

基于深度跳跃级联的图像超分辨率重建

袁昆鹏*, 席志红

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 针对模型 VDSR(very deep super resolution)收敛速度慢,训练前需要对原始图像进行预处理,以及网络中存在的冗余性等问题,提出了一种基于深度跳跃级联的单幅图像超分辨率重建(DCSR)算法。DCSR 算法省去了图像预处理,直接在低分辨率图像上提取浅层特征,并使用亚像素卷积对图像进行放大;通过使用跳跃级联块可以充分利用每个卷积层提取到图像特征,实现特征重用,减少网络的冗余性。网络的跳跃级联块可以直接从输出到每一层建立短连接,加快网络的收敛速度,缓解梯度消失问题。实验结果表明,在几种公开数据集上,所提算法的峰值信噪比、结构相似度值均高于现有的几种算法,充分证明了所提算法的出色性能。

关键词 机器视觉; 超分辨率; 深度学习; 跳跃级联; 梯度消失; 特征复用; 亚像素卷积; 冗余性

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201939.0715003

Image Super Resolution Based on Depth Jumping Cascade

Yuan Kunpeng*, Xi Zhihong

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

Abstract The very deep super resolution model has disadvantages: the convergence speed is low, the original image must be preprocessed before training, and the network redundancy must be reduced. This study proposes a single-image super resolution reconstruction method based on depth jumping cascade (DCSR). First, DCSR eliminates pre-processing, extracts the shallow features directly on the low-resolution image, and finally uses sub-pixel convolution to magnify the image. Second, each convolutional layer is fully utilized to extract the image features using the jump cascading block, thereby realizing feature reuse and network redundancy reduction. The jump cascading block of the network establishes a short connection directly from the output to each layer, speeding up the network convergence speed and alleviating the gradient disappearance problem. The experimental results show that on several public datasets, the peak-signal-to-noise ratio and the structural similarity of the algorithm are higher than those of existing algorithms, which fully demonstrates an excellent algorithm performance.

Key words machine vision; super resolution; deep learning; jumping cascade; gradient disappear; feature reuse; sub-pixel convolution; redundancy

OCIS codes 150.0155; 100.6640; 100.3010; 100.2000

1 引 言

图像的超分辨率重建技术就是应用软件通过一幅或多幅低分辨率(LR)图像重建出一幅高分辨率(HR)图像^[1-3],本文主要研究单幅图像超分辨率重建(SISR)。图像的超分辨率重建技术已经广泛应用于医学影像、军事侦查和视频监控等领域。目前,超分辨率重建技术可分为三类,分别是基于插值^[4]的方法、基于重建^[5]的方法和基于学习^[6-7]的方法。基于插值的方法有双线性插值、双立方插值和最近

邻插值等,这类方法虽然速度快,但是重建出的图像边缘会出现锯齿,导致图像不清晰。基于重建的方法有迭代反向投影法^[8]和凸集投影法等^[9],这类方法随着放大倍数的增加,不能很好地重建一些高频细节。基于学习的方法主要有邻域嵌入法^[10]、稀疏表示法^[11]和卷积神经网络(CNN)等,这类方法可构建低分辨率图像与高分辨率图像的映射关系,从而实现图像的重建。近几年,随着人工智能的迅速发展,基于CNN的图像超分辨率重建技术以其速度快、端对端学习效果好的特点,正在逐步成为主流方

收稿日期: 2019-01-08; 修回日期: 2019-03-06; 录用日期: 2019-03-22

基金项目: 国家自然科学基金(60875025)

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

法。最早将 CNN 用于图像超分辨率重建的是超分辨率卷积神经网络 (SRCNN)^[12], 该网络模型很简单, 仅使用三层卷积来模拟非线性映射, 从而实现图像的重建。Dong 等^[12-13]对 SRCNN 进行了改进并提出了快速超分辨率卷积神经网络 (FSRCNN), FSRCNN 使用更小的卷积核和更多的卷积层, 并通过反卷积层实现图像的放大。由于 FSRCNN 不需要将原始图像经过双三次差值放大到目标尺寸大小, 且使用了更小的卷积核, 所以在训练速度上有较大的提升。Shi 等^[14]提出了亚像素卷积神经网络 (ESPCN), ESPCN 同样不需要将原始图像经过双三次差值来进行图像放大, 直接在低分辨率图像上提取特征, 使用了 tanh 激活函数代替了 ReLU, 最后通过亚像素卷积层将像素进行重新排列以实现图像的重建。之后, Kim 等^[15]提出了 VDSR (very deep super resolution) 模型, 加深了网络深度, 使用残差网络结构 (ResNet^[16]), 加速了收敛过程, 同时避免了网络加深而引起的梯度消失问题, 提升了图像重建的质量。

上述几种方法仍存在一些缺点, 对于 SRCNN^[12], 其收敛速度慢, 卷积层数少, 无法充分获取更多的图像重建信息; 对于 VDSR^[15], 虽然其网络层数加深到 20 层, 但需要对低分辨率图像进行双三次差值, 因此增加了计算量, 且插值会导致一些图像的细节信息丢失。文献 [17] 表明, 在 ResNet^[16] 网络结构中, 训练时随机丢弃一些卷积层并不影响 ResNet^[16] 的测试性能, 说明神经网络中每一个卷积层学习到的图像特征很少, 导致网络中有大量的冗余信息。为了解决网络的冗余, Huang 等^[18]提出了 DenseNet, DenseNet 打破了通过加深网络层数和加宽网络结构来提高性能的惯性思维,

从特征图角度出发, 正常 N 层网络只有 N 层连接, 而 DenseNet 的每一层与之前所有层都相连, 所以 N 层网络有 $N(N+1)/2$ 层连接, 这种网络的优点是可以缓解梯度消失的问题, 加强特征的传播, 有效地减少网络冗余。基于这一结论可知, 之前的这几种方法都忽视了充分利用各卷积层提取到的图像特征, 采用递进的层级网络, 每一个卷积层只能获取到紧邻的上一个卷积层输出的图像特征, 仅采用来自 LR 空间的最后一个卷积层的特征映射进行图像的放大。VDSR^[15] 采用一个残差网络, 使用这种跳跃的短连接使得重建的卷积层不仅可以获得紧邻卷积层输出的图像特征, 还可以获得插值后低分辨率的图像特征。本文利用跳跃级联块 (JCD), 后面卷积层充分利用了前面卷积层的学习特征, 实现了特征重用, 减少了网络冗余, 从而更好地重建图像。

本文提出了基于深度跳跃级联的图像超分辨率重建方法 (DCSR), 该算法的优点包括:

- 1) 不需要对低分辨率图像进行双三次插值放大, 在重建网络中使用高效亚像素卷积进行图像的放大;
- 2) 使用跳跃级联块, 实现特征重用, 减少网络冗余, 加速参数更新的速度, 提高网络的收敛速度;
- 3) 通过建立的短连接, 缓解了由网络层数过深引起的梯度消失问题。

2 相关工作

2.1 SRCNN

SRCNN^[12] 算法需要使用双三次插值将低分辨率图像放大到所需的图像大小, 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, 其网络结构如图 1 所示。

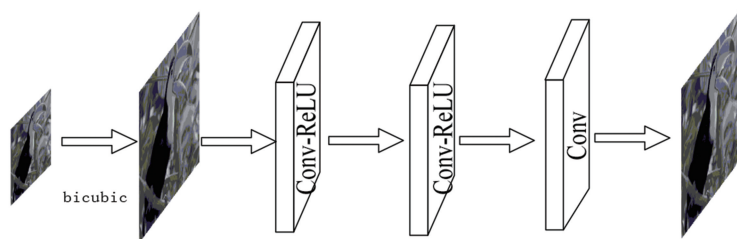


图 1 SRCNN 网络结构图

Fig. 1 Diagram of SRCNN network structure

在输入到网络之前, 需要利用双三次差值将低分辨率图像放大到目标图像尺寸, 首先经过第一层卷积 (9×9) 对图像块进行特征提取, 然后经过第二层卷积 (1×1) 将低分辨率的特征映射为高分辨率的

特征, 再经过最后一层卷积 (5×5) 对得到的高分辨率特征进行最终的重建。

2.2 VDSR

VDSR^[15] 借鉴了 ResNet^[16] 的思想, 使用残差

学习,防止了梯度消失,并加快了网络的收敛速度,

其网络结构如图 2 所示。

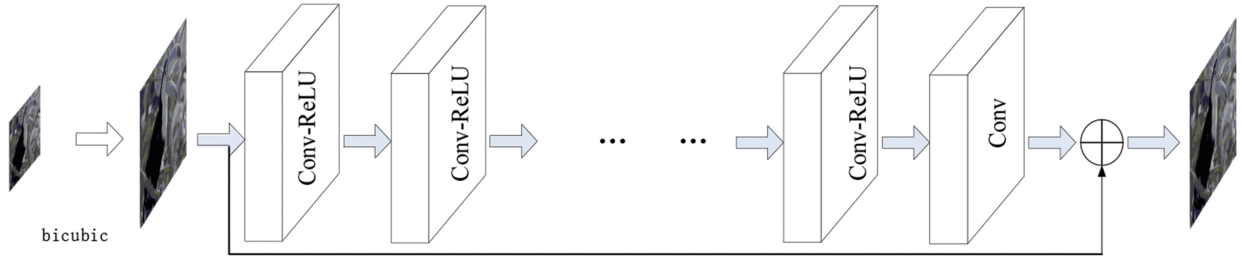


图 2 VDSR 网络结构图

Fig. 2 Diagram of VDSR network structure

网络层数加深到 20 层后,增大了感受野,由 SRCNN^[12]的 13×13 变为 41×41 ,每层卷积使用了更小的卷积核 3×3 ,为了充分利用图像边缘信息,在每次卷积前进行边界补零,使图像的大小保持不变。

3 本文算法

本文算法的网络结构如图 3 所示,分为低层特征提取网络(LFENet)、跳跃级联块网络(DCNet)和重建网络(ReconNet)三个部分。

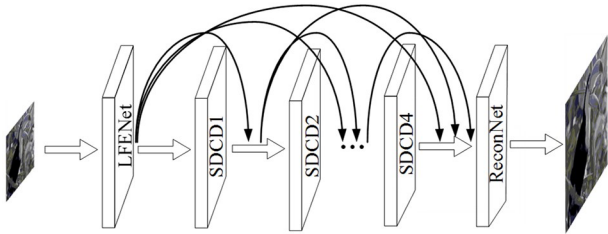


图 3 本文算法网络结构图

Fig. 3 Network structural diagram of proposed algorithm

3.1 低层特征提取网络

首先借鉴牛津大学 VGG 小组提出的基于卷积神经网络的视觉识别算法中所采用的 VGG-16^[19]模型,使用 3×3 卷积核进行低层特征提取。使用线性整流函数 PReLU 作为激活函数。为了充分利用图像的边缘信息,在每次卷积前进行补零操作,卷积层可表示为

$$F_0(Y) = \Phi(W_1 \times I_{LR} + b_1), \quad (1)$$

式中: Y 为低层特征提取网络的输出; I_{LR} 为输入的低分辨率图像; W_1 表示一组滤波器,其大小可表示为 $n_1 \times k_1 \times k_1 \times c_1$,其中 n_1 代表输出滤波器的数量,这里取 32, k_1 代表卷积核大小,这里 k_1 取 3, c_1 代表图像的通道数,图像正常有三个通道(Y, C_b, C_r),由于人眼对图像的亮度信息更敏感,所以只需要对 Y 通道进行处理, c_1 取值为 1,步长为 1,边界零填充也为 1; b_1 为偏移项; $\Phi(X)$ 表示激活函数,这里采

用带有参数的 PReLU 作为激活函数, $\Phi(X)$ 可表示为

$$\Phi(X) = \max(0, X) + \alpha_i \min(0, X), \quad (2)$$

其中

$$\Delta\alpha_{i+1} = \mu\Delta\alpha_i + \eta \frac{\partial \epsilon}{\partial \alpha_i}, \quad (3)$$

式中: X 为自变量; α_i 为可学习的参数,初始化为 0.25; $\Delta\alpha_i$ 为上一次更新的参数值,网络反向传播时,采用带动量的更新方式更新参数 α_i ; μ 和 η 分别为动量大小和学习率; ϵ 为损失函数。

3.2 跳跃级联块网络

在现有的 CNN 中,随着网络层数的加深,数据在传播过程中会出现梯度消失的问题,导致网络不能继续训练。对于 DenseNet^[18],靠近输入的层与靠近输出的层之间的连接越短,卷积神经网络可以设计得越深,精度更高,则可以更加有效地进行训练,这种短连接可以减轻梯度消失的问题。依照该结论,本文构造了跳跃级联块,由图 3 可以看出,在反向传播时,每一层的梯度都来自于后面的所有层,因此靠近输入层的梯度不会随着网络深度的增加而越来越小,从而使梯度更好地在网络中传输,降低了梯度消失的风险。虽然 VDSR^[15]通过建立一个跳跃的短连接结构在一定程度上促进了数据在卷积层间的流通,但还是未充分利用各层之间的图像特征。每一个跳跃级联块的连接方式如图 4 所示,每一层的输入都来自所有前面层的输出,从而实现特征的重复利用,减少网络冗余。与 VDSR^[15]不同的是,数据聚合方式不是对特征映射直接求和,而是采用连接的方式,将每个跳跃级联块采用相同的结构进行连接,进一步加强网络中信息的流通。这样能保证层与层之间信息最大程度的利用,加深了网络深度。为了防止过拟合,引入了 dropout,在训练的时候,按照一定的概率对每个神经元进行随机采样,相

当于给每个神经元加上了一个开关,因此只有一部分神经元在工作,减少了神经元之间的联系,并形成子网络,这样训练时参数就会大大减少,一个复杂的网络就变得简单。具体可表示为

$$\omega_{d,c} |_{\rho} = \omega_{d,c} \times \rho, \quad (4)$$

$$F_{d,c} = \sigma(\omega_{d,c} |_{\rho} [F_{d-1}, F_{d,1}, \dots, F_{d,c-1}] + b_{d,c}), \quad (5)$$

式中: $\omega_{d,c} |_{\rho}$ 表示权重是服从概率为 ρ 的伯努利分

布,设置 ρ 等于0.2; $\omega_{d,c}$ 和 $F_{d,c}$ 分别表示第 d 个JCD的第 c 层的一组滤波器,可表示为 $n_2 \times k_2 \times k_2 \times c_2$; n_2 代表滤波器的数量,这里取32; k_2 代表卷积核大小,这里取3; c_2 取值为1,步长为1,边界零填充为1; $b_{d,c}$ 表示偏移项; σ 表示激活函数PReLU; $F_{d,c}$ 表示第 d 个JCD的第 c 层的输出; $[F_{d-1}, F_{d,1}, \dots, F_{d,c-1}]$ 表示由第 $d-1$ 个JCD和第 d 个JCD的1,2,..., $c-1$ 层输出的特征映射的连接。

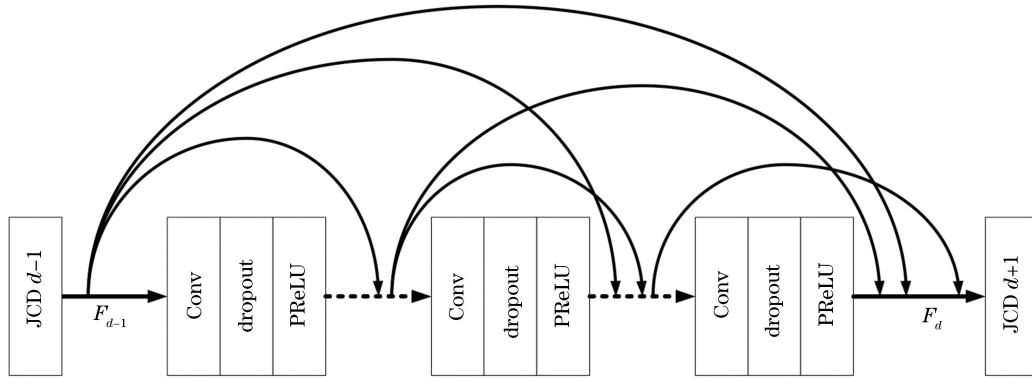


图4 跳跃级联块

Fig. 4 Jumping cascade block

3.3 重建网络

重建网络包括两个部分,首先使用一个 1×1 的卷积层对跳跃密集连接块网络进行一个特征映射的融合,然后将融合后的特征映射经过亚像素卷积层得到最终重建的高分辨率图像。经过 1×1 卷积层后特征融合的输出 F_{FF} 和最终重建的高分辨率图像 F_{HR} 可分别表示为

$$F_{FF} = H([F_1, F_2, \dots, F_D]), \quad (6)$$

$$F_{HR} = A(\omega_{HR} \times F_{FF} + b_{HR}), \quad (7)$$

式中: D 表示SDCD块个数; $[F_1, F_2, \dots, F_D]$ 表示 D 个SDCD块的输出特征映射的连接; H 表示 1×1 的卷积函数; $A(\cdot)$ 表示排列函数; ω_{HR} 表示一组大小为 $n_3 \times k_3 \times k_3 \times c_3$ 的滤波器, $n_3 = h^2$, h 表示放大倍数, k_3 代表卷积核大小,这里取3,步长为1,边界零填充为1; b_{HR} 表示偏移项。

图像经过亚像素卷积层时进行卷积和排列两个过程:1)经过卷积层,输出的特征图尺寸大小与输入图像一样,但是通道数变为 h^2 ,然后再将特征图像的 h^2 个像素按照特定的顺序排列成一个大小为 $h \times h$ 的区域,该区域对应高分辨率图像一个大小为 $h \times h$ 的子块,依照这种方式,就可以将一个大小为 $h^2 \times H \times W$ 的低分辨率特征图重新排列成一个大小为 $1 \times hH \times hW$ 的高分辨率图像,前面的卷积是在低分辨率图像上进行的,并且排列不需要卷积,

因此效率很高。

图5(a)~(c)分别是本文选取图像butterfly放大4倍时,经过低层特征提取网络、第一个跳跃级联网络和重建网络后的可视化结果,随着网络层数的加深,提取到的特征变复杂。

3.4 损失函数和评价标准

使用均方误差(MSE)作为损失函数 f_{MSE} ,MSE用来比较原始图像与重建高清图像之间的绝对误差,MSE的值越小,说明重建出的图像与原始图像越接近。 f_{MSE} 可表示为

$$f_{MSE} = \frac{1}{n} \sum_{i=1}^n \|F_{HR} - X_i\|^2, \quad (8)$$

式中: n 表示训练集的数目; X_i 表示原始清晰图像。使用最普遍的图像评价指标,即峰值信噪比(PSNR)和结构相似度(SSIM)^[20],对重建后的图像进行评价。PSNR的值越大,表明重建后的图像失真越小,效果越好。SSIM的值越接近于1,表明重建后的图像与原始图像越接近,重建图像质量越高。

4 实验结果与分析

4.1 实验环境

本实验基于ubuntu 16.04LTS操作系统,训练时使用深度学习框架caffe^[21],并搭建了cuda, cudnn

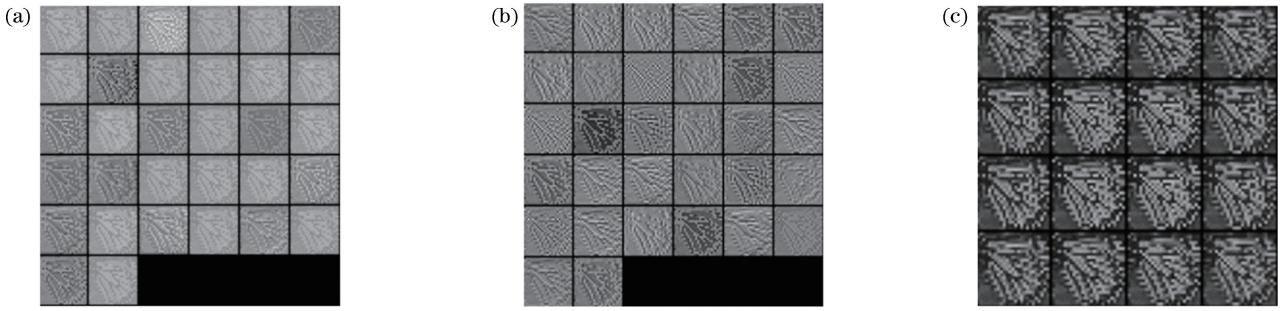


图 5 图像 butterfly 放大 4 倍后的可视化结果。(a)低层特征;(b)跳跃级联特征;(c)重建特征
Fig. 5 Visualization results of picture butterfly after 4-time magnification. (a) Low-level feature;
(b) hopping cascade feature; (c) reconstructing feature

用于实验的加速,测试使用 Matlab R2014a,电脑硬件配置 CPU I7 8700,显卡是 NVIDIA Geforce 1050TI。

4.2 数据集及优化方法

本实验采用的数据集是公开的 Timofte dataset^[22]和 Berkeley Segmentation Datasets^[23],将 Timofte dataset 中 91 张图片作为训练集,Set14 和 Set5 作为测试集,再将 Berkeley Segmentation Datasets 中的 BSD100 作为测试集。随着网络层数的加深,需要对训练集进行数据扩增,先将原始的 91 张训练集进行缩放,缩放到原来尺寸的 50%,60%,70%,80%,90%,再进行旋转,顺时针旋转 90°,180°,270°,扩增后的数据是原来的 24 倍;接着对图像进行预处理,先将训练集的图像进行 f 倍的下采样,得到对应的低分辨率的图像,再以步长 14 对图像进行裁剪,将其裁剪成 25×25 的子图像,将裁剪好的子图像作为网络的输入。

采用随机梯度下降法(SGD)结合动量作为优化方法,随机梯度下降法每次只选择一个样本来更新网络的参数,学习速度很快,但是其波动范围较大,会使收敛速度下降,所以加入了动量,可以有效地解决振荡,并加速收敛速度,该过程可表示为

$$W_{t+1} = W_t - V_{t+1}, \quad (9)$$

$$V_{t+1} = YV_t + \eta \nabla_w J(W_t), \quad (10)$$

式中: t 为迭代次数; W 表示权重; V 表示更新值; Y 表示动量,设置为 0.9; $J(W_t)$ 表示代价函数; $\nabla_w J(W_t)$ 表示代价函数的梯度; η 表示学习率,初始的学习率设置为 0.0001,每迭代 100 000 次,学习率降为原来的 1/10。

4.3 网络参数的设置选择

为了提升超分辨率重建后的效果,得到最优的网络参数设置,通过以下 5 个方面来展开实验,并画出测试集峰值信噪比的均值随迭代次数增加的

曲线。

4.3.1 层数设置

有研究表明,网络深度是影响网络性能的重要因素之一^[24]。适当增加网络层数可以提取更多的图像特征,便于后期的重建,为此进行了三组实验,实验以 $f=4$ 为例,采用 Set5 作为测试集,如图 6 所示,分别是跳跃密集连接块里面有 4(JCD4),6(JCD6),8(JCD8)个卷积层,网络对应深度为 21,27,35 层,初始阶段 35 层网络的 PSNR 值是最高的,随着迭代次数的增加,27 层网络与 35 层网络的差距越来越小,只保持了微小的优势,考虑到运算的复杂度和时间成本,最终选取 27 层的网络模型。

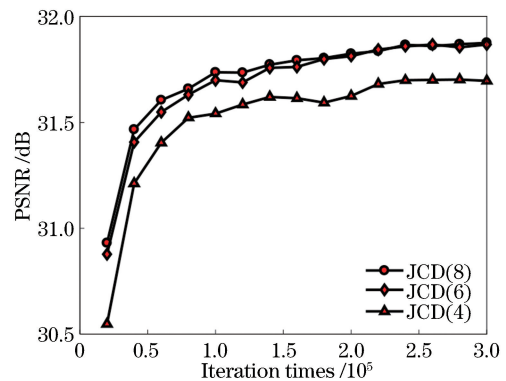


图 6 Set5 测试集下不同层数 PSNR 均值随迭代次数增长的曲线

Fig. 6 Relationship between average PSNR of different layers with number of iterations under Set5 test set

4.3.2 收敛速度

网络参数的更新速度越快,越有利于网络收敛。进行了 6 组实验,实验以 $f=4$ 为例,采用 Set5 作为测试集,如图 7 所示,分别采用了 15,21,27 层带跳跃密集连接块的网络和 15,21,27 层不带跳跃密集连接块的网络,由(9)、(10)式可得,权重的更新与损失函数的梯度有关,梯度越大,权重更新的速度越

快,越有利于网络的收敛。采用跳跃密集连接块的网络,在反向传播时,每一层的梯度都会接收到后面所有层的梯度,远大于同一层没有跳跃密集连接块的网络梯度,进而加速了参数更新的速度,提高了网络的收敛速度。与不同层数的网络模型和不带跳跃密集连接块的网络模型相比,本文模型的收敛速度更快,效果更好。

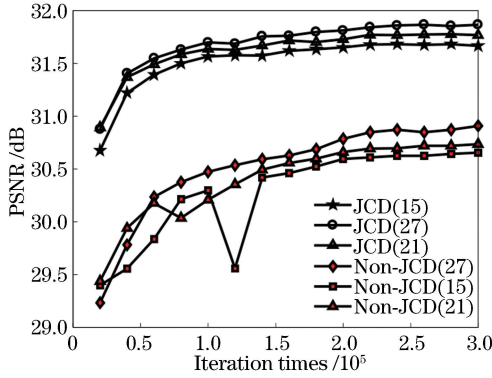


图 7 Set5 测试集下不同网络结构的迭代次数与 PSNR 均值的关系曲线

Fig. 7 Relationship between iteration times and average PSNR of different network structures under Set5 test set

4.3.3 激活函数

使用激活函数可以增加网络的非线性,提高网络模型的表达能力。进行两组实验,实验以 $f=4$ 为例,采用 Set5 作为测试集,分别使用 ReLU 函数与 PReLU 函数进行比较,ReLU 函数的表达式 $f(x)=\max(0,x)$,从这个表达式可以看出,负半轴的斜率为零,在训练过程中,一旦有数据进入负半区,就会发生神经元死亡,而 PReLU 函数 $\Phi(x)=\max(0,x)+\alpha_i \min(0,x)$, $\Delta\alpha_{i+1}=\mu\Delta\alpha_i+\eta\partial\epsilon/\partial\alpha_i$, PReLU 函数可以自适应地学习更新参数 α_i ,有效避免了神经元死亡。从图 8 中可以看出,PReLU 函数比 ReLU 函数的效果更好,收敛速度更快,所以选择 PReLU 函数作为本文的激活函数。

4.3.4 滤波器数目

滤波器数目的不同也会影响网络模型的效果。实验以 $f=4$ 为例,采用 Set5 作为测试集,如图 9 所示,分别测试滤波器数目为 8 和 16 时,可以发现随着滤波器数目的增大,网络模型的效果明显变好了。本文选择滤波器数目为 32 的网络模型。

4.3.5 运行时间

为了比较不同模型的运行时间,在相同的环境下进行了 5 组实验,实验以 $f=4$ 为例,采用 Set5 作为测试集,如图 10 所示,分别测试了 DCSR、VDSR^[15]、ESPCN^[14]、FSRCNN^[13] 和 SRCNN^[12],

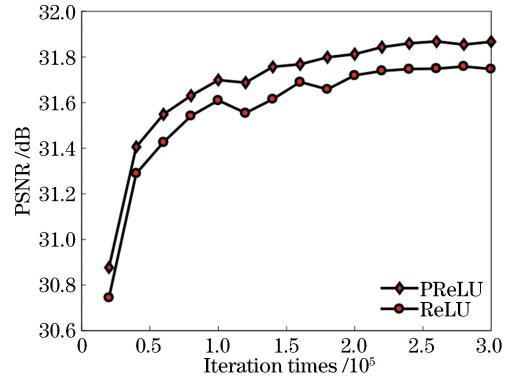


图 8 Set5 测试集下不同激活函数 PSNR 均值随迭代次数的变化曲线

Fig. 8 Relationship between average PSNR of different activation functions with number of iterations under Set5 test set

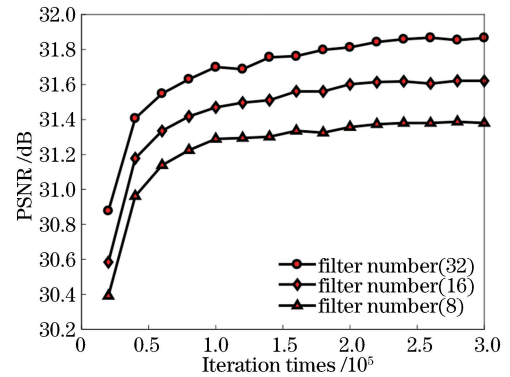


图 9 Set5 测试集下不同滤波器数目条件下的 PSNR 均值随迭代次数的变化曲线

Fig. 9 Relationship between average PSNR of different filter numbers with number of iterations under Set5 test set

可以发现模型 DCSR 的 PSNR 值最高,为 31.87 dB,说明在重建图像效果方面,本文的模型是最好的,这是因为本文的网络层数最多(27 层),运行时间为 0.756 s,比大部分模型都快,仅比 ESPCN 慢了 0.1 s。

4.4 实验结果分析

为了测试本文算法(DCSR)的性能,将其分别与 Bicubic、SRCNN^[12]、FSRCNN^[13]、ESPCN^[14] 和 VDSR^[15] 进行了比较,分别测试了图像放大 2, 3, 4 倍的效果,测试集采用了图像超分辨率重建领域常用的数据集,包括 Set5、Set14 和 BSD100,从中选取了边缘细节丰富的 3 张图像进行测试,并比较了图像放大后的细节。受数据集和迭代次数的影响,FSRCNN^[13]、VDSR^[15] 与本文算法的结果略有不同。从主观的视觉效果来看,Bicubic 的整体效果最模糊,看不清楚图像的细节信息。SRCNN^[12] 和 ESPCN^[14] 的整体效果比较好,但是图像边缘还是比

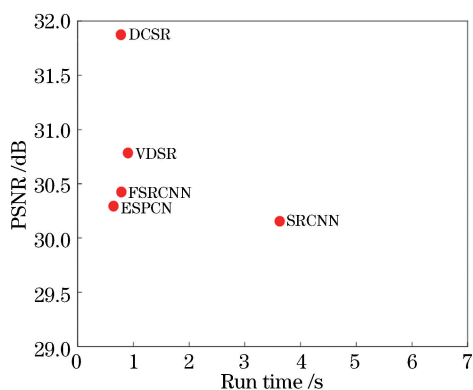


图 10 Set5 测试集下不同方法的运行时间与 PSNR 均值的关系曲线

Fig. 10 Relationship between run time and average PSNR of different methods under Set5 test set

较模糊,FSRCNN^[13]和VDSR^[15]的效果相对较好,细节恢复也相对完整,DCSR能恢复出更多的图像细节,边缘也更加锐利,辨识度较高。特别是放大倍数为4时,相比于其他算法,DCSR可以轻松发现视觉上的改进,图11中DCSR重建出的蝴蝶翅膀的边缘信息更加锐利,对比度更高,图12中DCSR重建出的How字母局部的高频信息更加丰富,边缘更加自然。图13中DCSR重建出的衣服细纹细节信息更清晰,主观效果更好。表1和表2为测试集Set5、Set14和BSD100下不同方法的PSNR和SSIM均值,可以看出,本文模型的PSNR值和SSIM值均是最高,从恢复出的图像效果来看,本文算法恢复出的图像最接近原始图像。

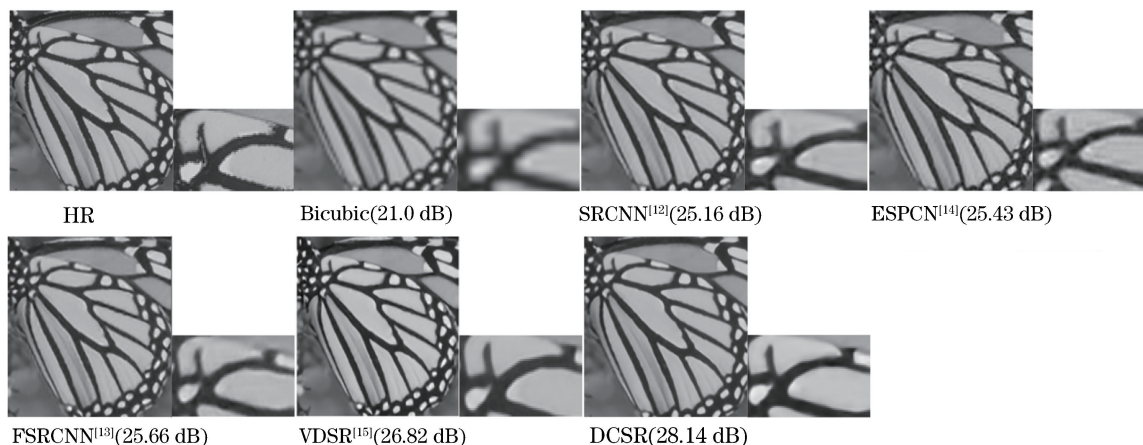


图 11 Butterfly 在不同算法下的效果对比图

Fig. 11 Comparison of butterfly images obtained by different algorithms



图 12 PPT 在不同算法下的效果对比图

Fig. 12 Comparison of PPT obtained by different algorithms

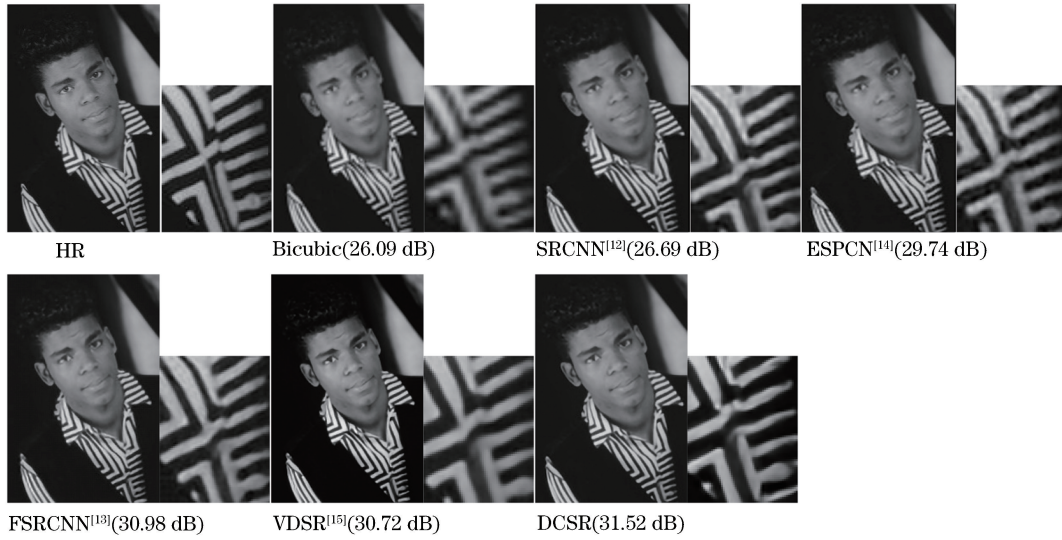


图 13 Man 在不同算法的对比图

Fig. 13 Comparison of man images obtained by different algorithms

表 1 在测试集 Set5, Set14, BSD100 下比较不同方法的 PSNR 均值

Table 1 Average PSNR of test sets Set5, Set14, and BSD100 under different algorithms

| Dataset | Scale | Bicubic | SRCNN ^[12] | ESPCN ^[14] | FSRCNN ^[13] | VDSR ^[15] | DCSR |
|---------|-------|---------|-----------------------|-----------------------|------------------------|----------------------|-------|
| Set5 | 2 | 33.68 | 36.19 | 36.38 | 36.45 | 37.34 | 37.70 |
| | 3 | 30.45 | 32.46 | 32.71 | 32.59 | 33.47 | 34.13 |
| | 4 | 28.46 | 30.15 | 30.29 | 30.42 | 30.78 | 31.87 |
| Set14 | 2 | 30.21 | 32.1 | 32.2 | 32.21 | 32.82 | 33.26 |
| | 3 | 27.51 | 28.99 | 29.12 | 29.12 | 29.51 | 29.97 |
| | 4 | 25.98 | 27.23 | 27.17 | 27.43 | 27.62 | 28.27 |
| BSD100 | 2 | 29.43 | 30.88 | 30.93 | 31.24 | 31.51 | 31.81 |
| | 3 | 27.08 | 28.06 | 28.16 | 28.25 | 28.43 | 28.79 |
| | 4 | 25.84 | 26.63 | 26.59 | 26.85 | 26.87 | 27.28 |

表 2 在测试集 Set5, Set14, BSD100 下比较不同方法的 SSIM 均值

Table 2 Average SSIM of test sets Set5, Set14, and BSD100 under different algorithms

| Dataset | Scale | Bicubic | SRCNN ^[12] | ESPCN ^[14] | FSRCNN ^[13] | VDSR ^[15] | DCSR |
|---------|-------|---------|-----------------------|-----------------------|------------------------|----------------------|--------|
| Set5 | 2 | 0.9306 | 0.9551 | 0.9568 | 0.9567 | 0.9580 | 0.9636 |
| | 3 | 0.8686 | 0.9110 | 0.9150 | 0.9122 | 0.9188 | 0.9321 |
| | 4 | 0.8102 | 0.8621 | 0.8629 | 0.8659 | 0.8750 | 0.8979 |
| Set14 | 2 | 0.8693 | 0.9576 | 0.9597 | 0.9634 | 0.9104 | 0.9644 |
| | 3 | 0.7744 | 0.8836 | 0.8873 | 0.8923 | 0.8271 | 0.8981 |
| | 4 | 0.7023 | 0.8207 | 0.8227 | 0.8273 | 0.7592 | 0.8446 |
| BSD100 | 2 | 0.8440 | 0.8801 | 0.8831 | 0.8867 | 0.8924 | 0.8964 |
| | 3 | 0.7401 | 0.7755 | 0.7811 | 0.7803 | 0.7928 | 0.7987 |
| | 4 | 0.6697 | 0.692 | 0.6943 | 0.7009 | 0.7186 | 0.7203 |

5 结 论

针对现有模型收敛速度慢,训练前需要对图片进行预处理,以及未充分利用各卷积层提取到的图像特征等问题,提出了一种基于跳跃密集连接的单幅图像超分辨率重建的方法,该算法充分利用各卷积层学习到的图像特征,加快了网络的收敛速度,并减少了网络的冗余性,且该算法使用亚像素卷积对图像进行放大。本文算法在评价指标和视觉效果上均优于现有算法,下一步将继续优化网络模型,在保证测试精度的同时,降低运行时间。

参 考 文 献

- [1] Zhang H, Zhang L, Shen H. 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] Bätz M, Eichenseer A, Seiler J, *et al.* Hybrid super-resolution combining example-based single-image and interpolation-based multi-image reconstruction approaches[C]//2015 IEEE International Conference on Image Processing (ICIP), September 27-30, 2015, Quebec City, QC, Canada. New York: IEEE, 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] Yang J, Wang Z, Lin Z, *et al.* Couple dictionary training for image super-resolution [J]. *IEEE Transactions on Image Processing*, 2012, 21 (8): 3467-3478.
- [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, June 27-July 2, 2004, Washington D C, USA. New York: IEEE, 2004: 1315043.
- [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] 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.
- [13] Dong C, Chen C L, Tang X. Accelerating the super-resolution convolutional neural network [M]//Leibe B, Matas J, Sebe N, *et al.* *Lecture notes in computer science*. Cham: Springer, 2016, 9906: 391-407.
- [14] 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]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1874-1883.
- [15] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 1646-1654.
- [16] He K M, Zhang X Y, Ren S Q, *et al.* Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [17] Huang G, Sun Y, Liu Z, *et al.* Deep networks with stochastic depth [M]//Leibe B, Matas J, Sebe N, *et al.* *Lecture notes in computer science*. Cham: Springer, 2016, 9908: 646-661.
- [18] Huang G, Liu Z, Maaten L V D, *et al.* Densely connected convolutional networks [C]//2017 IEEE

- Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2261-2269.
- [19] Qu Y Y, Lin L, Shen F M, *et al.* Joint hierarchical category structure learning and large-scale image classification [J]. IEEE Transactions on Image Processing, 2017, 26(9): 4331-4346.
- [20] 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.
- [21] Jia Y, Shelhamer E, Donahue J, *et al.* Caffe: convolutional architecture for fast feature embedding [EB/OL]. (2014-06-20) [2019-03-09]. <https://arxiv.org/abs/1408.5093>.
- [22] Timofte R, de Smet V, van Gool L. A+: adjusted anchored neighborhood regression for fast super-resolution[M]//Cremers D, Reid I, Saito H, *et al.* Lecture notes in computer science, Cham: Springer, 2014, 9006: 111-126.
- [23] 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 Eighth IEEE International Conference on Computer Vision, July 7-14, 2001, Vancouver, BC, Canada. New York: IEEE, 2001: 416-423.
- [24] Du B, Wang Z M, Zhang L F, *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.