表示。而在卷积神

经网络中,此部分可纳入到网络的基础优化中,卷积操作自动提取图像特征。形式上,本文的特征提取子网络表示式(5):

$$F_1(X) = \max(0, W_1 \otimes X + b_1) \quad (5)$$

其中: $X$ 为上采样后低分辨率图像; $W_1$ 和 $b_1$ 分别表示卷积权值和偏置,其中 $W_1$ 大小为 $3 \times 3 \times 64$ ;" $\otimes$ "表示卷积运算,运算添加0边界,步长为1,使输入、输出尺寸保持一致,防止产生边界降秩。同时使用 $\text{RELU}^{[33]}(\max(0, \times))$ 用于卷积特征激活。使用双层64个 $3 \times 3$ 卷积核的卷积层提取上采样后图像的特征,获取到图像特征细节。

### 2.1.3 编解码去噪结构

在特征去噪结构中,卷积与反卷积级联使用构造编解码结构,可以最大限度地消解图像特征噪声。卷积层保留了主要的图像内容,反卷积层则用来补偿细节信息,在达到良好去噪效果的同时较好地保留图像内容<sup>[34]</sup>。针对输入含有噪声图像特征和干净图像,本部分网络结构致力于学习一个残差即 $F(X) = Y - X$ 得到。形式上,这一层形式化表示为:

$$H_1(F_1) = \max(0, W_2 \otimes F_1 + b_2) \quad (6)$$

$$H_2(H_1) = \max(0, W_3 \ominus H_1 + b_3) \quad (7)$$

$$F_2(H_2) = H_2 + F_1 \quad (8)$$

其中: $F_1$ 为特征提取阶段输出; $W_2$ 和 $b_2$ 为卷积权值与偏置大小, $W_2$ 为 $3 \times 3 \times 64$ ;  $W_3$ 和 $b_3$ 为反卷积权值与偏置, $W_3$ 大小为 $3 \times 3 \times 64$ ;" $\ominus$ "表示反卷积运算。式(8)是一个跳跃连接<sup>[26]</sup>,正如残差网络的设计初衷,跳跃连接可以解决网络层数较深的情况下梯度弥散问题,同时有助于梯度的反向传播,加快训练过程。

卷积层逐渐缩小特征图大小,保留了主要的图像内信息,获取图像特征的抽象内容;反卷积层再逐渐增大特征图的大小,放大特征尺寸的同时恢复图像特征细节信息。同时采用跳跃连接加快训练过程,最终在确保编解码结构输入输出尺寸大小一致的同时,也保证了在移动端计算能力有限情况下的测试效率,获得去除噪声的图像特征图。

本部分卷积层全部使用64个 $3 \times 3 \times 64$ 的卷积核。但前一部分的卷积运算不添加0边界且步长为2,使特征输出尺寸变为输入的一半。而在后半部分的反卷积特征恢复中,卷积运算添加0边界、步长为2。将特征图恢复到原来大小,保证图像尺度完整性。

### 2.1.4 图像重建

在重建过程中,输入隐藏状态的特征图 $F_2$ ,输出超分辨率的重建图像,可以看作特征提取阶段的逆运算。在传统的SR方法中,通常平均化重叠高分辨率特征图以产生最终的完整图像。而在网络卷积中,使用卷积核 $W_c^d$ 作为一个反作用基底系数,将高维隐藏状态图像特征的每个位置看作高分辨率图像所对应像素不同维度的矢量形式。反之可以将特征图投影到图像域中,获得超分辨率重建图像。受此启发,定义了一个卷积层来生成最终的高分辨率图像:

$$F_3(F_2) = \max(0, W_4 \otimes F_2 + b_4) \quad (9)$$

$$F_4(F_3) = F_3 + F_1 \quad (10)$$

$$I^{\text{SR}} = W_c^d \otimes F_4 + b_5 \quad (11)$$

式(9)为形式化的空洞卷积, $W_4$ 为空洞卷积的卷积权值,设置rate分别为1、1.2、4。式(10)是一个跳跃连接,提高训练速度;式(11)为超分辨率图像重建, $I^{\text{SR}}$ 为输出的超分辨率重建图像, $W_c^d$ 为 $c$ 个大小为 $1 \times 1 \times 64$ 的卷积核, $c$ 代表着图像通道数量, $b_5$ 表示卷积偏置。

本部分前4层卷积使用64个 $3 \times 3 \times 64$ 的卷积核对图像特征进行高维提取运算,其卷积核卷积点间距分别为1、1.2、4,可以以更大感受野对图像特征进行运算。随后使用64个 $3 \times 3 \times 64$ 卷积核构成的卷积层,将初次提取到的图像特征与高维特征进行加和运算,保证了特征的充分利用。最后使用 $C$ ( $C$ 为高分辨率重建图像通道数量)个 $1 \times 1 \times 64$ 卷积核的卷积层重

建出需要的高分辨率图像。

本文通过将上述四种不同功能驱动模块(它们都与卷积网络具有相似的形式)联结在一起,形成一个具有去噪、增大感受野的混合深度卷积神经网络,用作图像超分辨率重建任务。

## 2.2 训练

### 2.2.1 损失函数与优化

对于任意给定一个训练数据集 $\{X_C^1, Y_C^1\}_{i=0}^{n-1}$ ,本文的目标是找到准确映射值 $Y = F(x)$ ,使获得的超分辨率重建图像 $F(x)$ 与真实高分辨率图像 $Y$ 之间的均方误差(Mean Squared Error, MSE)<sup>[35]</sup>最小,同时也有利于图像质量评价指标——峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)<sup>[36]</sup>的提高。尽管高PSNR并不代表重建图像的绝对优秀,但当使用替代评估指标(如SSIM(Structural Similarity))来评估模型时,仍然观察到满意的性能。

$$MSE = \frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} (P(i, j) - T(i, j))^2 \quad (12)$$

其中: $P(i, j)$ 、 $T(i, j)$ 分别表示预测图像与真实图像; $H$ 、 $W$ 分别为图像的高度和宽度。

$$f_{\text{loss}} = MSE + \beta \|\theta\|^2 \quad (13)$$

即:

$$f_{\text{loss}} = \frac{1}{N} \sum_{i=0}^{N-1} \|F(X) - Y\|^2 + \beta \|\theta\|^2 \quad (14)$$

在本文中,将 $MSE$ 作为一部分并入损失函数 $f_{\text{loss}}$ (式(14))。通过添加 $L_2$ 权重衰减对训练进行正则化, $\beta$ 表示权重衰减的乘系数,设为 $10^{-3}$ , $\theta$ 即为所求参数。这样在得到最终损失的条件上,可以微调节参数来尽可能地达到最优结果。

### 2.2.2 优化方式

本文使用Adam优化方法<sup>[37]</sup>更新权重矩阵,以将损失函数数值降至最低。其参数更新公式为:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t + \varepsilon}} \hat{m}_t \quad (15)$$

式(15)中, $\hat{m}_t$ 与 $\hat{v}_t$ 分别为:

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

其中: $\beta_1$ 设为0.9; $\beta_2$ 设为0.9999; $\varepsilon$ 设为 $10^{-8}$ ;学习率 $\eta$ 为0.001; $m_t$ 为梯度的第一时刻平均值; $v_t$ 为梯度的第二时刻方差值。

### 2.2.3 参数设置

在训练之前,需要初始化网络参数以供后续网络训练更新。对于权重初始化,结合Kin等<sup>[9]</sup>描述的方法。本文从平均值为零、标准偏差为0.001(偏差为0)的正态分布中随机抽取数值,初始化每层的卷积核权重 $W$ ,而偏置 $b$ 则全部置0。初始学习率被设置为0.001,之后每训练20轮递减1/2。

实验的硬件环境为AMD Ryzen 2700X八核的处理器; Nvidia GeForce GTX 1080 Ti GPU;软件环境为Ubuntu 16.04的操作系统。网络模型采用tensorflow框架构成。实验训练了网络参数100轮次(每轮次1000次迭代,迭代批次大小32),100个轮次后停止学习。

## 3 实验结果与分析

### 3.1 数据集

训练集 使用公开自然图像数据集BSD200<sup>[38]</sup>(200个图像)与T91数据集<sup>[39]</sup>(91个图像),其适用于大多数实验。在SRCNN/VDSR中已证明深度模型通常受益于海量数据的训练。实验中将291个训练图像进行旋转90度、镜像、倒置等操作,获得包含1164个图像的增强训练数据集;同时在增强训

训练数据集上裁剪 64×64 大小(裁剪步幅为 16)的子图送入网络,便于学习网络参数。将网络模型应用于任意大小的图像,因为它是基于完全卷积的,而裁剪可更加方便地优化网络模型参数。

测试集 使用通常用于基准测试<sup>[7,40]</sup>的数据集 Set5<sup>[41]</sup>(5 个图像)和 Set14<sup>[42]</sup>(14 个图像)、BSD100<sup>[38]</sup>(100 个图像)以及由 Huang 等<sup>[43]</sup>提供的城市图像 Urban100 数据集。

由于卷积网络的特殊性,其只能将图像放缩双倍大小。本文在测试数据集评价了两个尺度因子,包括×2和×4。

### 3.2 定性定量分析

#### 1) 编解码结构降噪。

在本文所提方法中,用编解码结构卷积提取图像特征,保留主要特征;用反卷积针对特征上采样,恢复图像细节,从而完成图像特征噪声过滤,进而达到图像特征降噪的目的。

如图 5(a)显示了有反卷积构建编解码结构与不具有反卷

积网络对比示例。当没有反卷积层时,网络模型为标准卷积网络。通过观察对比可发现含有反卷积编解码去噪部分的网络可以达到更高的精度(以 PSNR 为指标)。

#### 2) 空洞卷积。

空洞卷积不改变卷积运算本身的运算方式与参数,通过修改卷积核的构造以不同的方式使用过滤器参数。空洞卷积使用不同的扩张比例因子在不同的范围内应用相同的卷积核参数,以达到获取更多图像上下文信息的目的。显然,第一层的特征图包含不同的结构(例如,不同方向的边),而第二层的特征图在强度上主要不同。

如图 5(b)展示了具有空洞卷积与不具有空洞卷积网络对比示例,空洞卷积通过扩张卷积感受野促进了性能的提高。

最后,图 5(c)中展示了同时使用反卷积和空洞卷积与二者皆不使用的性能对比曲线,发现当二者同时使用时对性能提升效果显著。

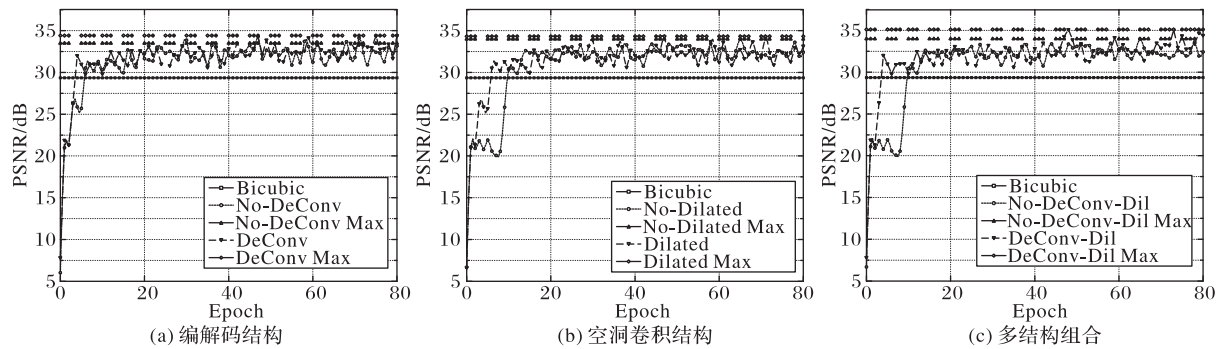


图 5 各网络结构性能结果分析

Fig. 5 Performance result analysis of different network structures

#### 3) 残差学习。

输入图像在网络计算中携带着大量的细节信息,但如果输入与输出之间存在许多卷积递归,就会产生梯度问题,如梯度消失/爆炸等。它们之间的映射关系学习是非常困难的,并且学习效率非常低。此时,可以通过设置残差学习快速解

决这个问题。

在网络中添加了两个跳跃连接,如图 4 所示:其一为特征去噪部分局部连接,协助梯度的反向传播,加快训练过程;另一个则为远程跳跃连接,加快训练过程,提高结果性能。具体对比分析结果见表 1。

表 1 残差学习测试(学习率:0.001)

Tab. 1 Residual learning testing (learning rate: 0.001)

| 残差结构                            | 倍数 | Epoch 为 10 |        | Epoch 为 20 |        | Epoch 为 40 |        | Epoch 为 60 |        | Epoch 为 80 |        |
|---------------------------------|----|------------|--------|------------|--------|------------|--------|------------|--------|------------|--------|
|                                 |    | PSNR/dB    | SSIM   | PSNR/dB    | SSIM   | PSNR/dB    | SSIM   | PSNR/dB    | SSIM   | PSNR/dB    | SSIM   |
| No-Residual                     | ×2 | 13.89      | 0.3757 | 14.22      | 0.3757 | 14.08      | 0.4229 | 14.05      | 0.3971 | 14.37      | 0.3929 |
|                                 | ×4 | 14.18      | 0.3942 | 13.20      | 0.3975 | 14.20      | 0.3844 | 13.44      | 0.3669 | 14.37      | 0.4035 |
| Single-Residual<br>(local)      | ×2 | 29.23      | 0.9018 | 30.96      | 0.7864 | 32.08      | 0.9124 | 32.60      | 0.9080 | 31.87      | 0.9008 |
|                                 | ×4 | 27.34      | 0.7244 | 28.15      | 0.7324 | 28.66      | 0.7438 | 28.41      | 0.7432 | 28.34      | 0.7387 |
| Single-Residual<br>(long range) | ×2 | 28.13      | 0.8982 | 31.48      | 0.8888 | 34.02      | 0.9177 | 32.96      | 0.9046 | 32.87      | 0.9118 |
|                                 | ×4 | 27.79      | 0.7323 | 27.54      | 0.7744 | 28.97      | 0.7497 | 28.16      | 0.7472 | 29.26      | 0.7549 |
| Multi-Residual                  | ×2 | 29.54      | 0.8988 | 33.59      | 0.9155 | 34.22      | 0.9194 | 35.38      | 0.9285 | 34.51      | 0.9228 |
|                                 | ×4 | 25.25      | 0.6766 | 26.30      | 0.7067 | 29.31      | 0.7743 | 28.64      | 0.7403 | 29.74      | 0.7715 |

表 1 实验采用全卷积网络,在数据集 Urban100 上针对比例因子 2 和 4 在不同训练轮次下的超分辨率重建定性定量分析结果,学习率设为 0.001,评测指标为 PSNR/SSIM。经过对比分析可以发现,含有跳跃连接的网络无论是训练速度还是评测指标都比非残差学习网络的结果更加优异,而在未含跳跃连接的网络中,不仅训练收敛困难,并且结果很差。而这也证明了残差学习在模型训练的表现良好。通过对比,发现含有单跳跃连接的网络中,远程连接比局部连接的结果性能更好,而训练收敛速度次之。在含有多跳跃连接的网络,无论是收敛速度还是结果性能都比单连接网络结果优异。

综上,使用残差学习所学习映射的效果要比直接学习效果相对较好,学习速度更高。

### 3.3 结果分析

将本文所提方法与其他几种已有先进图像超分辨率重建方法在多个公开数据集上进行图像重建结果对比,评估其特性与性能。用于对比的几种方法有:1)Bicubic:像素插值方法<sup>[44]</sup>;2)A+:调整锚定邻域回归法<sup>[40]</sup>;3)SRCNN:简单的端到端全卷积图像超分辨率方法<sup>[7]</sup>;4)VDSR:带有残差学习的深度全卷积图像超分辨率处理网络,全卷积网络图像超分辨率重建效果最好方法<sup>[9]</sup>;5)REDNet:由卷积层-反卷积层构成的对称的编解码框架,对图像超分辨与去噪表现出优异效果<sup>[11]</sup>。

在表 2 中,显示了多种优秀 SR 方法在多个公开数据集上进行测试结果。通过对比可发现,在这些数据集中,本文提出的方法明显优于其他传统 SR 方法如 A+等,且对比其他网络 SR 方法,在 PSNR 和 SSIM 上均有提高。

表 2 各 SR 方法性能对比

Tab. 2 Performance comparison of different SR methods

| 数据集      | 倍数 | Bicubic |        | A+      |        | SRCNN   |        | VDSR    |        | REDNet  |        | 本文方法    |        |
|----------|----|---------|--------|---------|--------|---------|--------|---------|--------|---------|--------|---------|--------|
|          |    | PSNR/dB | SSIM   | PSNR/dB | SSIM   | PSNR/dB | SSIM   | PSNR/dB | SSIM   | PSNR/dB | SSIM   | PSNR/dB | SSIM   |
| Set5     | ×2 | 33.66   | 0.9099 | 36.54   | 0.9544 | 36.66   | 0.9542 | 37.53   | 0.9587 | 37.43   | 0.9590 | 37.78   | 0.9598 |
|          | ×4 | 28.42   | 0.8104 | 30.28   | 0.8603 | 30.48   | 0.8628 | 31.35   | 0.8838 | 31.12   | 0.8794 | 32.13   | 0.9042 |
| Set14    | ×2 | 30.24   | 0.8688 | 32.28   | 0.9056 | 32.42   | 0.9063 | 33.03   | 0.9124 | 32.77   | 0.9125 | 33.39   | 0.9074 |
|          | ×4 | 26.00   | 0.7027 | 27.32   | 0.7491 | 27.49   | 0.7503 | 28.01   | 0.7674 | 27.58   | 0.7654 | 28.73   | 0.7700 |
| BSD100   | ×2 | 29.56   | 0.8431 | 31.21   | 0.8863 | 31.36   | 0.8879 | 31.90   | 0.8960 | 31.99   | 0.8974 | 32.13   | 0.9095 |
|          | ×4 | 25.96   | 0.6675 | 26.82   | 0.7087 | 26.90   | 0.7101 | 27.29   | 0.7251 | 27.40   | 0.7290 | 28.15   | 0.7324 |
| Urban100 | ×2 | 26.88   | 0.8403 | 29.20   | 0.8938 | 29.50   | 0.8946 | 30.76   | 0.9140 | 30.56   | 0.8623 | 31.48   | 0.9108 |
|          | ×4 | 23.14   | 0.6577 | 24.32   | 0.7183 | 24.52   | 0.7221 | 25.18   | 0.7524 | 24.80   | 0.7318 | 26.29   | 0.7370 |

图 6(a)返回的是 Set14 数据集第 008 号图像使用了放大因子为 4 的超分辨重建结果,观测到本文方法返回结果图在抗锯齿方面表现更好;图 6(b)返回的是在 BSD100 数据集第 43 号图像使用了超分因子为 2 的超分辨重建结果,观测到本

文的方法在船顶部分细节恢复得更加锐利与清晰;图 6(c)返回的是在 Urban100 数据集第 82 号图像使用了比例因子为 4 的超分辨重建结果。相比其他超分辨率方法,可以观测到本文的方法在人物腿部部分恢复得更清晰与连续。

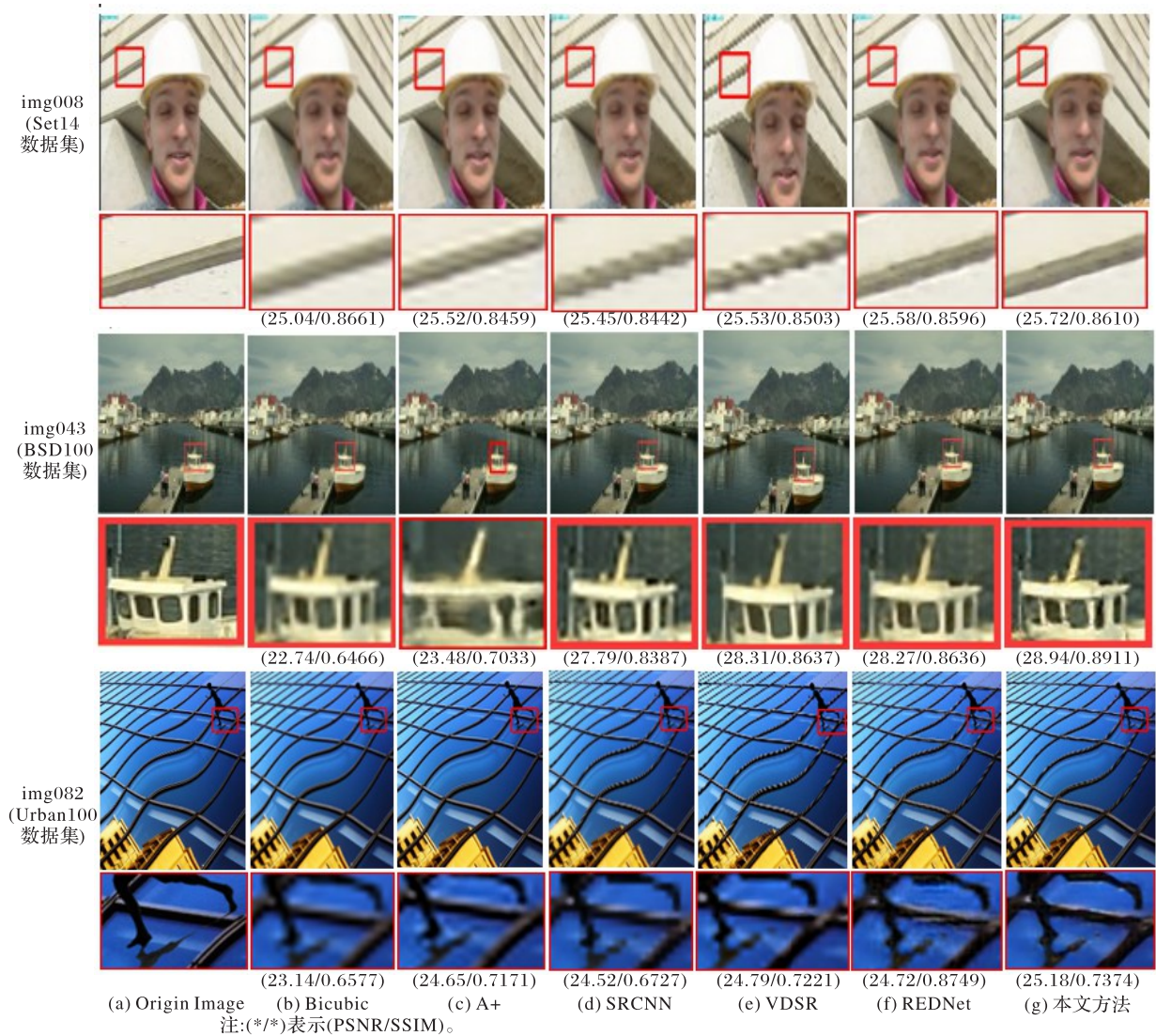


图 6 各 SR 方法重建结果对比

Fig. 6 Comparison of reconstruction results by different SR methods

通过对比发现,无论在客观数据还是在主观视觉效果对比上,本文所获取的重建图片效果明显优于其他先进 SR 方法,因此,本文所提方法相较于其他 SR 方法有明显性能优异之处。

## 4 结语

本文提出了一种带有去噪功能的混合深度卷积网络方法,用于图像超分辨率重建。首先,探索了卷积与反卷积构建

的编解码结构对图像特征噪声的消解效果;接着,探讨了空洞卷积对卷积感受野的提升作用;然后,发现残差学习不仅能更快优化网络,而且能提升图像重建效果。在Set5、Set14、Urban100以及BSD100等公开数据集上进行图像重建对比,表明所提方法比现有的方法在基准图像上有更好的效果。

但是本文所提算法在取得优异结果的情况下,也存在着一些不足。由于反卷积结构的限制性,本文只能将低分辨率图像重建为偶数倍的超分辨率图像,而奇数放大倍数则表现不好,在之后的研究中,将通过优化上采样方式来解决这一问题。

#### 参考文献 (References)

- [1] GLASNER D, BAGON S, IRANI M. Super-resolution from a single image [C]// Proceedings of the IEEE 12th International Conference on Computer Vision. Piscataway: IEEE, 2009:349-356.
- [2] IRANI M, PELEG S. Improving resolution by image registration [J]. CVGIP: Graphical Models and Image Processing, 1991, 53(3):231-239.
- [3] 孔玲莉,黄华,齐春,等. 图像超分辨率研究的最新进展[J]. 光学技术, 2004, 30(3): 374-377. (KONG L L, HUANG H, QI C, et al. Recent development in super-resolution [J]. Optica: Technique, 2004, 30(3): 374-377. )
- [4] 江静,张雪松. 图像超分辨率重建算法综述[J]. 红外技术, 2012, 34(1): 24-30. (JIANG J, ZHANG X S, A review of super-resolution reconstruction algorithms [J]. Infrared Technology, 2012, 34(1): 24-30. )
- [5] HA V K, REN J, XU X, et al. Deep learning based single image super-resolution; a survey [J]. International Journal of Automation and Computing, 2019, 16(4):413-426.
- [6] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [C]// Proceedings of the 13th European Conference on Computer Vision, LNCS 8689. Cham: Springer, 2014: 818-833.
- [7] 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, 2016, 38(2):295-307.
- [8] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 1637-1645.
- [9] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 1646-1654.
- [10] WANG X, YU K, WU S, et al. ESRGAN: enhanced super-resolution generative adversarial networks [C]// Proceedings of the 15th European Conference on Computer Vision, LNCS 11133. Cham: Springer, 2018: 63-79.
- [11] MAO X, SHEN C, YANG Y. Image restoration using convolutional auto-encoders with symmetric skip connections [C]// NIPS 2016: Advances in Neural Information Processing Systems 29. North Miami Beach, FL: Curran Associates Inc., 2016: 2802-2810.
- [12] ZHANG Y, LI K, LI K, et al. Image super-resolution using very deep residual channel attention networks [C]// Proceedings of the 15th European Conference on Computer Vision, LNCS 11211. Cham: Springer, 2018:294-310.
- [13] HUI Z, WANG X, GAO X. Fast and accurate single image super-resolution via information distillation network [C]// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018:723-731.
- [14] 邵保泰,汤心溢,金璐,等. 基于生成对抗网络的单帧红外图像超分辨率算法[J]. 红外与毫米波学报, 2018, 37(4): 427-432. (SHAO B T, TANG X Y, JIN L, et al. Single frame infrared image super-resolution algorithm based on generative adversarial nets [J]. Journal of Infrared and Millimeter Waves, 2018, 37(4): 427-432. )
- [15] 徐军,刘慧,郭强,等. 结合反卷积的CT图像超分辨重建网络[J]. 计算机辅助设计与图形学学报, 2018, 30(11): 2084-2092 (XU J, LIU H, GUO Q, et al. Super-resolution reconstruction of CT images using neural network combined with deconvolution [J]. Journal of Computer-Aided Design and Computer Graphics, 2018, 30(11):2084-2092. )
- [16] 朱福珍,刘越,黄鑫,等. 改进的稀疏表示遥感图像超分辨重建[J]. 光学精密工程, 2019, 27(3): 718-725. (ZHU F Z, LIU Y, HUANG X, et al. Remote sensing image super-resolution based on improved sparse representation [J]. Optical Precision Engineering, 2019, 27(3): 718-725. )
- [17] 邓承志. 图像稀疏表示理论及其应用研究[D]. 武汉:华中科技大学, 2008:18. (DENG C Z. Image sparse representation theory and its application [D]. Wuhan: Huazhong University of Science and Technology, 2008:18. )
- [18] ZEILER M D, KRISHNAN D, TAYLOR G W, et al. Deconvolutional networks [C]// Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2010:2528-2535.
- [19] 陈扬钦,钟平. 基于L1正则化反卷积网络的遥感图像表述与复原方法[J]. 数字技术与应用, 2011(5):158-160, 184. (CHEN Y T, ZHONG P. Remote sensing image representation and restoration method based on L1 regularized deconvolution network [J]. Digital Technology and Application, 2011(5): 158-160, 184. )
- [20] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, Piscataway: IEEE, 2017, 39(4):640-651.
- [21] XU L, REN J S J, LIU C, et al. Deep convolutional neural network for image deconvolution [C]// Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014:1790-1798.
- [22] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. [2019-09-12]. <https://arxiv.org/pdf/1511.06434.pdf>.
- [23] LECUN Y L, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [24] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]// Proceedings of the 25th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2012: 1097-1105.
- [25] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2019-09-12]. <https://arxiv.org/pdf/1409.1556.pdf>.
- [26] HE K, ZHANG X, REN S, et al. Deep residual learning for im-

- age recognition[C]// Proceedings of the 2016 Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016:770-778.
- [27] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions [EB/OL]. [2019-09-12]. <https://arxiv.org/pdf/1511.07122.pdf>.
- [28] KALCHBRENNER N, ESPEHOLT L, SIMONYAN K, et al. Neural machine translation in linear time [EB/OL]. [2019-09-12]. <https://arxiv.org/pdf/1610.10099.pdf>.
- [29] VAN DEN OORD A, DIELEMAN S, ZEN H, et al. WaveNet: a generative model for raw audio [EB/OL]. [2019-09-08]. <https://arxiv.org/pdf/1609.03499.pdf>.
- [30] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311-4322.
- [31] HERVÉ A, WILLIAMS L J. Principal component analysis [J]. WIREs Computational Statistics, 2010, 2(4):433-459.
- [32] RHEE S, KANG M G. Discrete cosine transform based regularized high-resolution image reconstruction algorithm [J]. Optical Engineering, 1999, 38(8): 1348-1356.
- [33] NAIR V, HINTON G E. Rectified linear units improve restricted Boltzmann machines [C]// Proceedings of the 27th International Conference on Machine Learning. Madison, WI: Omnipress, 2010: 807-814.
- [34] MAO X, SHEN C, YANG Y, et al. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections [C]// Proceedings of the 30th International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc., 2016: 2810-2818.
- [35] 沈恒范. 概率论与数理统计教程[M]. 北京:高等教育出版社, 2006: 56-58. (SHEN H F. Probability Theory and Mathematical Statistics Course [M]. Beijing: Higher Education Press, 2006: 56-58.)
- [36] TURAGA D S, CHEN Y, CAVIEDES J, et al. No reference PSNR estimation for compressed pictures [J]. Signal Processing: Image Communication, 2004, 19(2): 173-184.
- [37] KINGMA D P, BA J. Adam: a method for stochastic optimization [EB/OL]. [2019-09-12]. <https://arxiv.org/pdf/1412.6980.pdf>.
- [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 [C]// Proceedings of the 8th IEEE International Conference on Computer Vision. Piscataway: IEEE, 2002:416-423.
- [39] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19(11):2861-2873.
- [40] TIMOFTE R, DE SMET V, VAN GOOL L. A+: adjusted anchored neighborhood regression for fast super-resolution [C]// Proceedings of the 12th Asian Conference on Computer Vision, LNCS 9006. Cham: Springer, 2014:111-126.
- [41] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding [C]// Proceedings of the 2012 British Machine Vision Conference. Durham: BMVA Press, 2012: No. 135.
- [42] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations [C]// Proceedings of the 7th International Conference on Curves and Surfaces, LNCS 6920. Berlin: Springer, 2010:711-730.
- [43] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars [C]// Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015:5197-5206.
- [44] KEYS R. Cubic convolution interpolation for digital image processing [J]. IEEE Transactions on Acoustics Speech, and Signal Processing, 1981, 29(6):1153-1160.

This work is partially supported by the Zhejiang Key Research and Development Program (2019C03099), the Zhejiang Natural Science Foundation (LY19F020017).

**HU Xueying**, born in 1993, M. S. candidate. Her research interests include intelligent information processing, image processing.

**GUO Hairu**, born in 1977, Ph. D., associate professor. His research interests include intelligent information processing, image processing.

**ZHU Rong**, born in 1973, Ph. D., professor. Her research interests include intelligent information processing, machine learning, data mining.