

# 基于深层残差网络的加速图像超分辨率重建

席志红\*, 侯彩燕, 袁昆鹏, 薛卓群

哈尔滨工程大学信息与通信工程学院, 黑龙江 哈尔滨 150001

**摘要** 针对目前卷积神经网络的超分辨率算法存在卷积层数少、模型简单、计算量大、收敛速度慢以及图像纹理模糊等问题,提出了一种基于深层残差网络的加速图像超分辨率重建方法,该方法在提高图像分辨率的同时加快收敛速度。设计更深的卷积神经网络模型来提高精确度,通过残差学习并且使用 Adam 优化方法使网络模型加速收敛。在原始低分辨率图像上直接进行特征映射,只在网络的末端引入子像素卷积层,将像素进行重新排列,得到高分辨率图像。实验结果表明,在 set 5, set 14, BSD100 测试集上,所提算法的峰值信噪比与结构相似性指数均高于现有的几种算法,能够恢复更多的图像细节,图像边缘也更加完整且收敛速度更快。

**关键词** 图像处理; 超分辨率; 深度学习; 卷积神经网络; 残差网络; 子像素卷积

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201939.0210003

## Super-Resolution Reconstruction of Accelerated Image Based on Deep Residual Network

Xi Zhihong\*, Hou Caiyan, Yuan Kunpeng, Xue Zhuoqun

College of Information and Communication Engineering, Harbin Engineering University, Harbin, Heilongjiang 150001, China

**Abstract** An accelerated image super-resolution reconstruction algorithm based on deep residual network is proposed to solve some existing problems, such as few convolutional layers, simple model, large amount of calculation, slow convergence speed and fuzzy image texture. This method improves image resolution and accelerates convergence speed at the same time. First, a deep convolutional neural network model is proposed to improve accuracy, and accelerate convergence of network models is achieved by residual learning and Adam optimization method. Second, feature mapping is performed directly on the original low-resolution image, and the sub-pixel convolutional layer is introduced at the end of the network to rearrange the pixels, so a high-resolution image is obtained. Experimental results show that the proposed algorithm has higher peak signal-to-noise ratio and structural similarity index than those of existing algorithms on set 5, set 14 and BSD100 test sets, and can recover more image details. The image edges are complete and the convergence speed is fast.

**Key words** image processing; super-resolution; deep learning; convolution neural network; residual network; sub-pixel convolution

**OCIS codes** 100.6640; 100.3010; 100.2000

## 1 引 言

超分辨率图像重建技术作为图像处理领域的研究热点,是指利用一幅或者多幅低分辨率图像重建出一幅高分辨率图像<sup>[1-3]</sup>。作为一种无需改善硬件设备即可显著提高图像质量的方法,图像超分辨率重建技术在视频监控、高清晰数字电视、医学成像、遥感图像等领域中得到了广泛应用。图像超分辨率重建技术

可以分为基于插值<sup>[4]</sup>、基于重建<sup>[5]</sup>和基于学习<sup>[6-7]</sup>的方法。通过基于插值的方法得到的图像容易产生模糊、锯齿现象,常见的插值方法有最近邻插值和双三次插值方法等。基于重建的方法在利用低分辨率作为约束的前提下,结合图像的先验知识进行还原,主要包括配准与重建两个部分,常见的算法有迭代反向投影法<sup>[8]</sup>和凸集投影法<sup>[9]</sup>等。现在主流的方法是基于学习的方法,该方法通过提取高分辨率和低分辨率

收稿日期: 2018-05-03; 修回日期: 2018-08-04; 录用日期: 2018-09-25

基金项目: 国家自然科学基金(60875025)

\* E-mail: xizhihong@hrbeu.edu.cn

图像的特征学习两者之间的映射关系,从而重建高分辨率图像,Chang 等<sup>[10]</sup>提出邻域嵌入方法,之后, Yang 等<sup>[11]</sup>研究了基于稀疏编码的方法, Timofte 等<sup>[12-13]</sup>提出了固定邻域回归(ANR)和调整的固定邻域回归的方法。近几年,基于深度学习的超分辨率方法成为研究热点。2016年, Dong 等<sup>[14]</sup>率先提出基于超分辨率卷积神经网络(SRCNN)的方法,使用了三个卷积层,将图像特征从低分辨率空间非线性映射到高分辨率空间,重建效果优于其他传统方法。为了提高速度, Dong 等<sup>[15]</sup>对 SRCNN 算法进行改进,以低分辨率图像作为输入,在网络最后用反卷积层进行相应比例放大,得到高分辨率图像,提出了基于反卷积的快速图像超分辨率重建(FSRCNN)。之后, Shi 等<sup>[16]</sup>继续对 SRCNN 进行改进,同样是直接在低分辨率图像上进行卷积,通过子像素卷积层将特征图像进行重新排列,得到高分辨率图像,提出了高效子像素卷积神经网络(ESPCN)。为了提取更多的图像特征, Kim 等<sup>[17]</sup>加深了卷积层数,提出了一个 20 层的深度卷积神经网络模型(VDSR),提高了图像的重建质量。

但上述方法仍然存在一些问题,虽然 FSRCNN 和 ESPCN 提高了网络训练速度,但是卷积层数相对较少, ESPCN 只有三层卷积操作,无法充分提取图像更多的细节特征,并且网络模型没有对边界使用零填充,使输出图像在经过卷积操作后尺寸减小,无法充分利用图像的边缘信息。VDSR 虽然加深了网络深度,但是它的输入图像需要进行双三次插值的预处理,使得网络计算量加大、计算速度减慢且网络难以收敛。

针对以上问题,本文提出一种基于深层残差网络的加速图像超分辨率重建方法(DRSR),使用深层网络结构,直接在低分辨率图像上提取特征,由于不需要上采样插值,可以使用较小的卷积核级联更多的卷积层,为了加快网络收敛,避免训练过程中产生梯度弥散和梯度爆炸,使用多路径模式的局部残差学习与多权重的递归学习相结合的模型,并且使用 Adam 优化方法,最后在网络的末端通过子像素卷积层将像素进行重新排列,得到高分辨率图像。为了充分利用图像边缘信息,DRSR 对图像边缘采用零填充,使输出图像与输入图像具有相同的尺寸,相比于其他方法,DRSR 的超分辨率重建效果更好,并且大大加快了收敛速度。

## 2 相关工作

### 2.1 SRCNN 算法

SRCNN 算法使用双三次插值将低分辨率图像

升级到所需的大小,然后通过特征提取、非线性映射和重建,最终输出高分辨率图像。SRCNN 结构图如图 1 所示,图中 Conv 为卷积运算,ReLU 为线性整流函数。

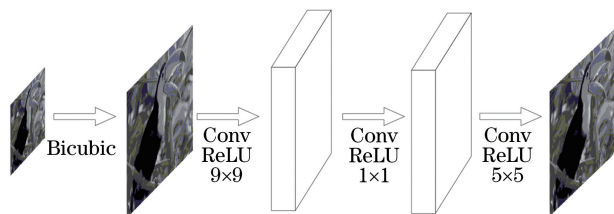


图 1 SRCNN 结构图

Fig. 1 Diagram of SRCNN structure

特征提取的目的是从低分辨率图像中有重叠地提取图像块并用高维向量表示,相当于采用一组滤波器对图像进行卷积操作,非线性映射阶段是将上一层的特征向量从低分辨率空间变换至高分辨率空间,图像重建是将重叠的高分辨率图像块利用平均化操作产生最终的图像,该过程可表示为

$$F_1(Y) = f(W_1 * I^{LR} + b_1), \quad (1)$$

$$F_2(Y) = f[W_2 * F_1(Y) + b_2], \quad (2)$$

$$I^{SR} = F_3(Y) = W_3 * F_2(Y) + b_3, \quad (3)$$

式中  $I^{LR}$  为输入的低分辨率图像,  $W_k$  ( $k=1, 2, 3$ ) 为滤波器,  $*$  表示卷积运算,  $b_k$  为偏差,  $f(x)$  为激活函数,  $F_3(Y)$  为最终的高分辨率图像, 即  $I^{SR}$ 。

上述三个操作结合在一起,形成一个卷积神经网络。参数  $\theta = \{W_1, W_2, W_3, b_1, b_2, b_3\}$ , 该模型对所有的权重和偏重都进行了优化,给定高分辨率数据集  $\{X_i\}$  及其对应的低分辨率数据集  $\{Y_i\}$ , 使用均方误差作为其损失函数, 即

$$L_{oss}(\theta) = \frac{1}{n} \sum_{i=1}^n \|F(Y_i; \theta) - X_i\|^2, \quad (4)$$

式中  $n$  为训练的样本数,  $\{Y_i\}$  需要  $\{X_i\}$  经过下采样后再通过双三次插值放大到与原图像相同的尺寸。

### 2.2 ESPCN 算法

ESPCN 直接在低分辨率图像上提取特征,不需要通过上采样插值,通过插值放大的图像本身会带来额外的干扰信息,在之后的特征提取阶段需要用相对较大的卷积核尺寸,使复杂度增加,而且会受到插值放大带来的额外干扰信息影响,通过在网络的末端用子像素卷积层将像素进行重新排列来得到高分辨率图像,从而大大减小计算复杂度,提高计算速度。ESPCN 结构图如图 2 所示。

卷积层表达式分别为

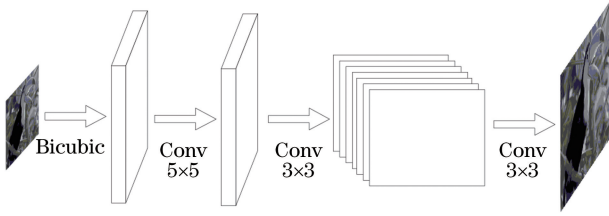


图 2 ESPCN 结构图

Fig. 2 Diagram of ESPCN structure

$$F_1(Y) = f(W_1 * Y^{LR} + b_1), \quad (5)$$

$$F_2(Y) = f[W_2 * F_1(Y) + b_2], \quad (6)$$

$$Y^{SR} = F_3(Y) = P[W_3 * F_2(Y) + b_3], \quad (7)$$

式中:  $P(x)$  为亚像素卷积函数, 即将低分辨率图像特征按照特定位置周期性地插入到高分率图像中;  $Y^{LR}$  为原始低分辨率图像;  $Y^{SR}$  为最终的高分辨率图像。输入、输出图像尺寸分别为  $25 \times 25$  和  $17 \times 17$ 。卷积层的网络参数设置为:  $n_1 = 64, k_1 = 5, n_2 = 32, k_2 = 3, n_3 = 9, k_3 = 3$ 。

### 3 DRSR 算法

本文算法 (DRSR 算法) 使用了较深的卷积网络, 用较小的卷积核进行级联, 直接在低分辨率图

像上提取特征, 由于增加了网络深度, 使用局部残差的网络结构来加快模型收敛, 提高精度, 最后用子像素卷积层将像素进行重新排列, 得到高分率图像, 本文算法的网络结构图如图 3 所示。与 VDSR 的全局残差学习不同, 本文算法使用的是多路径模式的局部残差学习与多权重的递归学习相结合的模式, 其中每个残差单元都共同拥有一个相同的输入, 即递归块中的第一个卷积层的输出, 每个残差单元都包含 2 个卷积层。在一个递归块内, 每个残差单元内对应位置相同的卷积层参数都是共享的。为避免卷积后图像信息丢失, 在卷积过程中使用零填充保证输出图像与输入图像尺寸相等<sup>[18]</sup>。在网络设计中, 池化层常用于图像分类的神经网络中, 通过对卷积后的结果进行下采样, 可减少部分参数, 获得旋转、平移、伸缩等不变性, 超分辨率算法的目标是增加图像的细节信息, 与池化层减少细节信息的目标相悖, 因此本文算法不使用池化层。

#### 3.1 深层神经网络结构

在未经过插值放大的情况下直接对原始低分辨率图像进行特征提取, 由于输入分辨率降低, 可以使

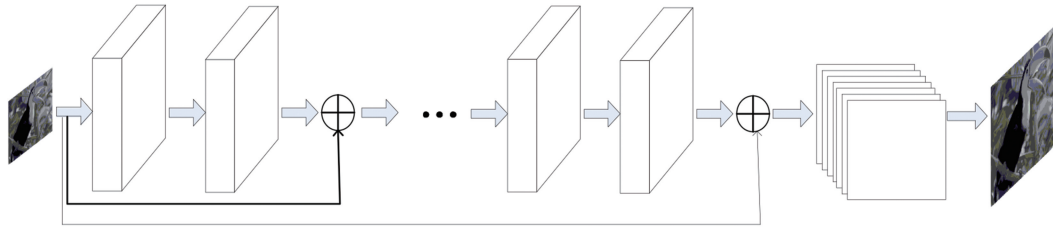


图 3 本文算法网络结构图

Fig. 3 Diagram of network structure of the proposed algorithm

用较小的卷积核来提取相同的信息, 一个  $9 \times 9$  的卷积核等同于 4 个  $3 \times 3$  的卷积核依次级联卷积, 但是由于额外引入了卷积层, 可以有效地提取更多的图像特征, 借鉴牛津大学 VGG 小组提出的基于卷积神经网络的视觉识别算法, 该算法采用 VGG-ILSVRC-16-layers 模型, 网络中所有卷积层均选取大小为  $3 \times 3$ 、数量为 32 的滤波器, 通道数量  $c$  仍然为 1, 卷积层表示为  $\text{Conv}(32, 3, 1)$ , 即

$$\begin{cases} F_1(Y) = f(W_1 * Y^{LR} + b_1) \\ F_2(Y) = f[W_2 * F_1(Y) + b_2] \\ \vdots \\ F_{n-1}(Y) = f[W_{n-1} * F_{n-2}(Y) + b_{n-1}] \end{cases}, \quad (8)$$

$$Y^{SR} = F_n(Y) = P[W_n * F_{n-1}(Y) + b_n], \quad (9)$$

式中  $W_1 = 32 \times 3 \times 3 \times 1, W_i = 32 \times 3 \times 3 \times 32 (i = 2, 3, \dots, n-1), W_n = r^2 \times 3 \times 3 \times 32$ 。

$P(x)$  的作用是将大小为  $r^2 \times H \times W \times C$  ( $H, W$  分别为图像的高度与宽度,  $C$  代表通道数) 的图像重新排列为大小为  $rH \times rW \times C$  的高分辨率图像  $Y^{SR}$ , 即将低分辨率图像特征按照特定位置周期性地插入到高分率图像中。通过使用子像素卷积层对图像进行重新排列, 在图像从低分辨率到高分率的放大过程中, 插值函数被隐含地包含在前面的卷积层中, 实现自动学习。只在最后一层对图像大小做变换, 由于前面的卷积运算在低分辨率图像上进行, 因此计算效率得到提高。

#### 3.2 残差网络模型

在计算机视觉里, 特征的等级随着网络深度的加深而变高, 增加网络深度可以有效地提取更多的特征, 然而梯度弥散/爆炸成为训练深层次网络的障碍, 导致网络难以收敛。为了加快网络收敛, 使用多

路径模式的局部残差学习与多权重的递归学习相结

合的模式,其结构图如图 4 所示。

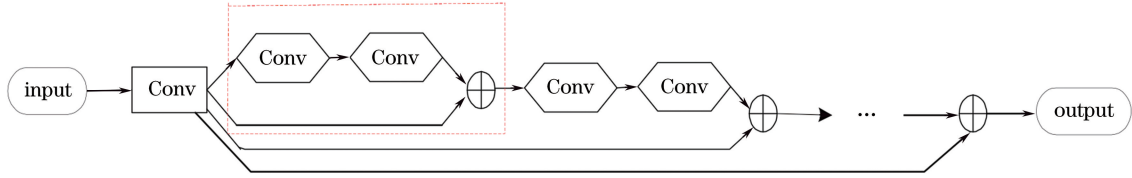


图 4 残差网络结构图

Fig. 4 Residual network structure

局部残差网络单元如图 4 中虚线框所示,如果设第一个残差单元的输入为  $x_l$ ,输出为  $x_{l+1}$ ,则残差单元的结构可以表示为

$$x_{l+1} = x_l + F(x_l, w_l), \quad (10)$$

$$F(x_l, w_l) = w_l \sigma(w_{l-1} x_{l-1}), \quad (11)$$

式中  $F(x)$  为卷积函数,  $\sigma(\cdot)$  为激活函数,  $w_l, w_{l-1}$  为权重。任意单元  $x_L$  可表示为

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, w_i). \quad (12)$$

对于标准普通网络,  $x_L$  可表示为

$$x_L = \prod_{i=l}^{L-1} w_i x_i. \quad (13)$$

由于求和的计算量远小于求积的计算量,因此残差网络可以加快收敛速度。

假设损失函数为  $C$ , 根据反向传播链式求导法则可以得到:

$$\frac{\partial C}{\partial x_l} = \frac{\partial C}{\partial x_L} \frac{\partial x_L}{\partial x_l} = \frac{\partial C}{\partial x_L} \left[ 1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, w_i) \right]. \quad (14)$$

(14) 式中  $\partial C / \partial x_L$  保证信息能直接传回任意层  $x_l$ ,  $1 + \frac{\partial}{\partial x_l} \sum_{i=l}^{L-1} F(x_i, w_i)$  保证网络不会出现梯度消失的现象。

局部残差网络结构是将局部残差单元顺序堆叠,不同的残差单元使用不同的输入,在此基础上使用了多路径结构,所有的残差单元共享相同的输入。与链式模式相比,这种多路径模式更利于学习,并且不容易过度拟合<sup>[19]</sup>。残差单元可表示为

$$x_{l+1} = g(x_l) = F(x_l, w) + x_0, \quad (15)$$

$$x_0 = f_0(x), \quad (16)$$

式中  $g(x_l)$  为残差单元的函数,  $x_0$  为递归块中第一层卷积的结果,  $f_0(x)$  为激活函数。由于残差单元是递归学习的,权重集合  $w$  在递归块内的残差单元之间共享。之后,通过叠加几个递归块得到最终的输出,再通过子像素卷积层得到高分辨率图像。

### 3.3 参数优化

深层卷积神经网络是一个端对端的映射,整个网络在训练的过程中,会通过反向传播算法优化网络,使其得到最优的参数。本文算法依旧使用均方误差作为损失函数,使用 Adam 代替随机梯度下降法来优化参数,使损失最小化。Adam 利用梯度的一阶矩估计和二阶矩估计动态调整每个参数的学习率,其优点在于经过偏置校正后,每一次迭代学习率都有一个确定范围,使得参数比较平稳。梯度的一阶矩估计  $m_t$  和二阶矩估计  $n_t$  可分别表示为

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t, \quad (17)$$

$$n_t = \beta_2 \cdot n_{t-1} + (1 - \beta_2) \cdot g_t^2, \quad (18)$$

式中  $\beta_1 = 0.9, \beta_2 = 0.999$ , 下标  $t$  表示当前时刻,下标  $t-1$  表示前一时刻。  $\hat{m}_t, \hat{n}_t$  是对  $m_t, n_t$  的校正,

$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \hat{n}_t = \frac{n_t}{1 - \beta_2^t}; -\frac{\eta}{\sqrt{\hat{n}_t} + \epsilon}$  对学习率形成一个动态约束,  $\beta_1$  和  $\beta_2$  为矩估计指数衰减速率;  $\theta$  为模型参数;  $\eta$  为步长。  $\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{n}_t} + \epsilon} \hat{m}_t, \epsilon = 10^{-8}$ 。

## 4 实验结果与分析

本文算法使用的系统为 Ubuntu 16.04, 在该系统下搭建了 caffe<sup>[20]</sup> 框架进行训练,在 MatlabR2014 下进行实验测试。电脑硬件配置为 Inter Core i5-4460@3.2 GHz, Nvidia GeForce GTX 750Ti。通过实验验证本文算法的优势,首先验证了深层网络的较浅层网络有更好的性能,然后通过对激活函数、运行速度、优化方法及滤波器数目的比较得到了最优的设置,最后通过实验表明局部残差网络可以使深层网络更快、更好地收敛,充分证明了本文算法无论在速度还是精度上都有明显优势。

### 4.1 数据集

在 SRCNN、ESPCN 等方法中,选择基于学习的超分辨率算法中广泛使用的 91 幅数据集作为训练数据。由于加深了网络深度,为了避免过拟合,需要扩大图像训练集,将这 91 幅图像先分别旋转  $90^\circ$ 、

180°和 270°之后,再将图像缩放为原来的 50%、60%、70%、80%和 90%,此时得到的训练数据是之前的 24 倍。将其中 1820(91×4×5)幅图像作为训练集,将 364(91×4×10)幅图像作为训练过程中的验证集,将 set 5、set 14、BSD100 作为测试集。将 1820 幅原始图像进行  $r$  倍下采样得到  $Y^{LR}$ ,然后按照步幅  $s=14$  将图像裁剪成大小为  $25 \times 25$  的子像素块,并将其与相应的裁剪后的高分辨率像素块组成训练数据,学习率设置为 0.0001,为了抑制振荡,动量设置为 0.9,优化方法选择 Adam,在确定范围内动态调整学习率。

### 4.2 层数设置

文献[21-22]表明,网络的深度是实现高分辨率图像的重要因素,增加网络深度可以提取更多的特征,然而深层网络会出现难以收敛和过拟合的情况,使网络训练的准确率下降<sup>[23]</sup>。本文算法增加了数据集并且采用局部残差网络结构,可以加速网络收敛以及有效地避免过拟合现象。当放大倍数为 3

时,分别训练和测试了层数为 6,10,12 的网络模型,即分别有 2,4,5 个残差单元。为了确保网络完全收敛,共迭代了  $10^6$  次。图 5(a)为 10 层和 12 层网络的训练集与验证集随迭代次数变化的关系曲线,从图中可以看出,最终 train loss 和 validation loss 都趋于平稳,近似于平行,说明网络已经完全收敛,没有出现过拟合,并且最终的 12 层网络的损失函数值更低,说明 12 层网络的效果更好。为进一步验证模型的有效性,选取 set 5 作为测试集对网络进行测试,并使用峰值信噪比 (PSNR) 与结构相似性指数 (SSIM) 的平均值作为客观评价指标,图 5(b)为测试集 set 5 随迭代次数的增长曲线,从图中可以看出,12 层网络的 PSNR 平均值虽然是最大的,但是相比于 10 层网络,二者差距很小,又对比了两者的 SSIM 平均值,最终结果说明了 12 层网络的 PSNR 与 SSIM 平均值均为最佳,因此最终选择了 12 层的网络模型。为了选取合适的层数,在表 1 中对比了不同层数的 SSIM 平均值。

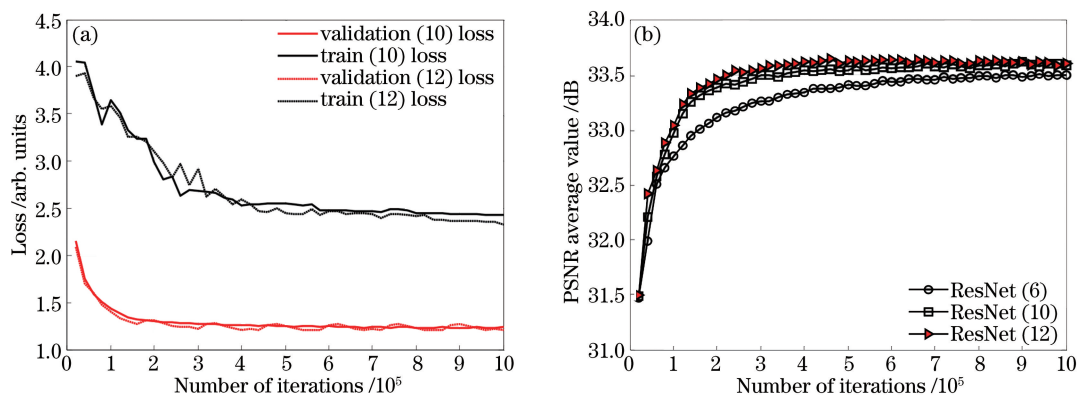


图 5 (a) 12 层网络的损失函数随迭代次数的变化;(b) set 5 的 PSNR 平均值在不同层数下随迭代次数的变化

Fig. 5 (a) Variation of loss function of 12-layer network with number of iterations; (b) variation of PSNR average value of set 5 with number of iterations under different layers

表 1 不同深度下各测试集的 PSNR/SSIM 均值

Table 1 PSNR/SSIM average values of test sets at different depths

Depth	PSNR average value		SSIM average value	
	set 5	set 14	set 5	set 14
6	33.48	29.65	0.9249	0.8936
10	33.58	29.67	0.9265	0.8941
12	33.60	29.69	0.9268	0.8944

### 4.3 激活函数比较

在网络训练过程中,激活函数的选择会影响模型的性能与速度。Sigmoid 函数的计算量较大,收敛缓慢且容易发生梯度消失的情况,已经逐渐不再使用;ReLU 的表达式为  $f(x) = \max(0, x)$ ,其计算复杂度低,但其在训练时很脆弱,极有可能会

导致神经元坏死;tanh 函数的表达式为  $f(x) = (1 - e^{-2x}) / (1 + e^{-2x})$ ,虽然其计算复杂度较高,但是其输出均值为 0,很好地解决了收敛速度慢与梯度容易消失的问题。如图 6 所示,网络层数为 12 时,在测试集 set 5 下将 tanh 与 ReLU 进行对比,可以看出 tanh 表现出的性能优于 ReLU,并且比 ReLU 更快地达到收敛。因此,最终选用 tanh 作为激活函数。

### 4.4 速度比较

网络结构的加深使模型复杂度增加,导致速度减慢。残差网络结构使得收敛速度加快,而且直接在低分辨率图像上进行卷积操作,减小了计算复杂度,使得计算速度有所提升。图 7 所示为在 set 5 测试集下,不同算法处理图像所用的平均时间,可以看

出,虽然本文算法使用了较深的网络模型,但是处理图像的速度得到提升,并且 PSNR 平均值明显增大,因此最终选取深度为 12 的网络模型。

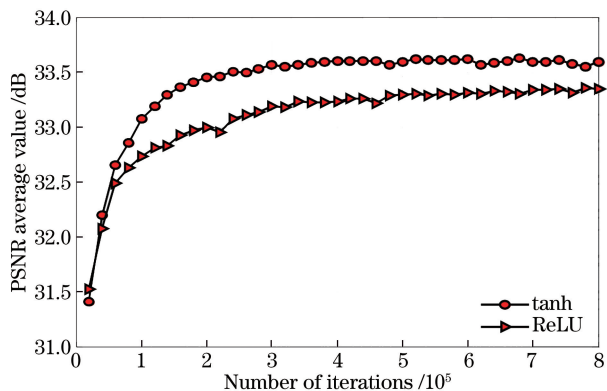


图 6 set 5 的 PSNR 平均值在不同激活函数下随迭代次数的变化

Fig. 6 Variation of PSNR average value of set 5 under different activation functions with number of iterations

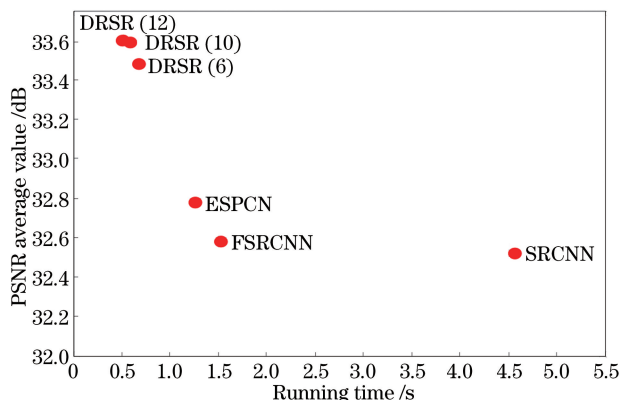


图 7 测试集 set 5 在不同算法下的运行时间与 PSNR 均值的关系图

Fig. 7 Relationship between running time and PSNR average value of set 5 under different algorithms

#### 4.5 优化方法比较

在 12 层(包含 5 个残差单元的深度)分别训练了优化方法为随机梯度下降法(SGD)、加速梯度下降法(NAG)和 Adam 的网络模型。从图 8 可以看出,Adam 优化方法和 NAG 明显优于 SGD,并且 Adam 优化方法也优于 NAG。因此,选择 Adam 优化方法。

#### 4.6 滤波器数目比较

为了评估不同滤波器数目对超分辨率性能的影响,在网络层数为 12 时分别训练了滤波器数目为 16,32,64 的网络模型,选择放大倍数为 3,并在 set 5 上进行测试,结果如图 9 所示。可以看出,增加滤波器数目虽然可以提高滤波性能,但是由于卷积层数较多,增大滤波器数目使得模型复杂度和计

算量大大增加,导致发生了过拟合现象,并且降低了计算速度,综合考虑超分辨率的性能与计算速度,选用滤波器数目为 32 的网络模型,以达到更好的效果。

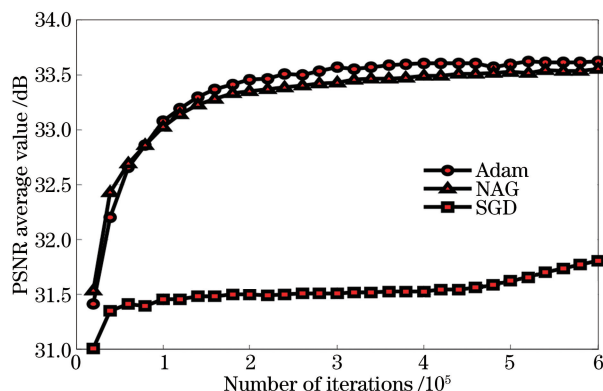


图 8 set 5 的 PSNR 平均值在不同优化方法下随迭代次数的变化

Fig. 8 Variation of PSNR average value of set 5 under different optimization methods with number of iterations

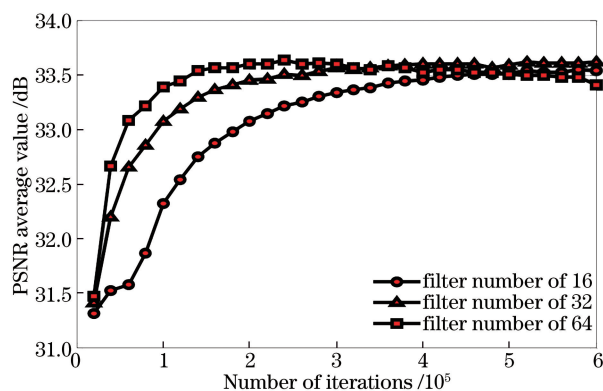


图 9 set 5 的 PSNR 平均值在不同滤波器数目下随迭代次数的变化

Fig. 9 Variation of PSNR average value of set 5 under different filter numbers with number of iterations

#### 4.7 残差网络设置

残差网络特殊的网络结构可以加快深度网络模型的收敛,为了验证其特性,分别将残差网络与标准网络进行实验比较,选取放大倍数为 3,测试集为 set 5,网络深度分别为 6,8,10,12(分别有 2,3,4,5 个残差单元),激活函数为 tanh 以及优化方法为 Adam 进行验证,其收敛曲线如图 10 所示。可以看出,残差网络的收敛速度要明显高于普通网络,而且 10 层和 12 层网络的收敛速度比 6 层和 8 层的更快,PSNR 平均值更高。由此可知,增加残差单元的个数可以使深层网络更快地收敛,更好地提升超分辨率重建的性能与速度。

#### 4.8 实验结果

将本文算法分别与 HE、Bicubic、ScSR<sup>[11]</sup>

(Image Super Resolution via Sparse Representation)、NE + LLE<sup>[10]</sup> (Super-Resolution Through Neighbor Embedding)、ANR<sup>[12]</sup>、SRCNN<sup>[14]</sup> 和 ESPCN<sup>[16]</sup> 进行比较。由于受实验环境以及迭代次数的影响,实验结果与 SRCNN<sup>[14]</sup>、

ESPCN<sup>[16]</sup> 等原文献略有不同,所有算法均在放大倍数为 2,3,4 的比例下进行实验,图 11、12 为放大倍数为 3 的情况下不同方法重建的效果图,为了进一步观察,对图像某些部分进行了局部放大,可以看出,双三次插值的效果最为模糊。图像经 ScSR

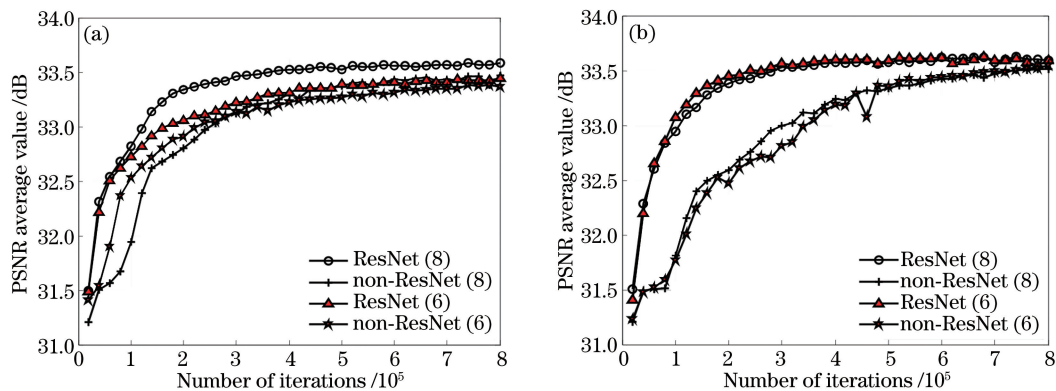


图 10 set 5 的 PSNR 平均值在不同网络模型下随迭代次数的变化。(a) 6 层和 8 层网络;(b) 10 层和 12 层网络

Fig. 10 Variation of PSNR average value of set 5 under different network models with number of iterations.

(a) Networks of 6-layer and 8-layer; (b) networks of 10-layer and 12-layer

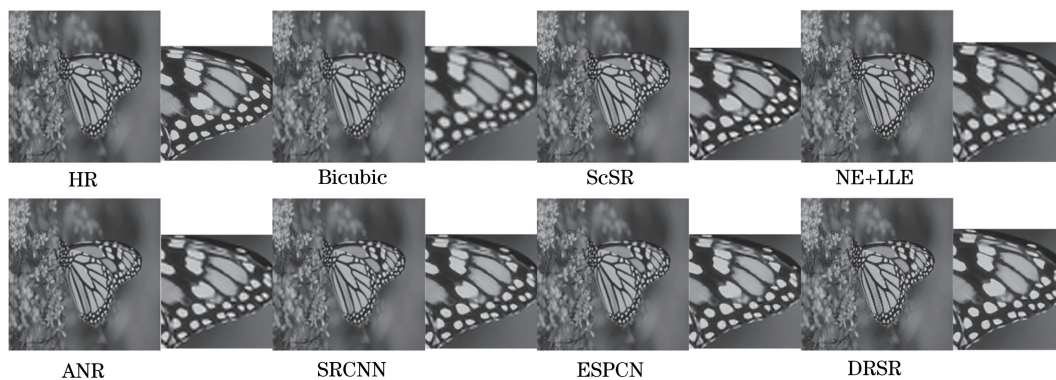


图 11 Monarch 在不同算法下的效果图

Fig. 11 Effect of Monarch under different algorithms

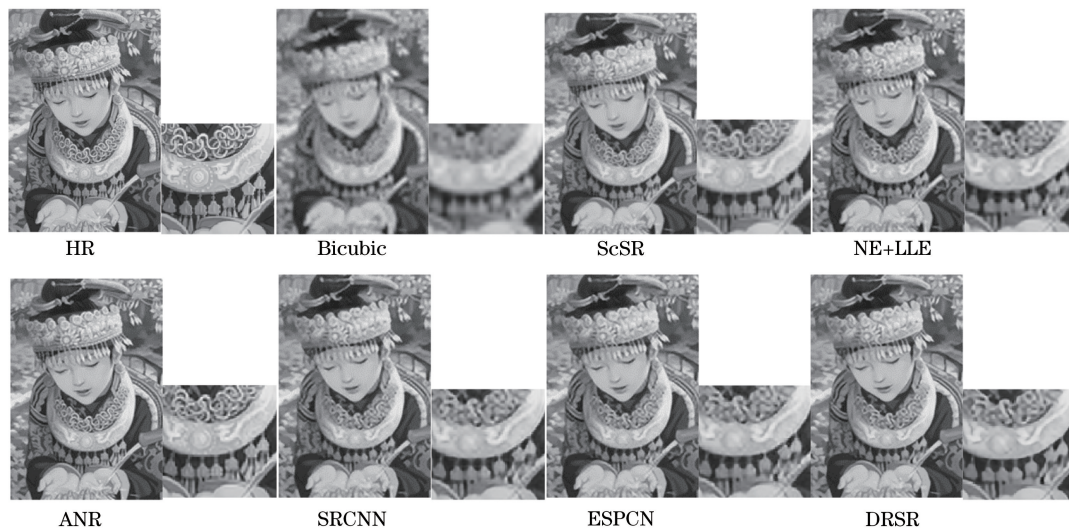


图 12 Comic 在不同算法下的效果图

Fig. 12 Effect of Comic under different algorithms

与 NE+LLE 处理后,虽然清晰度有所提升,但是图像细节效果仍有待提高,ANR, SRCNN, ESPCN 算法的整体视觉效果较清晰,边缘处也有很好的连续性,但易产生伪影现象。对比以上算法,本文算法可以恢复更多图像的细节,消除伪影现象,得到的图像边缘也更加锐利,视觉效果更好。

客观评价标准使用 PSNR 和 SSIM<sup>[24]</sup> 两个指标对算法进行评估。PSNR 用来衡量两幅图像之间的强度差,其计算式为

$$P_{\text{SNR}} = 10 \lg \frac{255^2}{M_{\text{SE}}}, \quad (19)$$

式中  $M_{\text{SE}}$  表示均方误差。 $P_{\text{SNR}}$  越大,图像的重建效果越好。

SSIM 能更好地表达恢复图像与参考图像之间的结构相似性,结构相似度的值越接近 1,说明原始标准图像块  $x$  和重建图像块  $\hat{x}$  之间越相似、重建图

像质量效果越好。SSIM 计算公式为

$$S_{\text{SIM}}(x, \hat{x}) = \frac{(2u_x u_{\hat{x}} + C_1)(2\sigma_{xx} + C_2)}{(u_x^2 + u_{\hat{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2)}, \quad (20)$$

式中  $u_x, u_{\hat{x}}$  分别表示图像  $x$  和  $\hat{x}$  的均值,  $\sigma_x, \sigma_{\hat{x}}$  分别表示图像  $x$  和  $\hat{x}$  的方差,  $\sigma_{xx}$  表示图像  $x$  和  $\hat{x}$  的协方差,  $C_1, C_2$  为常数。

不同测试集分别在放大倍数为 2, 3, 4 时的 PSNR 平均值和 SSIM 平均值如表 2 和表 3 所示。可以看出,本文算法的 PSNR 平均值与 SSIM 平均值较传统的双三次插值算法有明显提高,相比其他算法也均有所提高,在放大倍数为 2 时, set 5 测试集的 PSNR 平均值可以达到 37.41 dB, SSIM 平均值达到 0.9621, 相比于 Bicubic, PSNR 平均值提高了约 4 dB, SSIM 提高了 0.03。

表 2 测试集 set 5, set 14, BSD100 在不同算法下的 PSNR 平均值

Table 2 PSNR average value of set 5, set 14, and BSD100 under different algorithms

Data set	Scale	Bicubic	ScSR	NE+LLE	ANR	SRCNN	ESPCN	DRSR
set 5	×2	33.65	35.13	35.76	35.83	36.36	36.39	37.41
	×3	30.42	31.54	31.91	32.00	32.52	32.78	33.60
	×4	28.44	28.24	29.66	29.74	30.15	30.21	31.18
set 14	×2	30.21	31.36	31.78	31.81	32.21	32.21	32.95
	×3	27.51	28.36	28.59	28.64	29.03	29.13	29.69
	×4	25.97	25.97	26.78	26.83	27.23	27.17	27.83
BSD100	×2	29.41	29.41	30.39	30.43	30.89	30.93	31.55
	×3	27.07	27.67	27.78	27.81	28.11	28.27	28.54
	×4	25.84	25.84	26.39	26.41	26.63	26.59	27.00

表 3 测试集 set 5, set 14, BSD100 在不同算法下的 SSIM 平均值

Table 3 SSIM average value of set 5, set 14, and BSD100 under different algorithms

Data set	Scale	Bicubic	ScSR	NE+LLE	ANR	SRCNN	ESPCN	DRSR
set 5	×2	0.9355	0.9428	0.9537	0.9546	0.9566	0.9568	0.9621
	×3	0.8779	0.8851	0.9053	0.9062	0.9129	0.9162	0.9268
	×4	0.8185	0.8025	0.8516	0.8533	0.8621	0.8578	0.8846
set 14	×2	0.9348	0.9564	0.9575	0.9578	0.9591	0.9598	0.9630
	×3	0.8494	0.8649	0.8802	0.8777	0.8846	0.8874	0.8941
	×4	0.7799	0.7800	0.8152	0.8173	0.8206	0.8225	0.8361
BSD100	×2	0.8401	0.8702	0.8729	0.8742	0.8827	0.8832	0.8924
	×3	0.7301	0.7514	0.7702	0.7718	0.7778	0.7816	0.7917
	×4	0.6470	0.6470	0.6875	0.6896	0.6919	0.6942	0.7092

## 5 结 论

针对现有算法网络结构简单、收敛速度慢以及图像纹理模糊等问题,提出了一种深层残差卷积神经网络模型,使用较小的卷积核直接在低分辨率图

像上提取特征,在网络的末端通过子像素卷积层将像素进行重新排列,从而得到高分辨率图像,使用多路径模式的局部残差结构可加快模型收敛并提高训练速度。为避免卷积后图像信息丢失,在卷积过程中使用零填充以保证输出图像与输入图像尺寸相



等,实验中使用 Adam 优化方法。实验结果表明,本文算法在主观视觉效果和客观评价指标上均优于现有的其他几种算法,无论在速度还是精度上都有所提升。接下来的工作可以在保证速度的同时进一步优化网络模型,以提高图像超分辨率的精度,并且可以将算法应用于其他领域,如医学图像的超分辨率重建中。

## 参 考 文 献

- [1] Zhang H Y, Zhang L P, Shen H F. A blind super-resolution reconstruction method considering image registration errors[J]. *International Journal of Fuzzy Systems*, 2015, 17(2): 353-364.
- [2] Chen H G, He X H, Teng Q Z, *et al.* Single image super resolution using local smoothness and nonlocal self-similarity priors[J]. *Signal Processing: Image Communication*, 2016, 43: 68-81.
- [3] Li S M, Lei G Q, Fan R. Depth map super-resolution reconstruction based on convolutional neural networks[J]. *Acta Optica Sinica*, 2017, 37(12): 1210002.  
李素梅, 雷国庆, 范如. 基于卷积神经网络的深度图超分辨率重建[J]. *光学学报*, 2017, 37(12): 1210002.
- [4] Batz M, Eichenseer A, Seiler J, *et al.* Hybrid super-resolution combining example-based single-image and interpolation-based multi-image reconstruction approaches [C]. *IEEE International Conference on Image Processing (ICIP)*, 2015: 58-62.
- [5] Kim K I, Kwon Y. Single-image super-resolution using sparse regression and natural image prior[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(6): 1127-1133.
- [6] Lian Q S, Zhang W. Image super-resolution algorithms based on sparse representation of classified image patches[J]. *Acta Electronica Sinica*, 2012, 40(5): 920-925.  
练秋生, 张伟. 基于图像块分类稀疏表示的超分辨率重构算法[J]. *电子学报*, 2012, 40(5): 920-925.
- [7] Xiao J S, Liu E Y, Zhu L, *et al.* Improved image super-resolution algorithm based on convolutional neural network[J]. *Acta Optica Sinica*, 2017, 37(3): 0318011.  
肖进胜, 刘恩雨, 朱力, 等. 改进的基于卷积神经网络的图像超分辨率算法[J]. *光学学报*, 2017, 37(3): 0318011.
- [8] Irani M, Peleg S. Improving resolution by image registration [J]. *CVGIP: Graphical Models and Image Processing*, 1991, 53(3): 231-239.
- [9] Stark H, Oskoui P. High-resolution image recovery from image-plane arrays, using convex projections [J]. *Journal of the Optical Society of America A*, 1989, 6(11): 1715-1726.
- [10] Chang H, Yeung D Y, Xiong Y M. Super-resolution through neighbor embedding[C]. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004: 275-282.
- [11] Yang J C, Wright J, Huang T S, *et al.* Image super-resolution via sparse representation[J]. *IEEE Transactions on Image Processing*, 2010, 19(11): 2861-2873.
- [12] Timofte R, de Smet V, van Gool L. Anchored neighborhood regression for fast example-based super-resolution[C]. *IEEE International Conference on Computer Vision*, 2013: 1920-1927.
- [13] Timofte R, de Smet V, van Gool L. A+: adjusted anchored neighborhood regression for fast super-resolution[C]. *Asian Conference on Computer Vision*, 2015: 111-126.
- [14] Dong C, Loy C C, He K M, *et al.* Image super-resolution using deep convolutional networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.
- [15] Dong C, Chen C L, Tang X. Accelerating the super-resolution convolutional neural network[C]. *European Conference on Computer Vision*, 2016: 391-407.
- [16] Shi W Z, Caballero J, Huszár F, *et al.* Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1874-1883.
- [17] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1646-1654.
- [18] Sun C, Lü J W, Li J W, *et al.* Fast image super-resolution method based on deconvolution[J]. *Acta Optica Sinica*, 2017, 37(12): 1210004.  
孙超, 吕俊伟, 李健伟, 等. 基于去卷积的快速图像超分辨率方法[J]. *光学学报*, 2017, 37(12): 1210004.
- [19] Gong C, Tao D C, Liu W, *et al.* Label propagation via teaching-to-learn and learning-to-teach[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2017, 28(6): 1452-1465.
- [20] Jia Y Q, Shelhamer E, Donahue J, *et al.* Caffe: convolutional architecture for fast feature embedding [C]//*Proceedings of the 22nd ACM International Conference on Multimedia-(MM'14)*, 2014: 675-678.

- [21] Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2015: 5197-5206.
- [22] Du B, Wang Z, Zhang L, *et al.* Robust and discriminative labeling for multi-label active learning based on maximum correntropy criterion[J]. IEEE Transactions on Image Processing, 2017, 26(4): 1694-1707.
- [23] Du B, Xiong W, Wu J, *et al.* Stacked convolutional denoising auto-encoders for feature representation[J]. IEEE Transactions on Cybernetics, 2017, 47(4): 1017-1027.
- [24] Ye Y X, Shan J, Bruzzone L, *et al.* Robust registration of multimodal remote sensing images based on structural similarity[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(5): 2941-2958.