

Super-Resolution Reconstruction Algorithm for Cross-Module Based on Depth-Wise Separable Convolution

Lijuan Shang, Zilu Ying, Ying Xu*

School of Information Engineering, Wuyi University, Jiangmen Guangdong
Email: xuying117@163.com

Received: Apr. 3rd, 2018; accepted: Apr. 20th, 2018; published: Apr. 27th, 2018

Abstract

The deep convolutional neural network has achieved remarkable results in single-image super-resolution reconstruction. Most models use a single-flow structure, and it is difficult to the full flow of information. The cross-module is used to improve the information flow to obtain enough detailed information. Although this method can improve the network reconstruction performance, it increases parameters and difficulty of network training. Therefore, this paper proposes the cross module base on depth-wise separable convolutional (CM-DWSC) algorithm for super-resolution reconstruction. Using depth separable convolution instead of ordinary convolution in the cross module, and the BN layer that consumes the same amount of memory as the convolutional layer is removed, greatly reduces the computational complexity and model capacity. A series of improved cross-modules are stacked in each cascade sub-network to fuse complementary information, which facilitates the flow of information. We introduce a residual learning strategy at each stage to utilize fully low-resolution feature information to further improve the reconstruction performance. The evaluation of the benchmark data set shows that the proposed method is superior to the mainstream super-resolution method in the case of reducing network parameters.

Keywords

Deep Convolution Neural Network, Super-Resolution Reconstruction, Depth-Wise Separable Convolution, Improved Cross-Module

基于深度可分离卷积的交叉模块的超分辨率重构算法

商丽娟, 应自炉, 徐颖*

*通讯作者。

五邑大学信息工程学院, 广东 江门
Email: xuying117@163.com

收稿日期: 2018年4月3日; 录用日期: 2018年4月20日; 发布日期: 2018年4月27日

摘要

深度卷积神经网络进行单图像超分辨率重构方面已取得显著成果。大多数模型采用单流结构, 不利于信息充分流动。采用交叉模块进行改善信息流动以获取足够的细节信息, 然该方法虽然可以提高网络重构性能, 但增加大量的网络训练参数, 增大了网络训练难度, 因此, 本文提出了基于深度可分离网络的交叉模块(Cross-Module Based on Depth-Wise Separable Convolution, CM-DWSC)超分辨率重构算法, 在交叉模块中使用深度可分离卷积代替普通卷积, 同时, 去掉与卷积层消耗等量内存的BN层, 极大降低计算复杂度及模型容量。在每个级联子网络中堆叠多个改进交叉模块以便融合互补信息, 有利于信息间流动, 在每个阶段引入残差学习策略, 充分利用低分辨率特征信息, 进一步提升重建性能。在基准数据集的评估表明, 本文方法能够在减少网络参数情况下优于主流的超分辨率方法。

关键词

深度卷积神经网络, 超分辨率重构, 深度可分离卷积, 改进交叉模块

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目前, 基于卷积神经网络(Convolution Neural Network, CNN)模型在单图像超分辨(Single image super-resolution, SISR)方面已经取得了巨大的成功。大多数 CNN 模型通过训练大量参数而学习一种有效的从低分辨率(Low-Resolution, LR)到高分辨率(High-Resolution, HR)图像的非线性映射关系。超分辨率(Super-Resolution, SR)的端到端学习架构去掉了繁琐的特征提取前的图像预处理过程和后续重叠的 HR 图像聚合过程。利用深度网络模型强大学习能力, 能够有效解决 SR 这一病态逆问题[1] [2]。2010年, Yang [3]等根据稀疏编码解决退化问题的有效性, 提出稀疏编码的 SRR 算法。该算法通过学习高、低分辨率图像块相同先验信息来重建 HR 图像。受 Yang 算法启发, Dong 等[4]最早提出利用卷积神经网络对图像进行超分辨率重构(Super Resolution Using Convolutional Neural Networks, SRCNN), 该网络利用卷积神经网络学习低分辨率到高分辨率的端到端映射关系, 重构超分辨率图像。随后, Dong 等[5]在 SRCNN 方法基础上提出一种快速超分辨率(Fast Super-Resolution by Convolutional Neural Networks, FSRCNN)方法, 该方法以低分辨率图像作为网络输入, 通过减小特征图维度和卷积核尺寸, 减少网络训练参数, 提高了网络运算速度。后来, Shi 等[6]提出一种利用亚像素卷积层的超分辨率(Super-Resolution Using Efficient Sub-Pixel CNN, ESPCN)重构方法, 该方法在最后一层重新对特征图进行排列, 得到高分辨率图像, 大幅提高了网络运算速度。以上 CNN 模式的一个共同之处在于它们的网络层数少于 5, 实验表明, 这些深层结构在 4 或 5 层并没有达到更好的效果, 训练更深层网络的难度导致“越深越好”思想在 SR 中实现遇到重重阻碍。2014 年生成对抗网络模型[7]被提出, 该网络模型包括生成模型和判别模型。Ledig [8]将生

成对抗网络应用到图像超分辨率重建上。由于是由噪声生成的图片，训练过程中存在不稳定性，得到的图片可能会和期望图片相差较大。同时受到 ImageNet 中利用非常深网络成功的启发[9] [10] [11]，Kim 等 [12] [13]提出两个非常深的卷积神经网络，分别是极深网络超分辨率重构(Super-Resolution Using Very Deep Convolutional Networks, VDSR)和递归网络(Deeply-Recursive Convolutional Network, DRCN)超分辨率重构。这两个网络模型均引入残差思想，使深度网络能够应用于 SR 领域。而随着深度加深，网络的感受野变得越来越宽，使网络不仅仅具备局部特征，也能捕捉到更加全局的特征。近年，毛等[14]提出一个 30 层卷积自动编码器网络命名为 RED30 用于图像恢复，它使用对称跳过连接来帮助训练。胡等[15]提出单图像级联交叉网络进行超分辨率重构，该网络通过将不同阶段子网络所生成图像按照不同比例融合重构超分辨率图像，以上几个模型优于之前介绍的深度学习方法和非 DL 方法，在 SR 中表现出“越深越好”的思想。但深度网络模型结构复杂，参数量巨大，模型容量大，计算复杂，网络训练难度加大。为了进一步提高网络性能，本文提出基于深度可分离卷积的交叉模块进行超分辨率重构，在每个子网络中，堆叠多个可分离卷积交叉模块以融合互补信息从而有效改善跨层的信息流，同时，又达到降低网络计算复杂度的问题，提高网络训练效率。

2. 理论基础

2.1. 卷积神经网络

典型 CNN 是一个多层的神经网络，主要由输入层、卷积层、池化层、全连接层和输出层组成，其结构如图 1 所示。

CNN 中一个典型层包含三级运算，在第一级运算中，多个卷积核并行计算产生一组线性的激活响应，第二级运算是每一个线性激活响应进行非线性激活函数处理，增加网络非线性性。对于第三级运算，利用池化层来调整某一层输出。

2.2. 残差神经网络

在堆叠层上采用残差学习策略，构建残差块定义如下：

$$f(x) = F(x, \{W_i\}) + x \tag{1}$$

其中 x 表示输入， $f(x)$ 表示输出。函数 $F(x, \{W_i\})$ 表示学到的残差映射。一个构建块如图 2 所示。

图中一个残差块包含 2 层， $F = W_2\sigma(W_1x)$ ，其中 σ 表示 ReLU， $f(x)+x$ 操作表示输入数据通过跳跃连接与映射进行逐像素加法，然后将获取的结果经过非线性激活函数处理。图中跳跃连接过程中并未添加额外参数量以及计算复杂度。残差函数 F 形势是灵活多变的，可以两层或者三层，甚至是更多层，

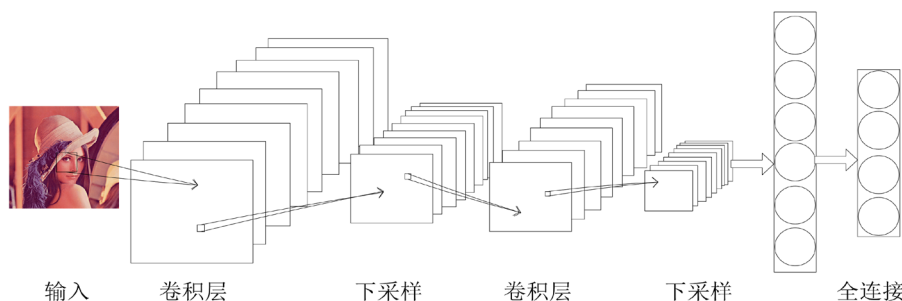


Figure 1. Diagram of convolutional neural network architecture
图 1. 卷积神经网络结构示意图

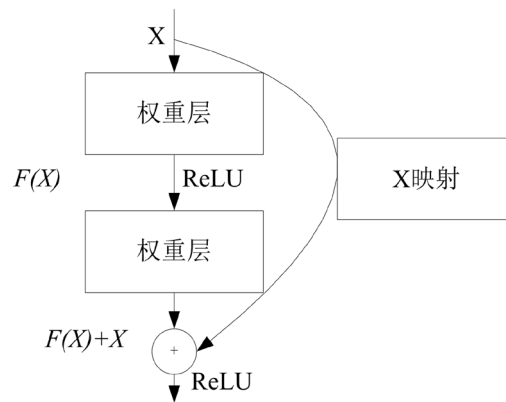


Figure 2. Residual block model
图 2. 残差块模型

若只有一层，则就退化为简单的线性函数： $y = w_i x + x$ ，不具有任何优势。同时，这种结构不仅可以应用到全连接层，就卷积层而言依然可以适用， $F(x, \{W_i\})$ 表示多个卷积层堆叠映射，在通道之间的两个特征图逐像素级进行求和。

3. 本文方法

3.1. 基于级联交叉网络超分辨率重构

基于级联多尺度交叉网络超分辨率重构方法[15]，在每个级联子网络中，堆叠多个多尺度交叉模块以融合互补多尺度的信息从而有效改善跨层的信息流。同时，在每个阶段引入残差特征学习，进一步提升重构性能。整个网络结构如图3所示。

该网络包括特征提取、子网络特征学习及图像重构三部分，第一部分，使用一层卷积层从初始图像块中提取图像底层特征，得到特征图 F_l 可以表示为：

$$F_l = \max(\alpha, W_l * Y + B_l) \quad (2)$$

其中 Y 表示高分辨率图像， $*$ 表示卷积操作， B_l 表示偏置， W_l 表示卷积核。第二部分是由一系列多尺度交叉模块级联成的子网络结构，子网络有多个并行映射来整合不同分支信息，同时，创造新的映射传递中间分支信息到后面重构模块。第三部分是利用卷积层重构来自不同子网络的高分辨率图像块，将获得高分辨率图像块按所学习的权重参数进行融合，生成最终的高分辨率图像。

3.2. 基于深度可分离卷积的交叉模块

3.2.1. 深度可分离卷积

受残差网络瓶颈结构[16]启发，本文方法采用深度可分离卷积[17]代替残差结构中的卷积，深度可分离卷积操作与分组卷积和 Google 提出的 inception 系列具有极大相关性，其是在普通卷积的基础上引用了分组卷积思想，并针对在不同分组间的输入特征图进行相互独立卷积操作。将普通卷积分解为一个逐深度卷积和一个标准 1×1 卷积，逐深度卷积对应着每一个输入特征图通道， 1×1 标准卷积是将逐深度卷积所提取特征进行融合，运用深度可分离卷积比普通卷积减少了所需要的网络参数，普通卷积同时考虑通道和区域，而可分离卷积先只考虑区域，再考虑通道融合，实现了通道和卷积区域的分离。可分离卷积结构示意图如图4所示。

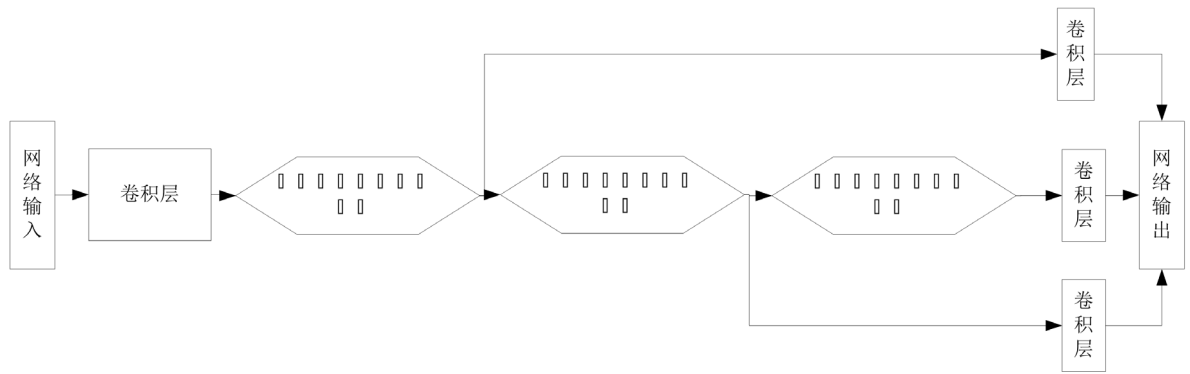


Figure 3. Cascade network structure
图 3. 级联网络结构

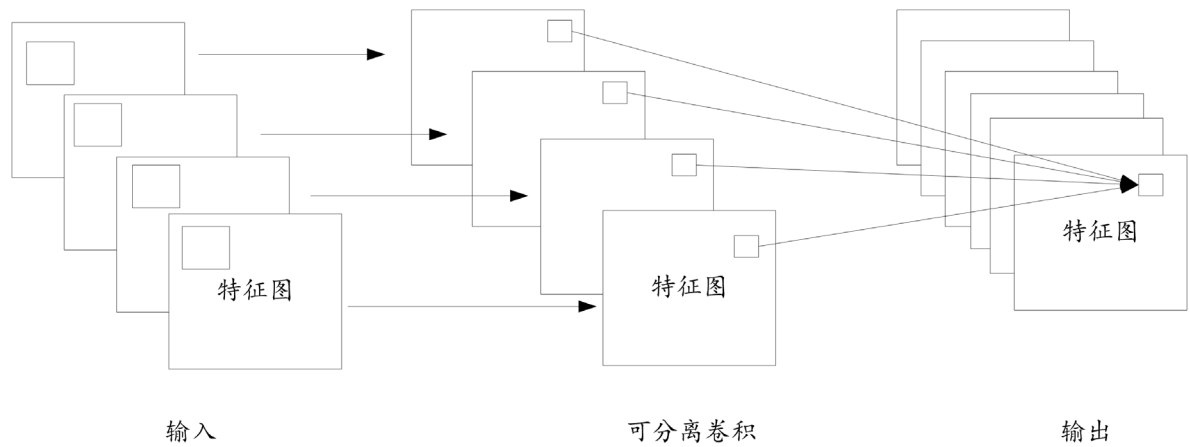


Figure 4. Diagram of depth-wise separable convolution
图 4. 深度可分离卷积示意图

假设一个普通卷积层的输入特征图为 X , 大小为 $C \times H \times W$, 其中 C 表示输入特征图通道数, H 和 W 分别是输入特征图的高和宽; 经大小为 $C \times K \times K \times N$ 卷积核 F 卷积操作后输出特征图为 Y , 大小为 $N \times H' \times W'$, 其中 K 是卷积核的大小, N 是输出特征图通道数, H' 和 W' 是输出特征图的高和宽。普通卷积操作可以表示为:

$$Y_{n,h,w} = \sum_{i,j,c} F_{c,i,j,n} * X_{c,h'+i-1,w'+j-1} \quad (3)$$

其参数量和计算复杂度分别为:

$$C \times K \times K \times N \quad (4)$$

$$C \times K \times K \times N \times H' \times W' \quad (5)$$

深度可分离卷积中, 区域卷积和通道是可分离的, 每个卷积核对应只一个输入特征图, 深度可分离卷积操作可以表示为:

$$Y_{n,h,w} = \sum_{i,j} F_{c,i,j} \odot X_{c,h'+i-1,w'+j-1} \quad (6)$$

\odot 表示可分离卷积操作符, 逐深度卷积后, 在深度可分离卷积后面加入一个输出特征图为 N 的 1×1 普通卷积, 其参数量和计算复杂度分别为:

$$C \times K \times K + C \times 1 \times 1 \times N \quad (7)$$

$$C \times K \times K \times H' \times W' + C \times H' \times W' \times N \quad (8)$$

利用公式(4)和公式(7)可计算在参数量上,深度可分离卷积是普通卷积的 $\frac{1}{N} + \frac{1}{K^2}$ 倍,同理可得,在计算复杂度上,深度可分离卷积是普通卷积的 $\frac{1}{N} + \frac{1}{K^2}$ 倍。当卷积核尺寸较大时,可以充分体现深度可分离卷积在模型优化方面的优势。

采用普通卷积操作时,针对不同卷积核提取不同的特征,卷积核与特征图像卷积后输出一个特征图,而如果选用可分离卷积,就可以将输入特征图与输出特征图通过卷积操作进行一对一连接。图4可以看出,每一个通道用一个卷积核进行卷积之后,得到对应通道输出,然后再利用 1×1 卷积核进行特征融合,通过特征提取和特征融合分离可以有效降低计算复杂度和模型容量。

3.2.2. 改进交叉模块

残差网络通过恒等映射跳跃连接模式使深度网络在图像超分辨率重构领域可以实现,考虑到不同的感受域的卷积网络的可以提供不同的上下文信息,这对SR来说非常重要的。采用交叉模块(Cross Module, CM)来充分获取上下文中有用信息,交叉模块结构如图5(a)所示。

在CM中有两个残差模块进行并行处理,在残差分支上,对输入进行平均处理,将处理结果传输到各自分支去。每个残差分支在CM模块中包含两个卷积层,每个卷积层之后采用ReLU激活函数。网络是由一系列交叉模块堆叠而成,为了减少网络计算复杂度及网络训练参数,本文选用可分离卷积代替普通卷积,可分离卷积相较普通卷积,极大降低了网络的计算复杂度,减少了网络参数量。改进后交叉模块如图5(b)所示,由于BN层[18]会致使部分细节信息丢失有价值信息,和消耗与卷积层等量内存,致使训练速度减慢,所以,本文不引入BN层,利用这种设计,这些分支结构可以提取到丰富的互补的背景信息,在多个尺度上进一步合并和运行组合映射。在每一个卷积后面,采用ReLU激活函数,为网络引入非线性因素,提高网络的学习能力。通过这种信息融合结构,CM可以利用各种各样的情境信息来推断缺失的高频信息。同时,CM有助于丰富信息流过不同的模块,也支持迭代训练期间的反向传播。

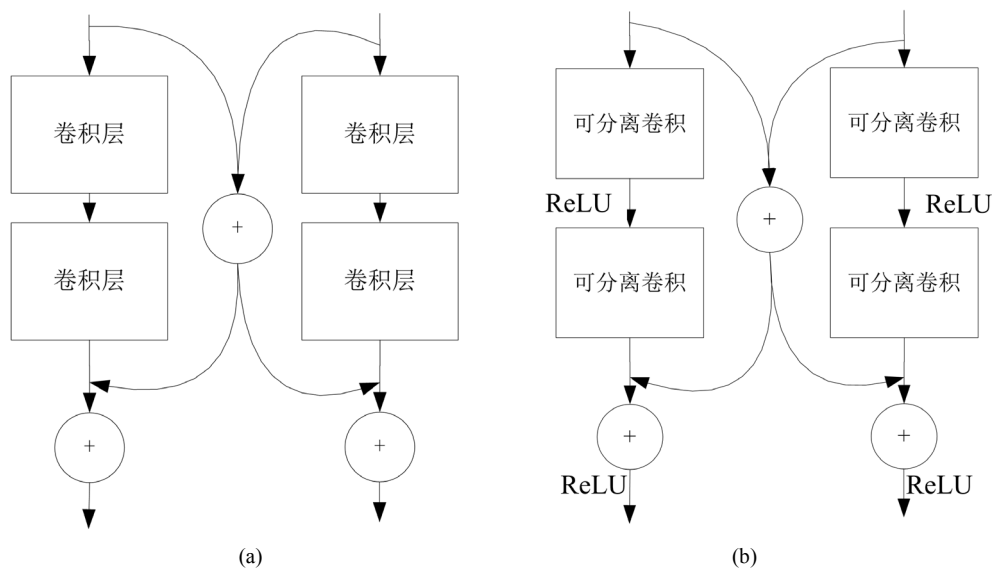


Figure 5. (a) Diagram of cross-module; (b) Diagram of improved cross-module
图5. (a) 交叉模块示意图; (b) 改进交叉模块示意图

4. 实验结果

文在 Timofte 数据集中进行超分辨率重构。为防止网络过拟合, 采用尺度缩放和旋转两种方式对训练图像进行增强。具体地, 对在数据集中每幅图像进行降采样, 缩放到原图像尺寸的 0.6、0.7、0.8、0.9 倍。然后将数据集中图像按顺时针分别旋转 90°、180°、270°。通过图像尺度缩放和旋转处理后, 原数据库扩增 20 倍。数据增强效果如图 6 所示。

为得到满足要求的低分辨率图像及增强网络对不同尺度超分辨率的泛化能力, 网络训练前, 本文对增强后数据集进行尺度变换预处理(处理过程如图 7 所示), 将不同尺度因子作用下的数据集混合在一起作为网络输入训练得到的模型, 不仅支持多种尺度因子图像超分辨率重构, 并且在多种尺度因子的情况下参数共享。

实验结果与分析

本文采用峰值信噪比(Peak Signal to Noise Ratio, PSNR)对重构算法进行评价, 假设输入图像表示为 $f(x,y)$, 其大小为 $M \times N$ 。 $g(x,y)$ 表示估计图像, 期望其尽可能逼近与输入图像 $f(x,y)$, 定义 MSE 为:

$$MSE = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [g(x,y) - f(x,y)]^2 \quad (9)$$

PSNR 表示为:

$$PSNR = 10 \times \log(255^2 / MSE) \text{ (dB)} \quad (10)$$

PSNR 值越大, 说明图像失真越小, 重构效果越好。

为测验所提图像 SR 方法有效性, 本文实验在 SR 领域中基准测试集 Set5、Set14、BSD100 上进行测试,



Figure 6. Data enhancement
图 6. 数据增强

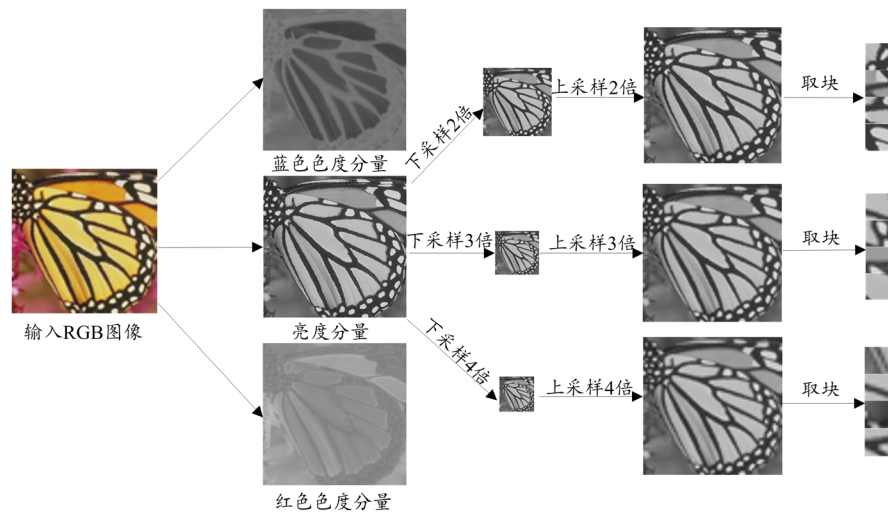


Figure 7. Image preprocess
图 7. 图像预处理

Table 1. Average PSNR comparison of super resolution algorithm on the Set5, Set14, B100 datasets
表 1. 超分辨率算法在 Set5、Set14、B100 数据集上 PSNR 平均值

数据集	尺度因子	双三次插值	SRCNN [4]	FSRCNN [5]	ESPCN [6]	VDSR [12]	本文方法
Set5	3	30.29	32.35	33.06	32.39	33.66	34.14
Set14	3	27.54	29.01	29.37	29.09	29.78	30.11
B100	3	27.21	28.21	28.41	28.26	28.83	29.02
Set5	4	28.42	30.15	30.55	30.27	31.53	31.94
Set14	4	26.00	27.44	27.50	27.39	28.02	28.35
B100	4	25.96	26.69	26.90	26.81	27.29	27.48

Table 2. Comparison of network computational complexity and network parameters in different convolution strategy
表 2. 不同卷积方式网络计算复杂度及网络参数比较

属性	本文方法	标准卷积	可分离卷积
计算复杂度		990 MFLOPS	125 MFLOPS
网络参数		2.11 G	0.27 G

并将测试结果与 Bicubic、SRCNN [4]、FSRCNN [5]、ESPCN [6]、VDSR [12]等主流方法进行对比。

从表 1 可看出，本文方法针对不同测试集进行超分辨率重构均能获得较好的实验效果，能够很好实现图像超分辨率重构，同时，本文网络能够适应不同缩放尺度，从而更好实现图像超分辨率重构。现有主流方法存在网络可重用性较差，需分别训练针对不同尺度因子网络，本文将输入网络图像进行不同尺度缩放处理，使网络具有能够很好学习不同缩放尺度的特性，具有可增强网络泛化能力的优点。并且，通过采用可分离卷积代替普通标准卷积，能够极大地减少了网络的计算复杂度及网络参数量，结果如表 2 所示，对比选用不同卷积方式时，网络的计算复杂度及网络参数的变化，可以看出，可分离卷积在优化网络模型中具有很重要的作用。综合所述，本文方法重构效果优于其他主流超分辨率重构算法。

5. 结论

本文深入研究了深度级联交叉模块的子网络超分辨率重建算法，级联一系列由交叉模块堆叠而成的

子网, 以级联监督形式逐步学习图像特征, 使其更接近本地真实图像特征。并针对该网络存在不足进行研究, 提出基于深度可分离卷积的交叉模块超分辨率重构算法, 在每个子网络中, 堆叠多个可分离卷积交叉模块以融合互补信息从而有效改善跨层的信息流。同时, 又达到降低网络计算复杂度的问题, 提高网络训练效率。虽然深度可分离网络能够减少网络模型及计算复杂度, 由于所有阶段子网都有相同的结构和相同的目标, 跨越级联模型有可能采用递归连接方式进行改善, 实现网络参数共享。在未来的工作中, 将探索如何共享级联内参数, 从而控制模型参数。另一方面, 将该模型扩展到其他图像退化领域。

基金项目

国家基金项目 No. 61372193, No. 61771347; 广东高校优秀青年教师培训计划资助项目 No. SYQ2014001; 广东省特色创新项目 No. 2015KTSCX143, 2015KTSCX145, 2015KTSCX148; 广东省青年创新项目 No. 2015KQNCX172, No. 2016KQNCX171; 江门市科技计划项目 No. 201501003001556, No. 201601003002191。

参考文献

- [1] 曾坤. 基于学习的单幅图像超分辨率重建的若干关键问题研究[D]: [博士学位论文]. 厦门: 厦门大学, 2015.
- [2] 李展. 空间超分辨率图像重建算法研究[D]: [博士学位论文]. 广州: 华南理工大学, 2012.
- [3] Yang, J., Wright, J., Huang, T.S., *et al.* (2010) Image Super-Resolution via Sparse Representation. *IEEE Transactions on Image Processing*, **19**, 2861-2873. <https://doi.org/10.1109/TIP.2010.2050625>
- [4] Dong, C., Chen, C.L., He, K., *et al.* (2014) Learning a Deep Convolutional Network for Image Super-Resolution. *Computer Vision—ECCV 2014*. Springer International Publishing, 184-199.
- [5] Dong, C., Chen, C.L. and Tang, X. (2016) Accelerating the Super-Resolution Convolutional Neural Network. *European Conference on Computer Vision*. Springer: Cham, 391-407.
- [6] Shi, W., Caballero, J., Huszar, F., *et al.* (2016) Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. 1874-1883.
- [7] Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., *et al.* (2014) Generative Adversarial Nets. *International Conference on Neural Information Processing Systems*, MIT Press, 2672-2680.
- [8] Ledig, C., Wang, Z., Shi, W., *et al.* (2016) Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, 21-26 July 2017, 105-114.
- [9] He, K., Zhang, X., Ren, S., *et al.* (2015) Deep Residual Learning for Image Recognition. *Computer Vision and Pattern Recognition*, **1**, 770-778.
- [10] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. *Computer Science*. arXiv:1409.1556 [cs.CV]
- [11] Szegedy, C., Liu, W., Jia, Y., *et al.* (2015) Going Deeper with Convolutions. *IEEE Conference on Computer Vision and Pattern Recognition*, Boston, 7-12 June 2015, 1-9.
- [12] Kim, J., Lee, J.K. and Lee, K.M. (2015) Accurate Image Super-Resolution Using Very Deep Convolutional Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 1646-1654.
- [13] Kim, J., Lee, J.K. and Lee, K.M. (2015) Deeply-Recursive Convolutional Network for Image Super-Resolution. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 1637-1645.
- [14] Mao, X.J., Shen, C. and Yang, Y.B. (2016) Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections. arXiv:1603.09056.
- [15] Hu, Y., Gao, X., Li, J., *et al.* (2018) Single Image Super-Resolution via Cascaded Multi-Scale Cross Network. arXiv: 1802.08808.
- [16] He, K., Zhang, X., Ren, S., *et al.* (2015) Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 770-778.
- [17] Chollet, F. (2017) Xception: Deep Learning with Depthwise Separable Convolutions. In: *IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, 1800-1807.
- [18] Ioffe, S. and Szegedy, C. (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *International Conference on Machine Learning*, 448-456.