

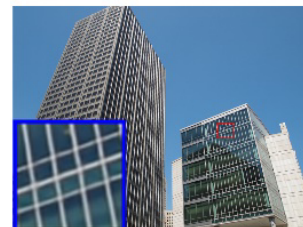


DOI: 10.12086/oe.2019.180489

多路径递归网络结构的单帧 图像超分辨率重建

沈明玉, 俞鹏飞, 汪荣贵, 杨娟*, 薛丽霞

合肥工业大学计算机与信息学院, 安徽 合肥 230009



摘要: 卷积神经网络在单帧图像超分辨率重建任务中取得了巨大成功, 但是其重建模型多是基于单链结构, 层间联系较弱且不能充分利用网络提取的分层特征。针对这些问题, 本文设计了一种多路径递归的网络结构(MRCN)。通过使用多路径结构来加强层之间的联系, 实现特征的有效利用并且提取丰富的高频成分, 同时使用递归结构降低训练难度。此外, 通过引入特征融合的操作使得在重建的过程中可以充分利用各层提取的特征, 并且自适应的选择有效特征。在常用的基准测试集上进行了大量实验表明, MRCN 比现有的方法在重建效果上具有明显提升。

关键词: 卷积神经网络; 超分辨; 分层特征; 多路径; 特征融合

中图分类号: TB872

文献标志码: A

引用格式: 沈明玉, 俞鹏飞, 汪荣贵, 等. 多路径递归网络结构的单帧图像超分辨率重建[J]. 光电工程, 2019, 46(11): 180489

Image super-resolution via multi-path recursive convolutional network

Shen Mingyu, Yu Pengfei, Wang Ronggui, Yang Juan*, Xue Lixia

School of Computer and Information, Hefei University of Technology, Hefei, Anhui 230009, China

Abstract: Convolutional neural network (CNN) has recently achieved a great success for single image super-resolution (SISR). However, most deep CNN-based super-resolution models use chained stacking to build the network, which results in the fact that the relationship between layers is weak and does not make full use of hierarchical features. In this paper, a multi-path recursive convolutional network (MRCN) is designed to address these problems in SISR. By using multi-path structure to strengthen the relationship between layers, our network can effectively utilize features and extract rich high-frequency components. At the same time, we also use recursive structure to alleviate training difficulty. In addition, by introducing the operation of feature fusion into the model, our network can make full use of the features extracted from each layer in the reconstruction process and select the effective features adaptively. Extensive experiments on benchmarks datasets have shown that MRCN has a significant performance improvement against existing methods.

Keywords: convolutional neural network; super-resolution; hierarchical features; multi-path; feature fusion

Citation: Shen M Y, Yu P F, Wang R G, *et al.* Image super-resolution via multi-path recursive convolutional network[J]. *Opto-Electronic Engineering*, 2019, 46(11): 180489

收稿日期: 2018-09-17; 收到修改稿日期: 2018-12-28

基金项目: 国家自然科学基金资助项目(61672202)

作者简介: 沈明玉(1963-), 男, 副教授, 主要从事模式识别、网络与信息安全的研究。E-mail: shenmy@126.com

通信作者: 杨娟(1983-), 女, 讲师, 主要从事深度学习、智能信息处理的研究。E-mail: yangjuan6985@163.com

1 引言

单帧图像超分辨重建技术可以从一张低分辨率图像中恢复出包含更多高频细节的高分辨率图像,因此被广泛用于医学影像、卫星遥感以及视频监控等领域。目前出现的超分辨率算法主要包括三类:基于插值的算法^[1-2]、基于重建的算法^[3-4]和基于学习的算法^[5-8]。

近年来,随着深度学习的广泛应用,基于深度学习的算法逐渐成为主流的图像重建算法。Dong 等^[9-10]首先提出了超分辨重建卷积神经网络,论证了卷积神经网络(convolutional neural network, CNN)可以通过端到端的方式学习低分辨率图像和高分辨率图像之间的映射,其原理等价于稀疏编码^[11]。后来, Kim 等^[12]加深了网络的深度,构建了深度超分辨网络(very deep network for super-resolution, VDSR),并采用残差连接和梯度裁剪的方式降低了训练的困难,其后在 VDSR 的基础上引入了递归监督的学习策略构建了深度重建网络(deep reconstruction-classification networks, DRCN)^[13],降低网络的参数。为了进一步降低网络的训练难度, IRCNN^[14]和 MSSRNet^[15]分别基于 VGG 和 GoogLeNet 网络结构,引入扩张卷积的技术,在扩大感受野的同时进一步降低了网络参数,取得了较好的效果。

上述提及的网络结构多为单路径前馈结构,即当前层的信息主要来自于其直接前继层,忽略其他前继层的影响,如 VDSR^[12]、DRCN^[13]和 IRCNN^[14]等,其都是通过层之间的链式堆叠构建网络结构,层间联系较弱。后来又出现了 CNN 的一些变体,如 SRResNet^[16],实现了跨一层传递信息,随后又出现了 DRRN^[17]、RED^[18]等,在加深网络深度的同时,利用跳跃连接实现了跨多层传递特征,使得当前层可以融合多层提取的特征获得新的特征,提高了网络的学习能力,加强了层间联系。在此基础之上, Tai 等^[19]又提出了具有更深层结构的 MemNet,其使用记忆块去获得并融合网络多层提取的特征,通过密集连接将所有记忆块联系起来,加强了块间联系。随着网络深度的增加,带有不同感受野的分层特征被逐层提取出来。但是,上述这些方法大都忽略了分层特征的作用以及对分层特征的多次学习,导致网络不能充分的提取高频成分,使得在重建过程中不能很好的恢复高频信息。MemNet^[19]虽然实现融合多个记忆块提取的特征完成重建,但是对于块内的卷积层仍然是单链结构,这样虽然降低了网络的复杂度,但却不利于充分利用已学习到的分层

特征,降低了网络学习能力。为了解决这个问题,本文引入了多路径递归结构,使得当前层提取的信息能够被后面连续的多层共享,提升网络的学习能力。还利用网络中所有层提取的特征同时用于重建高分辨率图像,保证了特征的丰富度,采用完全自适应的方式选择有效特征,减轻了无用特征在重建过程的干扰。

另一方面,这些网络都需要对原始的低分辨率图像进行插值处理,并不能真正实现端到端的重建,该操作不仅增加了计算复杂度,还导致了原始的低分辨率图像一些细节的丢失。针对此问题, Dong 等^[20]直接将低分辨率图像作为输入,在网络最后使用反卷积完成上采样,网络的整个操作均在低分辨率空间进行,降低了训练难度。Shi 等^[21]在 ESPCN 中提出了亚像素卷积层,进一步降低了计算复杂度。此两种方法的网络结构非常简单,不能很好地完成低分辨率到高分辨率之间复杂的映射。本文将模型的特征提取以及非线性映射集中在低分辨率空间,同时改进了网络的结构,使之更好地模拟出端到端之间的映射关系。且没有直接从低分辨空间生成高分辨率图像,而是利用一个卷积层作为过渡,先完成特征的上采样,接着在高分辨率空间恢复图像,这样进一步增强网络的非线性能力。

本文提出了一种多路径递归卷积网络(multipath recursive convolutional network, MRCN)去完成图像重建。网络使用低分辨图像作为输入,利用状态转移来模拟多路径递归结构中特征在向后传递过程中的变化,在状态转移的过程中提取新的高频信息,增强网络的非线性映射能力;同时通过状态转移,实现对某层特征的多次学习,并保存所有学习到的特征;然后通过自适应的方式增强有效特征,抑制无用的特征;最后,使用 ESPCN 中提出的亚像素卷积完成上采样。

2 算法描述

2.1 MRCN 网络模型

本文的网络结构如图 1(a)所示,主要由三部分构成:特征提取网络、多路径递归结构、特征融合过程。 x 和 y 表示 MRCN 的输入和输出,网络的目的是学习低分辨图像 x 到高分辨率图像 y 之间的映射函数。首先,在特征提取网络中利用两个卷积层提取网络的浅层特征,第一个卷积层从 x 中提取特征表示为 F_{-1} ,则:

$$F_{-1} = f_{\text{EXT1}}(x) \quad (1)$$

这里 f_{EXT1} 表示卷积操作, F_{-1} 用于进一步提取浅层特征以及学习全局特征残差,接下来有:

$$F_0 = f_{EXT2}(F_{-1}), \quad (2)$$

其中： f_{EXT2} 为卷积操作， F_0 为提取的浅层特征，将作为多路径递归的输入和用于学习局部特征残差任务。

假设多路径递归存在 M 次完全状态转换，即 M 轮次递归过程，每次完全状态转换包含 K 个卷积层，取 $K=4$ 。将一个卷积层的操作称为一次状态转移，每次完全状态转换包含 K 次状态转移的过程，则有：

$$F_m = f_m(F_{m-1}), \quad (3)$$

其中： $m \in [1, 2, \dots, M]$ ， F_m 表示第 m 次完全状态转换后的输出，它是由状态转移而提取的特征连接而成，具体过程如 2.2 节所述。 f_m 表示第 m 次完全状态转换操作的集合，此外，本文中每轮完全状态转换操作涉及的卷积层具有相同的权重。在通过式(3)中提取了多路径递归中所有的分层特征后，需要进行两次特征融合得到对重建有益的特征，如下：

$$F_{GR} = f_{TF}(F_1, F_2, \dots, F_M), \quad (4)$$

其中： f_{TF} 表示两次特征融合操作，在 2.3 节将会详细展开， F_{GR} 表示得到的全局特征残差。然后利用全局残差学习获得上采样之前网络提取的特征图 F_{GF} ：

$$F_{GF} = F_{-1} + F_{GR}. \quad (5)$$

本文模型在低分辨率空间中完成图像特征的提取以及非线性映射，然后用亚像素卷积层完成特征图的上采样，最后利用一个卷积层在高分辨空间中完成重建。则 MRCN 的输出可以表示为

$$y = F_{MRCN}(x), \quad (6)$$

其中 F_{MRCN} 表示本文提出的网络。

给定训练集 $\{x^i, \tilde{x}^i\}_{i=1}^N$ ，其中 N 为训练样本的数量， \tilde{x}^i 表示低分辨率图像块 x^i 对应的高分辨率图像块。对于损失函数，本文使用 Charbonnier 损失函数，其可以缓解均方误差带来的图像过平滑的问题。令 θ 表示网络的参数，损失函数 L 表示如下：

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \rho(\tilde{x}^i - F_{MRCN}(x^i)), \quad (7)$$

其中 $\rho(t) = \sqrt{t^2 + \varepsilon^2}$ 是 Charbonnier 惩罚函数，根据文献[22]，本文将 ε 设置为 $1E-3$ 。

2.2 多路径递归

Huang 等^[23]设计了密集块结构，与其他网络结构不同，在每一个密集块中，其每一层都与其所有的前驱层直接相连，提高了各层特征的复用率，增强了层间联系。该模块提升了信息和梯度在网络中的传输效率，并且还有正则化的效果。随后，Chen 等^[24]在 DPN 中深度地分析了这种连接方式的优点，前驱层提取的特征不仅可以充分被后继层复用，而且探索了全新的特征，保证了提取特征的丰富性。本文基于密集连接和递归策略提出了多路径递归结构，提取了图像的深层特征信息，增强了网络的非线性映射能力。首先，如图 1(b)的 State 阶段，利用密集连接生成初始状态，增长率 $g=32$ 。该阶段用于生成特征相关的 P 个部分， P 的值与状态转移次数 K 相同也为 4。因此将特征提取网络提取的浅层特征 F_0 作为输入特征时，则有：

$$H_0 = \max(0, w_0 * F_0), \quad (8)$$

其中： H_0 表示特征相关的第一个部分， $*$ 表示卷积操

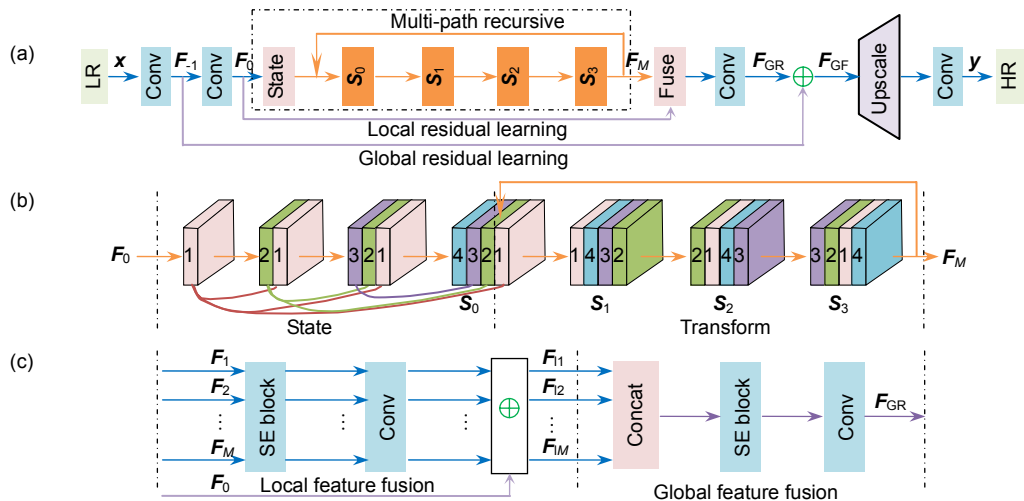


图 1 网络结构。(a) 本文的基本网络结构；(b) 多路径递归过程，其利用状态转移来模拟，“State”表示生成组成状态的部分，“Transform”表示递归中状态转移，对应(a)中橙色部分；(c) 特征融合，对应(a)中的 Fuse
Fig. 1 Basic architectures. (a) The architecture of our proposed multi-path recursive convolutional network; (b) The multi-path recursive structure. State transitions are used here to simulate this process, where “State” represents the generation of different states and “Transform” represents state transitions in recursion; (c) Feature fusion, corresponding to the “fuse” in (a).

作,那么由密集连接得到的剩余的 $P-1$ 个部分如下:

$$H_p = \max(0, w_p * [H_{p-1}, H_{p-2}, \dots, H_0]) , \quad (9)$$

其中: $p \in [1, \dots, P-1]$, H_p 表示初始状态为 S_0 的第 $p+1$ 个部分, $[H_{p-1}, H_{p-2}, \dots, H_0]$ 表示将其中的特征图进行通道连接操作, w_p 表示第 p 个卷积层权重,所以初始状态 $S_0 = [H_{p-1}, H_{p-2}, \dots, H_0]$, 与图 1(b)中的 S_0 一致。

用式(8)和式(9)得到初始状态 S_0 的各部分后,就可以在 Transform 阶段完成状态的转移,本文通过图 1(b) $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_0$ 状态转移又回到初始状态表示一轮递归的结束,即为一次完全状态转移,其状态转移通过卷积操作完成。图 2 是图 1(b)中的 Transform 阶段的具体过程,从图中可以看出每一次状态转移都使用前一次的状态生成自己特有的部分,然后与前面新生成的 $P-1$ 个部分进行通道连接作为下一次状态转移的输入,即对特征图最早生成的部分进行一次更新,在图 1(b)中 $S_1 \rightarrow S_2$, 在图 2 中表现为 H_1 的更新,完成一次完全状态转移相当于对初始状态 S_0 的一次完全更新。多路径结构将完全更新的特征图用于最后的重建。在第 m 轮完全状态转移过程中,状态 S_i 转移到下一状态 $S_{(i+1) \bmod P}$ 的过程可以表述如下:

$$S_i = f_{m,i}(S_i) , \quad (10)$$

$$S_{(i+1) \bmod P} = [H_i, H_{i-1}, \dots, H_0, H_{p-1}, H_{p-2}, \dots, H_{i+1}] , \quad (11)$$

其中: $i \in [0, 1, \dots, K-1]$, $f_{m,i}$ 表示第 m 轮完全状态转移中第 i 次状态转移。在状态转移过程中不断探索新的高频信息并保留提取的分层特征。最后将所有的分层特征用于重建。由式(10)和式(11)可知,将所有更新的 H_p 连接起来即为每一轮完全状态转移的输出 F_m :

$$F_m = [H_{p-1}, H_{p-2}, \dots, H_0] 。 \quad (12)$$

2.3 特征融合

通过状态转移以及递归轮转,得到了各卷积层提取的分层特征,然后利用局部特征融合和全局特征融合方式选择分层特征,如图 1(c)所示。受到 SENet^[25] 的启发,本文在特征融合过程中使用 SENet 中的 Squeeze 和 Excitation 操作去自动学习每个通道的重要程度,然后依照这个重要程度去增强有用的特征并抑制对当前任务用处不大的特征。如图 3 所示,为了获

得每个通道的全局空间信息,本文采取与 SENet 相同的方法, Squeeze 函数式表示为

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) , \quad (13)$$

其中: z_c 表示第 c 个通道的通道描述符, F_{sq} 表示全局平均池化, u_c 表示输入的第 c 个通道, H 和 W 表示输入的高和宽。Excitation 函数描述如下:

$$e = F_{ex}(z) = \sigma(W_2 * (\max(0, W_1 * z))) , \quad (14)$$

其中: F_{ex} 表示 Excitation 函数, e 则反应每个通道的重要程度, W_1 和 W_2 分别表示 1×1 的卷积, σ 表示 Sigmoid 函数。最后,采取基于通道的像素级别的乘积运算 F_{scale} 得到第 c 个通道的最终输出 \tilde{x}_c :

$$\tilde{x}_c = F_{scale}(u_c, e_c) = e_c \cdot u_c 。 \quad (15)$$

局部特征融合被用于融合一次完全状态转移产生的所有新的特征,与 MemNet 使用门单元去自适应的选择有效特征不同,本文首先利用 SE block 将每一轮完全状态转移得到的所有分层特征进行自适应加权处理,提升重要特征的作用,并且也用于解决密集连接带来的特征冗余的问题,然后使用 1×1 的卷积进一步控制输出信息,融合前面的分层特征,同时也减少了特征通道的数量。 F_m 表示第 m 轮完全状态转移得到的所有分层特征,局部特征融合的输出 F_{lm} 表示为

$$F_{lm} = W(f_{SE}(F_m)) + F_0 , \quad (16)$$

其中: f_{SE} 表示图 3 所示结构,即式(13),式(14)和式(15)所述过程。 W 表示一个权重层, $m \in [1, 2, \dots, M]$ 。这里为了进一步改进网络中的信息流,增强梯度在网络中的传递,本文还引入了局部残差学习。

全局特征融合用于融合所有分层特征的信息共同完成重建,其整个过程与局部特征融合相似,首先将所有的局部融合的特征连接起来,利用 SE block 进行选择,然后使用 1×1 的卷积进一步融合信息,生成全局特征残差,公式如下:

$$F_{GR} = W(f_{SE}([F_{11}, F_{12}, \dots, F_{LM}])) , \quad (17)$$

其中: F_{GR} 表示学习到的全局特征残差, $[F_{11}, F_{12}, \dots, F_{LM}]$ 表示连接每一轮局部融合的特征从而将所有的局部的特征放到全局进行处理。

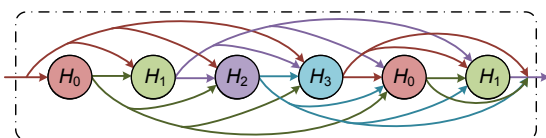


图 2 多路径递归过程,其颜色对应图 1(b)中各部分
Fig. 2 The process of multi-path recursive. Its color corresponds to each part in Fig. 1(b)

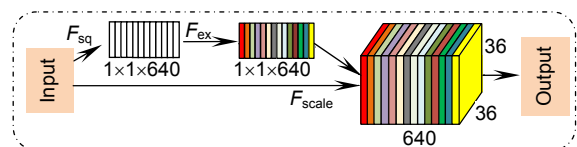


图 3 SE block 结构图
Fig. 3 The architecture of SE block

3 实验结果及分析

首先介绍用于训练和测试的数据集和实验的基本设置；其次描述实验的具体细节，包括三个部分：多路径递归结构、损失函数以及特征融合单元；最后将MRCN的重建效果与经典方法进行比较。用峰值信噪比(peak signal to noise ratio, PSNR)和结构相似性(structural similarity index measurement, SSIM)作为客观评价指标。

3.1 实验设置

数据集：本文使用了 Timofte 等^[26]发布的用于图像恢复的高质量数据集 DIV2K 训练模型，其包含 800 张训练图像。对于测试，使用了四个广泛使用的基准数据集作为测试集，分别为 Set5 测试集^[27]、Set14 测试集^[28]、BSD100 测试集^[29]和 Urban100 测试集^[30]。由于人眼对亮度通道较敏感，故实验中只对亮度通道 Y 进行处理，色度通道 Cb、Cr 使用插值的方式放大。

训练设置：为了缩短训练时间，减小存储复杂度，将训练图像裁剪成 108×108 大小的图像块，步长为 108，并将其下采样 2×、3×和 4×分别用于训练。训练网络使用 Adam 优化器，批量数据规模设置为 64，权重衰减系数为 10^{-4} ，初始学习率为 0.1，每隔 100 轮学习率减小一倍，共训练 200 轮。除了特征融合模块使用的卷积核大小为 1×1 之外，其余的均为 3×3。对于权重初始化，使用了 He 等^[31]提出的方法，它被证明适用于激活函数为 ReLu 的网络。网络中每一个卷积层都不含有偏置项。实验硬件环境为 CPU：7700K，GPU：1070，内存：8 G，深度学习架构使用 PyTorch。

3.2 实验分析

1) 多路径递归结构

本文将 VDSR 的结构作为基准模型，然后构建多

路径结构，如图 1 所示，其显示的是状态数目 $s=4$ 时的结构图。这里分别取 $s=1,2,\dots,8$ ，其中 $s=1$ 即为 VDSR 结构，图 4(a)曲线表示随着 s 值的增加，PSNR 的变化情况，从图中可以看出，随着 s 的增大，在开始时，PSNR 增加的很快，而后增幅减缓，并且使用多路径结构的峰值信噪比要高于单链前馈结构网络。

其次，本文将递归结构引入到多路径前馈网络中，缓解了由网络深度增加引起的参数量大幅提升的问题。本文测试了递归轮数，也即完全状态转移次数对 PSNR 的影响，从图 4(b)可以看出，递归次数越多，PSNR 值越高。增加完全状态转移次数就增加了网络的深度，即通过更大的感受野以及更多的非线性映射来提高网络性能。但是递归的结构保证了参数量不会大幅增加。

2) 损失函数分析

均方误差或者 L2 损失是超分辨率重建领域广泛使用的损失函数。然而，该损失函数会导致重建的图像过于平滑，Zhao 等^[32]的研究也发现在 PSNR 和 SSIM 两个度量指标下，在 L2 损失下训练的网络并不能保证有最好的效果。在他们的实验中，用 L1 损失训练的网络比用 L2 损失训练的网络性能更好。而本文采用的是 L1 损失函数的一种变体 Charbonnier 损失函数^[33]，其对于奇异点的处理具有更好的鲁棒性。

在所有实验环境均相同的情况下，本文分别使用 Charbonnier 损失函数、L1 损失函数和均方差损失函数去训练网络，得到其收敛状况如图 5 所示。从图中可以看出，L1 损失函数和 Charbonnier 损失函数在收敛后的效果优于均方差损失函数。本文计算了 L1 损失函数和 Charbonnier 损失函数训练的各轮的平均 PSNR 及方差分别为 33.89 和 33.92 以及 0.3355 和 0.3051，由此可见 Charbonnier 损失函数的效果略好于 L1 损失函

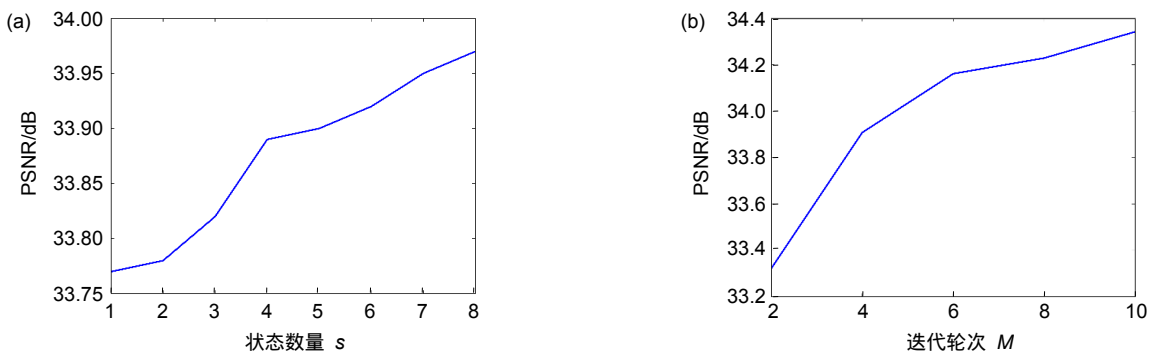


图 4 多路径递归结构对网络性能的影响。(a) 状态数量对重建效果的影响；(b) 递归轮数对重建效果的影响
Fig. 4 The effect of multi-path on network performance. (a) The effect of the number of states on SR; (b) The effect of the number of recursive-rounds on SR

数，且收敛过程相比于 L1 损失函数更加稳定。

3) 特征融合

DRRN 使用卷积层来融合特征，MemNet 通过门单元来融合以及选择特征，其本质也是卷积操作，但是 SENet 为选择特征提供了新的思路，其通过自适应的方式选择有效特征，抑制无效特征，有效地缓解了特征融合过程中无用特征的干扰。如图 6 所示，本文分别训练了两个网络，其中一个使用 SE block 模块，为了保证公平性，这里将 SE block 替换成卷积层构成一个对比的网络，从图中可以看出，存在 SE block 的

网络效果优于单纯使用卷积进行融合的网络，且收敛过程相对来说更加平稳。

3.3 基准测试

本文比较了 MRCN 与 A+^[8], SRCNN^[9-10], VDSR^[12], DRCN^[13], DRRN^[17]和 MemNet^[19]的重建效果，考虑到参数的性能和数量，MRCN 选择 K=8 的包含 10 次完全状态转移的网络作为最终的结构。测试数据集为 Set5, Set14, BSD100, Urban100。表 1 为 2×、3×和 4×模型的比较，红色表示方法的最佳性能，蓝色表示

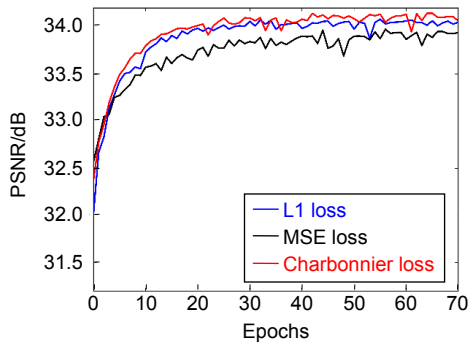


图 5 损失函数的选择对网络性能的影响

Fig. 5 The impact of the choice of loss function on network performance

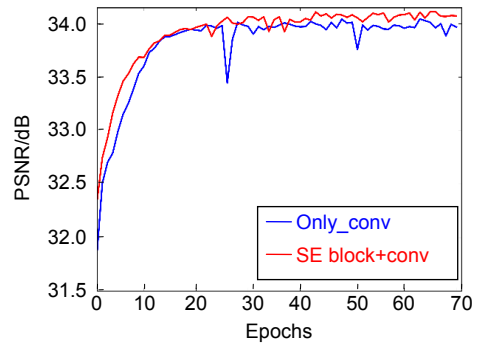


图 6 SE block 对网络性能的影响

Fig. 6 The impact of SE block on network performance

表 1 数据集 Set5、Set14、BSD100 以及 Urban100 在比例因子为 2×、3×和 4×下的平均 PSNR/SSIM

Table 1 Average PSNR/SSIMs for scale factors of 2×, 3× and 4× on datasets Set5, Set14, BSD100 and Urban100

Scale	Method	Set5	Set14	BSD100	Urban100
		PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
2×	Bicubic	33.66 /0.9299	30.24/0.8688	29.56/0.8431	26.88/0.8403
	A+ ^[8]	36.54/0.9544	32.28/0.9056	31.21/0.8863	29.20/0.8938
	SRCNN ^[9-10]	36.66/0.9542	32.42/0.9063	31.36/0.8879	29.50/0.8946
	VDSR ^[12]	37.53/0.9587	33.03/0.9124	31.90/0.8960	30.76/0.9140
	DRCN ^[13]	37.63/0.9588	33.04/0.9118	31.85/0.8942	30.75/0.9133
	DRRN ^[17]	37.74/0.9591	33.23/0.9136	32.05/0.8973	31.23/0.9188
	MemNet ^[19]	37.78/0.9597	33.28/0.9142	32.08/0.8978	31.31/0.9195
MRCN(ours)	37.89/0.9602	33.37/0.9163	32.12/0.8985	31.36/0.9231	
3×	Bicubic	30.39/0.8628	27.55/0.7742	27.21/0.7385	24.46/0.7349
	A+ ^[8]	32.58/0.9088	29.13/0.8188	28.29/0.7835	26.03/0.7973
	SRCNN ^[9-10]	32.75/0.9090	29.28/0.8209	28.41/0.7863	26.24/0.7989
	VDSR ^[12]	33.66/0.9213	29.77/0.8314	28.82/0.7976	27.14/0.8279
	DRCN ^[13]	33.82/0.9226	29.76/0.8311	28.80/0.7963	27.15/0.8276
	DRRN ^[17]	34.03/0.9244	29.96/0.8349	28.95/0.8004	27.53/0.8378
	MemNet ^[19]	34.09/0.9248	30.00/0.8350	28.96/0.8001	27.56/0.8376
MRCN(ours)	34.24/0.9267	30.16/0.8413	29.06/0.8022	27.60/0.8391	
4×	Bicubic	28.24/0.8104	26.00/0.7027	25.96/0.6675	23.14/0.6577
	A+ ^[8]	30.28/0.8603	27.32/0.7491	26.82/0.7087	24.32/0.7183
	SRCNN ^[9-10]	30.48/0.8628	27.49/0.7503	26.90/0.7101	24.52/0.7221
	VDSR ^[12]	31.35/0.8838	28.01/0.7674	27.29/0.7251	25.18/0.7524
	DRCN ^[13]	31.53/0.8854	28.02/0.7670	27.23/0.7233	25.14/0.7510
	DRRN ^[17]	31.68/0.8888	28.21/0.7721	27.38/0.7284	25.44/0.7638
	MemNet ^[19]	31.74/0.8893	28.26/0.7723	27.40/0.7281	25.50/0.7630
MRCN(ours)	31.83/0.8904	38.31/0.7732	27.44/0.7301	25.52/0.7641	

方法的次最佳性能。从表中可以看出, MRCN 比其他方法均有较好的效果。对于 2×数的重建, 本文网络在 PSNR 上平均提高了 0.072 dB, SSIM 平均提高 0.0032, 对于 4×数的重建, 在 PSNR 上平均提高了 0.05 dB, SSIM 平均提高 0.0012。

为了进一步验证本文算法的有效性, 这里对比本文方法和其他方法重建所得图像的视觉效果。图 7 和图 8 分别表示放大 3 倍时重建图像“img096”和“ppt3”的局部细节放大图。从图 7 中可以看出, 本文的 MRCN 方法较好地恢复了建筑的线条信息, 并且较好地反应了线条之间的层次信息, 且 PSNR/SSIM 均高于其他方法。图 8 中, 本文的重建方法较好地恢复了文本的轮廓信息。这表明了本文算法与其他算法相比有更好的

重建表现, 对不同的图像特征能够生成更准确的超分辨率估计。

4 结 论

本文提出了一种多路径递归结构的超分辨率重建网络, 利用卷积神经网络学习低分辨率图像到高分辨率图像的映射。利用多路径结构实现了对分层特征的多次学习, 提高了分层特征利用率, 同时也提升了网络学习特征的能力, 然后结合递归学习使得在网络深度提升的同时网络参数不会出现大幅增加, 接着利用特征融合处理所有的分层特征, 并通过 SE block 选择有效特征完成重建任务。实验结果表明本文的方法比其他经典方法具有更好的重建效果。

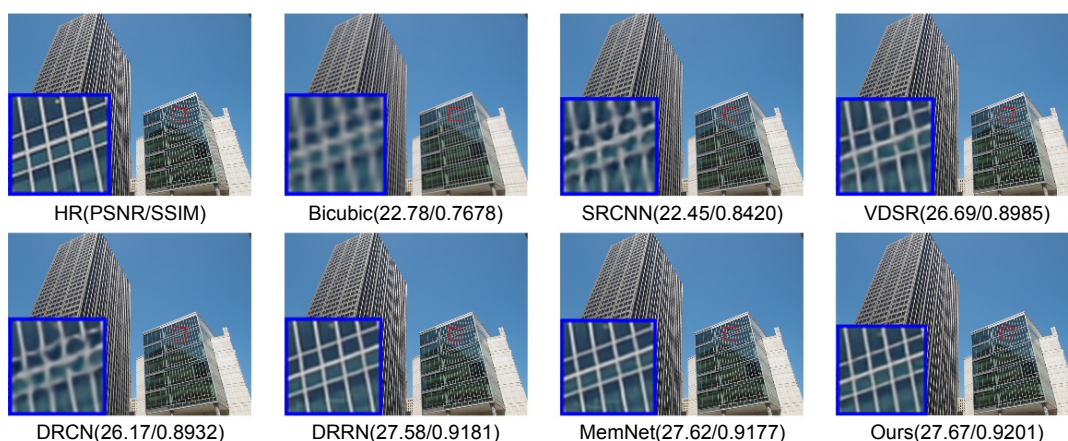


图 7 Urban 数据集中的“img096”图片 3×重建
Fig. 7 “img096” from Urban with an upscaling factor of 3×



图 8 Set14 数据集中的“ppt3”图片 3×重建
Fig. 8 “ppt3” from Urban with an upscaling factor of 3×

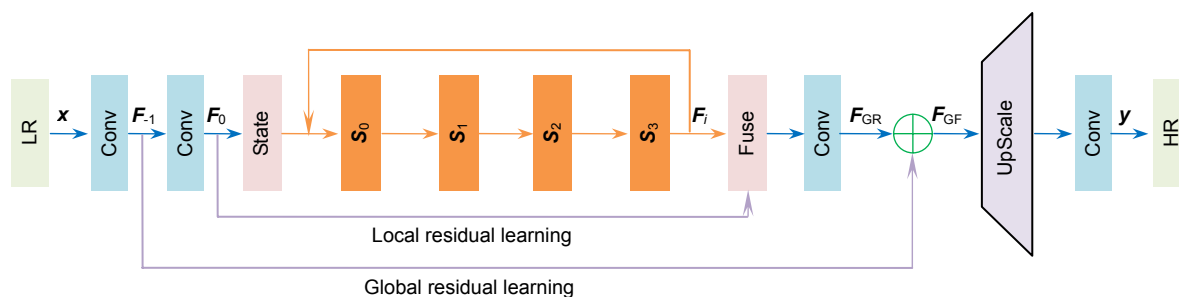
参考文献

- [1] Li X, Orchard M T. New edge-directed interpolation[J]. *IEEE Transactions on Image Processing*, 2001, **10**(10): 1521–1527.
- [2] Zhang L, Wu X L. An edge-guided image interpolation algorithm via directional filtering and data fusion[J]. *IEEE Transactions on Image Processing*, 2006, **15**(8): 2226–2238.
- [3] Dai S Y, Han M, Xu W, et al. SoftCuts: a soft edge smoothness prior for color image super-resolution[J]. *IEEE Transactions on Image Processing*, 2009, **18**(5): 969–981.
- [4] Sun J, Xu Z B, Shum H Y. Image super-resolution using gradient profile prior[C]//*Proceedings of 2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008: 1–8.
- [5] Wu C Z, Hu C S, Zhang M J, et al. Single image super-resolution reconstruction via supervised multi-dictionary learning[J]. *Opto-Electronic Engineering*, 2016, **43**(11): 69–75.
吴从中, 胡长胜, 张明君, 等. 有监督多字典学习的单幅图像超分辨率重建[J]. *光电工程*, 2016, **43**(11): 69–75.
- [6] Wang R G, Wang Q H, Yang J, et al. Image super-resolution reconstruction by fusing feature classification and independent dictionary training[J]. *Opto-Electronic Engineering*, 2018, **45**(1): 170542.
汪荣贵, 汪庆辉, 杨娟, 等. 融合特征分类和独立字典训练的超分辨率重建[J]. *光电工程*, 2018, **45**(1): 170542.
- [7] Timofte R, De Smet V, Van Gool L. Anchored neighborhood regression for fast example-based super-resolution[C]// *Proceedings of 2013 IEEE International Conference on Computer Vision*, 2013: 1920–1927.
- [8] 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. *Computer Vision—ACCV 2014*. Cham: Springer, 2014: 111–126.
- [9] Dong C, Loy C C, He K M, et al. Learning a deep convolutional network for image super-resolution[C]//*Proceedings of the 13th European Conference on Computer Vision*, 2014: 184–199.
- [10] 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.
- [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] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1646–1654.
- [13] Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1637–1645.
- [14] Zhang K, Zuo W M, Gu S H, et al. Learning deep CNN Denoiser prior for image restoration[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 2808–2817.
- [15] Shi W Z, Jiang F, Zhao D B. Single image super-resolution with dilated convolution based multi-scale information learning inception module[C]//*Proceedings of 2017 IEEE International Conference on Image Processing*, 2017: 977–981.
- [16] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 105–114.
- [17] Tai Y, Yang J, Liu X M. Image super-resolution via deep recursive residual network[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 2790–2798.
- [18] Mao X J, Shen C H, Yang Y B. 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*, 2016: 2810–2818.
- [19] Tai Y, Yang J, Liu X M, et al. MemNet: a persistent memory network for image restoration[C]//*Proceedings of 2017 IEEE International Conference on Computer Vision*, 2017: 4549–4557.
- [20] Dong C, Loy C C, Tang X O. Accelerating the super-resolution convolutional neural network[C]//*Proceedings of the 14th European Conference on Computer Vision*, 2016: 391–407.
- [21] 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]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 1874–1883.
- [22] Lai W S, Huang J B, Ahuja N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution[C]// *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 5835–5843.
- [23] Huang G, Liu Z, van der Maaten L, et al. Densely connected convolutional networks[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 2261–2269.
- [24] Chen Y P, Li J A, Xiao H X, et al. Dual path networks[C]//*Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017: 4470–4478.
- [25] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, doi: 10.1109/TPAMI.2019.2913372.
- [26] Timofte R, Agustsson E, van Gool L, et al. NTIRE 2017 challenge on single image super-resolution: methods and results[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017: 1110–1121.
- [27] 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*, 2012: 135.1–135.10.
- [28] 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*, 2010: 711–730.
- [29] Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars[C]//*Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 5197–5206.
- [30] 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*, 2001: 416–423.
- [31] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification[C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*, 2015: 1026–1034.
- [32] Zhao H, Gallo O, Frosio I, et al. Loss functions for neural networks for image processing[OL]. arXiv:1511.08861[cs.CV], 2015.
- [33] Bruhn A, Weickert J, Schnörr C. Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods[J]. *International Journal of Computer Vision*, 2005, **61**(3): 211–231.

Image super-resolution via multi-path recursive convolutional network

Shen Mingyu, Yu Pengfei, Wang Ronggui, Yang Juan*, Xue Lixia

School of Computer and Information, Hefei University of Technology, Hefei, Anhui 230009, China



The architecture of our proposed multi-path recursive convolutional network

Overview: Single image super-resolution is widely used in security monitoring, satellite remote sensing imagery, and medical image processing. It aims at restoring a high-resolution image from corresponding degraded low resolution LR-image. Recently, Dong et al. first discovered that convolutional neural networks can accomplish super-resolution by end-to-end manner, opening the door for deep learning in the field of super-resolution. And a series of new network model were proposed. Although these models have achieved good performance, the existing problems cannot be ignored. First, with the increase of network depth, many models fail to take into account the effect of hierarchical features on super-resolution, and the extracted features of each layer can only be learned once, which cannot be fully utilized. Second, the many models use pre-processing methods to get the target size, which not only increases the computational complexity, but also destroys the information carried by the original image. In response to this problem, ESPCN based on subpixel convolution and FSRCNN based on deconvolution are proposed. However, their structure is too simple to complete the exact mapping. Third, most methods use the mean square error (MSE) to optimize the model, resulting in overly smooth images.

To solve these problems, we propose a multipath recursive network (MRCN). We use multi-path structure to extract features and improve the ability of non-linear mapping, which accelerates the transfer of feature and gradient in the network. Then we use recursive methods to reduce network parameters. Finally, all the features were merged to complete super-resolution. Compared with other models, our network mainly has the following differences. First, different from the traditional single-chain structure, our network adopts a multi-path structure, which enables the extracted features of each layer to be learned multiple times, improving feature richness, and the reconstructed image contains more high-frequency information. Second, most models use the last layer of the network to complete reconstruction, while our network uses all the features extracted from the network to complete reconstruction together. At the same time, we use the nature of SENet to select the effective features of these features adaptively and suppress the useless features. Third, we use the Charbonnier loss function to alleviate the problem that the reconstructed images are too smooth due to MSE, and the performance of the network can be improved. A large number of experiments on the benchmark set show that our method is superior to the existing methods in reconstruction performance.

Citation: Shen M Y, Yu P F, Wang R G, *et al.* Image super-resolution via multi-path recursive convolutional network[J]. *Opto-Electronic Engineering*, 2019, 46(11): 180489

Supported by National Natural Science Foundation of China (61672202)

* E-mail: yangjuan6985@163.com