

## 图像超分辨率卷积神经网络加速算法\*

刘超<sup>1,2</sup>, 张晓晖<sup>1</sup>, 胡清平<sup>1</sup>

(1. 海军工程大学 兵器工程学院, 湖北 武汉 430033;

2. 军事科学院 系统工程研究院, 北京 100044)

**摘要:** 为了实现模型的实时和嵌入式运行, 提出了一种轻量级的卷积神经网络结构。通过采用较小的滤波器尺寸和引入深度可分离卷积, 可大量减少模型参数, 提高模型非线性表达能力; 在网络末端引入子像素卷积层, 直接从原始低分辨率图像学习到高分辨率图像的映射, 计算成本为原来的  $1/k^2$  ( $k$  为放大因子)。在 Set5 数据集上的实验表明, 所提模型的速度较经典的图像超分辨率重建算法速度提高了 25.8 倍, 能够在通用 GPU 上实时运行, 峰值信噪比平均提高了 0.17 dB, 并且参数只有它的 35%。

**关键词:** 卷积神经网络; 超分辨率重建; 深度可分离卷积; 子像素卷积

中图分类号: TP391.4 文献标志码: A 文章编号: 1001-2486(2019)02-091-07

## Image super resolution convolution neural network acceleration algorithm

LIU Chao<sup>1,2</sup>, ZHANG Xiaohui<sup>1</sup>, HU Qingping<sup>1</sup>

(1. College of Weaponry Engineering, Naval University of Engineering, Wuhan 430033, China;

2. System Engineering Research Institute, Academy of Military Science, Beijing 100044, China)

**Abstract:** In order to realize the real-time and embedded operation of the model, a lightweight convolution neural network structure was proposed. By using a smaller filter size and introducing depthwise separable convolution, a large number of model parameters can be subtracted and the nonlinear capability can be improved. The sub-pixel convolution was introduced at the end of the network, then the mapping was learned directly from the original low-resolution image (without interpolation) to the high-resolution one, the cost is  $1/k^2$  as much as before ( $k$  is the magnification factor). Experimental results on Set5 show that the proposed model is more than 25.8 times faster than the classical super resolution reconstruction algorithm and can run in real-time on a common GPU; and the proposed method with only 35% parameters of SRCNN (super resolution convolution neural network), improves the PSNR (peak signal to noise ratio) with value of 0.17 dB.

**Keywords:** convolution neural network; super-resolution reconstruction; depthwise separable convolution; sub-pixel convolution

单帧图像超分辨率重建旨在从给定的低分辨率图像中重建出高分辨率图像, 该技术能够提供视觉效果更好的图像并提供更多的图像信息<sup>[1]</sup>。最近的超分辨率算法主要是基于深度学习的方法, 此类方法利用先验知识学习低分辨率和高分辨率图像之间的映射关系, 可以显著增强图像细节。2014年, 香港中文大学 Dong 等成功地将深度学习引入到图像超分辨率重建问题中, 提出基于卷积神经网络的图像超分辨率重建算法 (Super Resolution Convolution Neural Network, SRCNN) 取得了较好的效果<sup>[2]</sup>。

随后, 基于深度学习的方法被大量改进, 2016年, Kim 等针对 SRCNN 方法感受野较小的问题,

提出使用更多的卷积层来增加网络感受野, 同时为了减少网络参数, 引入递归神经网络, 提出的深度递归卷积网络 (Deeply-Recursive Convolutional Network, DRCN) 方法得到的效果比 SRCNN 有了较大提高<sup>[3]</sup>。在 SRCNN 和 DRCN 中, 低分辨率图像一般先通过上采样插值扩充到与高分辨率图像同样大小, 再作为网络输入, 这就意味着卷积操作是在较高的分辨率上进行, 相比于在低分辨率图像上计算, 会降低效率, 于是 SHI 等提出了一种在低分辨率图像上直接计算卷积得到高分辨率图像的高效亚像素卷积网络 (Efficient Sub-Pixel Convolutional Neural network, ESPCN) 方法<sup>[4]</sup>。上述方法对于大尺度放大因子重建会存

\* 收稿日期: 2017-12-15

基金项目: 国家部委基金资助项目 (4010501050401)

作者简介: 刘超 (1989—), 男, 助理研究员, 博士, E-mail: generaladolph@163.com;

张晓晖 (通信作者), 女, 教授, 博士, 博士生导师, E-mail: 1547300454@qq.com

在图像模糊问题, Ledig 等将生成对抗网络引入超分辨率重建中,提出了一种基于生成对抗网络的超分辨率重建超分辨率重建生成对抗网络 (Super-Resolution Generative Adversarial Network, SRGAN) 方法,该方法能够提高大尺度放大因子的重建质量<sup>[5]</sup>。为了提高运行速度,2017 年, Dong 等通过引入  $1 \times 1$  卷积和转置卷积对经典的 SRCNN 进行了改进,提出了一种快速超分辨率卷积神经网络 (Fast Super-Resolution Convolutional Neural Network, FSRCNN) 模型,取得了很好的加速和复原效果,但模型参数冗余,难以进行嵌入式部署<sup>[6]</sup>。2017 年,武汉大学肖进胜等通过分析 SRCNN 卷积核尺寸和数量,提出了一种改进的 SRCNN,重建效果得到了一定的提高,但模型参数大量增加<sup>[7]</sup>。

虽然目前这些模型取得了较好的效果,但由于模型参数冗余,计算量依旧很大,难以实时运行和进行嵌入式部署。

通过深入分析当前超分辨率重建网络结构,发现限制其运行速度主要有以下三个方面:首先,作为预处理步骤,大部分基于深度学习的方法需要使用双线性插值将原始低分辨率图像上采样到期望的大小,以形成输入。因此,模型的计算复杂度随着高分辨率图像(而不是原始低分辨率图像)的空间尺寸二次增长。对于放大因子  $n$ ,插值的低分辨率图像的卷积计算成本是原始低分辨率图像计算成本的  $n^2$  倍。如果直接从原始低分辨率图像学习网络,则计算速度快  $n^2$  倍。其次,目前,模型参数主要来源于卷积操作,标准卷积需要同时进行特征提取和特征融合的工作,从参数利用角度来讲,效率低且效果不佳。相反,2017 年,谷歌提出的深度可分离卷积 (depthwise separable convolution) 把两步分离开来,从深度方向把不同的通道之间相互独立开,先进行特征提取,再进行特征融合,这种方式可以充分利用模型参数进行表示学习,在使用更少参数的情况下取得更好的

效果<sup>[8]</sup>。因此,利用深度可分离卷积替换标准卷积能够降低模型计算量,同时提升模型表达能力。再次,大的卷积核会导致模型参数和计算量的暴增(参数数量随着卷积核尺寸二次增长),但是为了获取更大的感受野,部分模型采用大尺寸卷积核<sup>[9]</sup>。然而多个  $3 \times 3$  卷积核级联也可以获得与对应大尺寸卷积核同样大小的有效感受野,并且包含了更多的非线性,同时减少了网络参数(如,利用 2 个  $3 \times 3$  卷积核的组合比 1 个  $5 \times 5$  卷积核的效果更佳,同时参数量  $(3 \times 3 \times 2 + 1 < 5 \times 5 \times 1 + 1)$  被降低)。

根据上述分析,本文设计了一种简洁高效的网络结构。

### 1 基于卷积神经网络的图像快速超分辨率重建模型

针对当前超分辨率模型参数较多、计算量较大等问题,本文设计了一种轻量级的超分辨率卷积神经网络结构 (Light Super-Resolution Convolutional Neural Network, LSRCNN),如图 1 所示。

整个重建模型主要包含特征提取、非线性映射和重建三层网络单元结构:特征提取网络单元结构紧跟在输入层后面,利用常规卷积提取大量低分辨率特征;非线性映射网络单元结构采用深度可分离卷积进行非线性映射,并引入  $1 \times 1$  卷积进行维度压缩和扩展;重建网络单元结构通过子像素卷积层进行上采样操作。为了便于理解,将常规卷积层表示为  $Conv(f_i, n_i, c_i)$ ,深度可分离卷积层表示为  $SepConv(f_i, n_i, c_i)$ ,子像素卷积层表示为  $SubConv(f_i, n_i, c_i)$ ,其中变量  $f_i, n_i, c_i$  分别表示卷积核的大小、数量的和输入通道数量(当前层的输入通道数量和上一层的卷积核数量表示同一量,即  $c_i = n_{i-1}, i$  为层数)。

由于整个网络包含几十个变量,不可能对其进行逐一研究,因此本文提前给不敏感变量分配

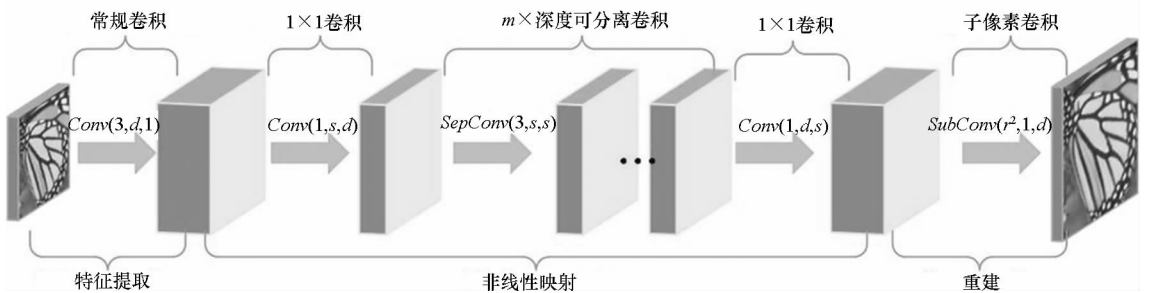


图 1 基于卷积神经网络的图像快速超分辨率重建模型

Fig. 1 Image fast super-resolution reconstruction model based on convolution neural network

一个合理的值,并保持敏感变量(当变量的轻微变化可能显著影响性能时,称为变量敏感)未设置。下面将对不同网络单元结构进行详细设计,并确定对应结构的敏感变量和非敏感变量。

### 1.1 特征提取网络单元

大部分基于神经网络的超分辨率模型的特征提取网络单元采用1个或多个卷积核进行低分辨率特征提取。与之相似,本文的特征提取网络单元采用1个常规卷积进行低分辨率特征提取,但在输入图像层面有所不同的是,LSRCNN在不进行插值的情况下对原始低分辨率图像进行特征提取。设低分辨率输入表示为 $Y$ ,通过与第一组卷积核进行卷积,将每个输入图像块表示为高维特征向量。由前面分析可知,卷积核尺寸的增加会引起参数的二次增长,由于 $3 \times 3$ 已经能够足够覆盖图像特征,因此设置 $f_1 = 3$ ;对于输入通道数量,根据文献[2]的分析,在YCBCR空间中,利用亮度通道进行模型训练能够取得更好的效果,因此设置 $c_1 = 1$ ;低分辨率图像的特征向量(底层特征)对最终结果很关键,由于卷积核数量 $n_1$ 表示了低分辨率特征通道数量,可将其定义为第一个敏感变量 $d$ ,因此,第一层可以表示为 $Conv(3, d, 1)$ ,图2为特征提取单元输出的低分辨率特征。

### 1.2 非线性映射网络单元

非线性映射是影响超分辨率重建性能最重要的部分,一般紧跟在特征提取步骤后面,用于将高维低分辨率特征直接映射到高分辨率特征空间。然而,由于低分辨率特征维度 $d$ 通常非常大,映射步骤的计算复杂度相当高,因此减少非线性映射网络单元模型参数成为压缩超分辨率模型的

关键。

谷歌的Inception结构通过应用 $1 \times 1$ 层来节省计算成本<sup>[10]</sup>,基于同样的考虑,本文在进行特征映射前,通过引入 $1 \times 1$ 层来减少低分辨率特征维度 $d$ 。因此设置卷积核大小 $f_2 = 1$ ,通过采用较少的卷积核数量 $n_2 = s \ll d$ 使低分辨率特征尺寸从 $d$ 减小到 $s$ 。其中, $s$ 是确定收缩水平的第二个敏感变量,因此第二层可以表示为 $Conv(1, s, d)$ 。通过该策略可以大量减少参数数量。

由上述分析可知,相对于常规卷积,深度可分离卷积能够在减少参数的同时,提高映射能力。为了充分利用其优势,本文非线性映射单元除了用于通道压缩和膨胀的 $1 \times 1$ 卷积,其余均采用深度可分离卷积。通过该策略,可以大大减少参数数量,同时增加网络的非线性能力。作为性能和网络规模之间的折中,本文采用中等滤波器尺寸 $f_3 = 3$ ,为了保持网络性能不下降,使用多个 $3 \times 3$ 层来替代单个宽度。因此,映射层数量 $m$ 是另一个敏感变量,这里所有的映射层采用相同的滤波器数量 $n_3 = s$ 。那么非线性映射部分可以表示为 $m \times SepConv(3, s, s)$ 。

由于映射图层的卷积核数量 $n_3 = s$ 比较小,如果直接从这些低维特征生成高分辨率图像,则最终的恢复质量将会很差。因此,需要在映射部分之后添加一个扩展层,以扩展高分辨率的特征维度,为了保持与收缩层的一致性,本文还采用 $1 \times 1$ 卷积,其数量与低分辨率特征提取层的数量相同,则扩展层可表示为 $Conv(1, d, s)$ 。

### 1.3 重建网络单元

重建网络单元主要对非线性映射层获取的特征进行上采样和聚合,可采用双三次或双线性插



图2 特征提取单元的输出特征

Fig. 2 Output feature of the feature extraction unit

值、上池化 + 卷积方式和转置卷积等方式。但双三次或双线性插值内核固定,不能够很好地恢复细节,转置卷积虽然可以学习一组输入特征图的上采样内核,但学习过程中需要补零操作,导致图像边缘存在失真。2016 年,文献[4]提出了一种全新的上采样方式——子像素卷积层(sub-pixel convolution),该方式能够减弱转置卷积补零操作带来的影响,并减小运算量。因此,本文在网络末端引入子像素卷积层。

子像素卷积层的实质是,通过在常规卷积层后面添加相位移层改变图像尺寸,在利用其进行上采样的过程中,插值函数被隐含地包含在前面的卷积层中,可以被自动学习。对于放大因子为  $k=r$  的网络,子像素卷积层的卷积核数量应该设置为  $r^2$ ,然后将每个像素的  $r^2$  个通道重新排列成一个  $r \times r$  区域,对应高分辨率图像中的一个  $r \times r$  大小的子块,从而大小为  $r^2 \times H \times W$  的特征图像被重新排列成  $1 \times rH \times rW$  大小的高分辨率图像。设放大因子为  $r$ ,则  $n_5 = r^2$ ;子像素卷积的卷积部分在低分辨率图像上进行,因此可采用较小尺寸的卷积,这里为了尽量压缩模型参数,设置  $f_5 = 1$ ,那么子像素卷积层可以表示为  $SubConv(r^2, 1, d)$ 。

#### 1.4 网络总体结构和性能分析

将上述三部分连接起来,形成一个完整的 LSRCNN 网络  $Conv(3, d, 1) - Conv(1, s, d) - m \times SepConv(3, s, s) - Conv(1, d, s) - SubConv(r^2, 1, d)$ ,这里每个卷积后面都跟有 ReLU 激活函数。控制网络参数和速度的三个敏感变量分别为低分辨率特征维度  $d$ 、非线性映射单元深度可分离卷积数量  $s$  以及映射深度  $m$ 。为简单起见,本文将 LSRCNN 网络称为  $LSRCNN(d, s, m)$ 。参数数量(不考虑偏置的情况下)为:  $18d + 2sd + m(9s + s^2)$ ,设  $S_{LR}$  为低分辨率图像大小,则计算复杂度为:

$$O[18d + 2sd + m(9s + s^2)]S_{LR} \quad (1)$$

经典卷积神经网络算法 SRCNN 网络的参数数量(不考虑偏置的情况下)为  $f_1^2 n_1 + n_1 f_2^2 n_2 + n_2 f_3^2$ ,则,计算复杂度为:

$$O[f_1^2 n_1 + n_1 f_2^2 n_2 + n_2 f_3^2]S_{HR} \quad (2)$$

其中,  $\{f_i\}_{i=1}^3$  和  $\{n_i\}_{i=1}^3$  分别是三层滤波器尺寸和滤波器数量。 $S_{HR}$  是高分辨率图像的大小。可以观察到复杂度与高分辨率图像的大小成正比,中间层对网络参数贡献最大。

本文对  $LSRCNN(48, 16, 1)$  和 SRCNN 进行比较,通过上式计算可知,  $LSRCNN(48, 16, 1)$  的参

数量为 2800, SRCNN 参数量为 8032,其参数量仅有 SRCNN 的 35%;对于放大因子为 3 的网络,提高运算速度  $8032/2800 \times 3^2 = 25.8$  倍。值得注意的是,这种加速并不以降低性能为代价。相反,  $LSRCNN(48, 16, 1)$  性能在 Set5 数据集上为 32.53 dB,优于 SRCNN 的 0.17 dB。

#### 1.5 损失函数

本文采用均方误差(Mean Square Error, MSE)作为损失函数,优化目标表示为:

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^n \|F(Y^i; \theta) - X^i\|_2^2 \quad (3)$$

其中,  $Y^i$  和  $X^i$  是训练数据中的第  $i$  个低分辨率和高分辨率子图像对,  $F(Y^i; \theta)$  是具有参数  $\theta$  的  $Y^i$  的网络输出。网络通过最小化估计  $Y$  和真实高分辨率图像块  $G$  之间的误差找到最优参数  $\theta = [W_j, B_j]$ ,对所有参数使用标准反向传播的随机梯度下降进行优化。

## 2 实验结果与分析

### 2.1 实验细节

数据集:91 数据集具有良好的质量,被广泛用作基于学习的超分辨率重建方法的训练集<sup>[11]</sup>,在实验中本文采用 91 图像数据集作为训练数据集。为了更好地和其他模型结果进行比较,没有采用数据增强技术。使用 Set5<sup>[12]</sup> 和 Set14<sup>[13]</sup> 数据集作为测试数据集。

训练样本:为了生成训练数据,本文首先用所需的放大因子  $k$  对原始训练图像进行采样,以形成低分辨率图像(由于不同语言的 Resize 函数内部实现机制有所差异,为了和其他模型比较,这里的采样操作采用 MATLAB 中的 Resize 函数)。然后将低分辨率训练图像裁剪成具有步幅  $l$  的一组尺寸为  $f_{sub} \times f_{sub}$  的子图像。相应的高分辨率子图像(尺寸为  $kf_{sub} \times kf_{sub}$ )也从高分辨率图像中裁剪出来。这些低分辨率/高分辨率(Low Resolution/High Resolution, LR / HR)子图像对是基本训练数据。

对于填充问题,为了确保提取的低分辨率特征轮廓清晰,对特征提取层不进行填充。实验中发现非线性映射层是否填充对结果影响不大,这里为了保证映射层输入输出尺寸相同,采用零填充,由于卷积核尺寸均为 3,填充尺寸设置为  $P = 1$ ,最终网络输出尺寸为  $k(f_{sub} - 2) \times k(f_{sub} - 2)$ ,而不是  $kf_{sub} \times kf_{sub}$ 。因此,需要在高分辨率图像上裁剪  $2k$  边界像素。最后,对于典型的放大因子 3,本文将 LR / HR 子图像的大小设置为  $18 \times 18 /$

48 × 48。

训练参数:优化算法采用学习率可以自适应优化的 Adam 算法<sup>[14]</sup>,初始学习率  $lr = 0.001$ ;权重初始化采用 xavier 初始化方法<sup>[15]</sup>;Batchsize 设置为 32;周期数设置为 80。

训练环境:训练过程均采用 Keras 库实现<sup>[16]</sup>,训练平台采用 CPU 为 intel i7 4710MQ, GPU 为 NVIDIA GTX 860,内存为 16 G 的笔记本电脑。

### 2.2 不同敏感参数研究

为了测试 LSRCNN 结构的属性,本文设计了一组包含三个敏感变量的控制实验(即低分辨率特征维度  $d$ 、非线性映射单元深度可分离卷积数量  $s$  以及映射深度  $m$ )。这里分别设置  $d = 48, 64$ ;  $s = 16, 24$  和  $m = 1, 2$ , 因此总共进行了不同组合的 12 次实验。

这些实验在 Set5 数据集上的平均 PSNR 值如表 1 所示。本文分别在水平和垂直方向上分析实验结果,首先固定  $d$  和  $m$ ,分析  $s$  对模型性能的影响,可以看到,随着  $s$  尺寸的增大,网络具有更好的性能。这种趋势也可以从图 3(a)所示的收敛曲线观察到。但是当网络参数冗余, $s$  的增加并不能提高性能,甚至会引发过拟合(见表 1 中第四行  $s$  的变化规律)。其次固定  $s$ ,并检查  $d$  和  $m$  的影响,通常更好的结果需要更多的参数,但是更多的参数可能引起网络过拟合,并不能保证更好的结果(见表 1 中第 2 列的变化规律)。这种趋势也反映在图 3(b)中,可以看到网络最后几乎汇聚到一起。从所有的结果中,能够找到性能和参数之间的最佳平衡  $LSRCNN(48, 24, 2)$ ,在适度参数情况下,获得了最好的效果。由于选择的 batchsize 比较小,因此收敛曲线会存在比较明显的跳跃。

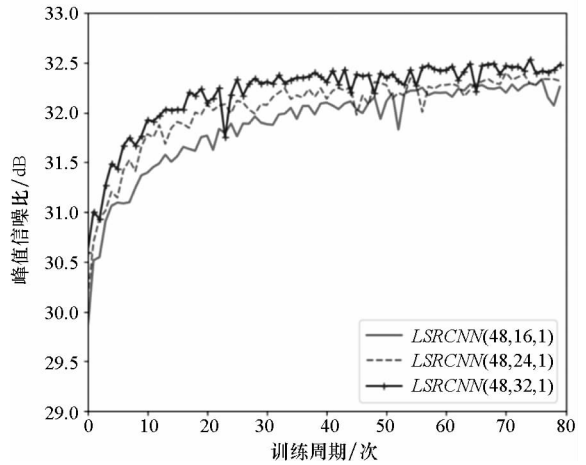
表 1 不同敏感参数在 Set5 测试集上的 PSNR 比较

Tab. 1 Comparison of PSNR for different sensitive parameters in Set5 test set

设置	$s = 16$	$s = 24$	$s = 32$
$d = 48, m = 1$	32.53	32.58	32.69
$d = 48, m = 2$	32.56	32.63	32.70
$d = 64, m = 1$	32.61	32.66	32.64
$d = 64, m = 2$	32.53	32.66	32.55

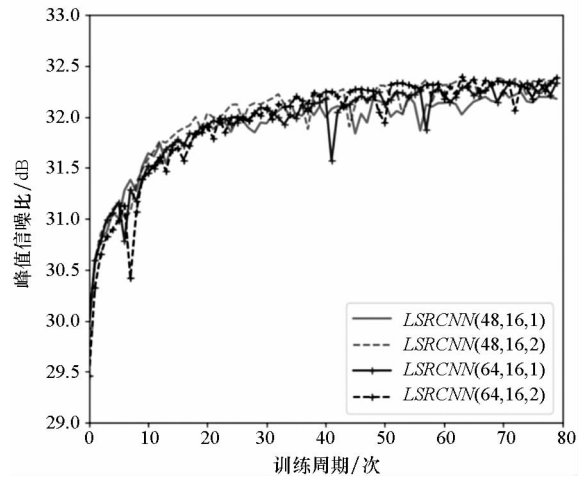
### 2.3 与其他方法比较

本文将提出的方法与传统的锚定邻域回归方法 A + <sup>[17]</sup> 以及三种基于外部数据库学习的超分



(a) 固定  $d = 48, m = 1$ , 改变  $s$

(a) Fixed  $d = 48, m = 1$ , change  $s$



(b) 固定  $s = 24$ , 改变  $d$  和  $m$

(b) Fixed  $s = 24$ , change  $d$  and  $m$

图 3 不同敏感参数的 PSNR 收敛曲线

Fig. 3 PSNR convergence curve of different sensitive parameters

辨率重建算法进行比较,即 SRCNN<sup>[2]</sup>、改进的 SRCNN<sup>[7]</sup>、FSRCNN-s<sup>[6]</sup>方法。A + 方法的实现基于它们发布的源代码,其余三种基于深度学习的方法根据文献内容在 Keras 框架上进行了编写,为了对恢复质量进行比较,所有模型都对没有增强的 91 数据集图像进行了训练,因此结果可能与对照论文略有不同。本文选择了两个代表性的 LSRCNN 模型  $LSRCNN(48, 16, 1)$  (简称为 LSRCNN-s) 和  $LSRCNN(48, 24, 2)$  (简称为 LSRCNN)。所有模型的运行时间通过 NVIDIA GTX 860 GPU 上的 Python 进行测试。表 2 列出了放大因子为 3 的定量结果(PSNR、模型参数和测试时间)。首先,分析测试时间,可以看到,提出的 LSRCNN 是最快的方法,比改进的 SRCNN 至少快 15 倍(在 GPU 上测试时间),LSRCNN-s 最快

可以达到 50 帧/s 的处理速度。其次,LSRCNN 也具有更少的参数,相比于动辄具有几万甚至几十万参数的模型,LSRCNN-s 仅有 2800 个参数,按照 32 bit 的精度进行存储,只需要 10 KB 空间,可以满足嵌入式部署。最后,LSRCNN 在具有更少参数的情况下,PSNR 依旧优于 A +、SRCNN、改进的 SRCNN 以及 FSRCNN 等算法。

图 4 和图 5 分别为 Set5 数据集和 Set14 数据集中典型图像在放大因子为 3 条件下不同模型的测试效果,可以看到本文算法相比其他算法在边缘和对比特度保持方面效果更好,与原图效果最接近,且取得了最好的 PSNR。

### 3 结论

本文针对当前基于深度学习超分辨率模型存

在参数量大、运行速度慢等问题,通过深入分析影响当前网络模型参数数量和运算性能的因素,结合深度模型压缩技术取得的最新成果,提出了一种更为高效的网络结构,通过减少卷积核尺寸、将子像素卷积层放置在网络末端进行上采样、引入  $1 \times 1$  卷积进行维度压缩以及采用最新的深度可分离卷积,可以大幅度降低网络参数,提高模型非线性表达能力,从而实现高速运行而不损失重建分离卷积,可以大幅度降低网络参数,提高模型非线性表达能力,能够实现高速运行,而不损失重建质量。实验结果表明,本文提出的方法在参数更少、运算速度更快的情况下,取得了更优的超分辨率效果。模型大小仅有 10 KB,可以达到 50 帧/s 的处理速度,能够部署到 FPGA 等嵌入式设备并实时运行。

表 2 不同模型在测试集上的 PSNR、参数数量和测试时间

Tab. 2 The PSNR, number of parameters and test time for different models on the test set

测试集	指标	BI	A +	SRCNN	改进的 SRCNN	FSRCNN	LSRCNN	LSRCNN-s
Set5	PSNR/dB	30.39	32.59	32.39	32.33	32.64	32.64	32.55
	参数数量	—	64 000	8032	38 016	12 464	4752	2800
	时间/s	—	0.4	0.21	0.38	0.14	0.03	0.02
Set14	PSNR/dB	27.54	29.13	29.00	29.05	28.97	29.10	29.04
	参数数量	—	64 000	8032	38 016	12 464	4752	2800
	时间/s	—	0.79	0.42	0.78	0.33	0.06	0.04

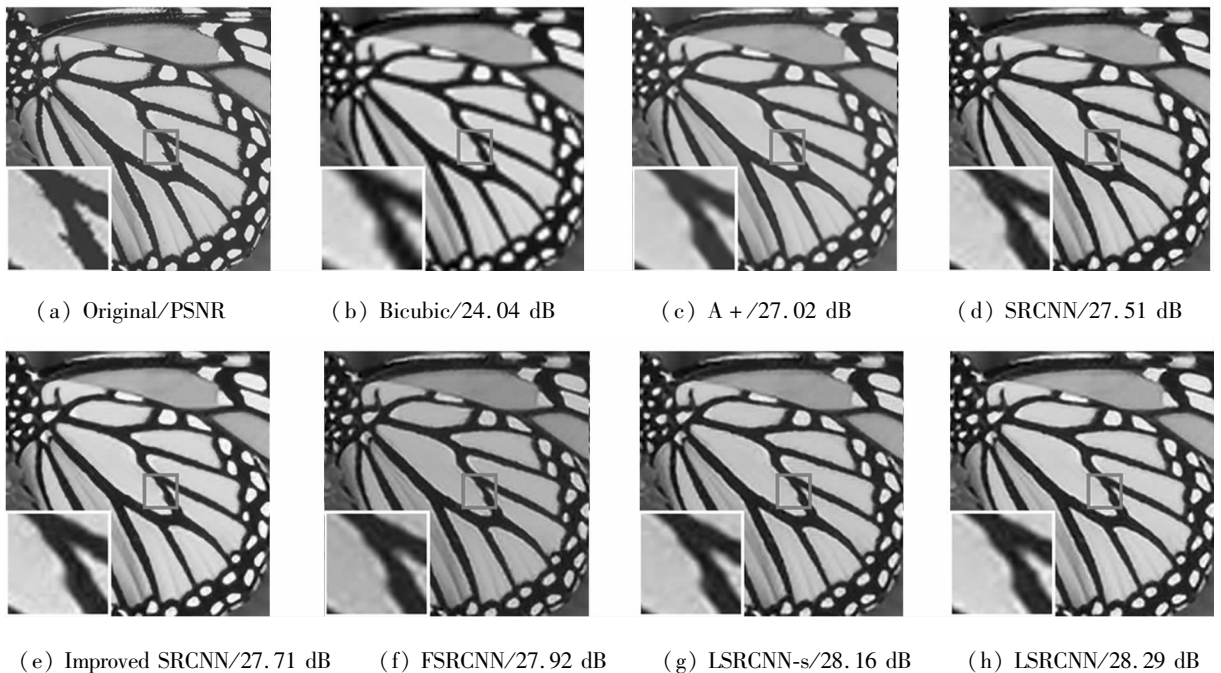


图 4 测试集 Set5 中的蝴蝶图像在放大因子为 3 条件下重建效果

Fig. 4 Reconstruction effect of butterfly image from the Set5 dataset with an upscaling factor 3

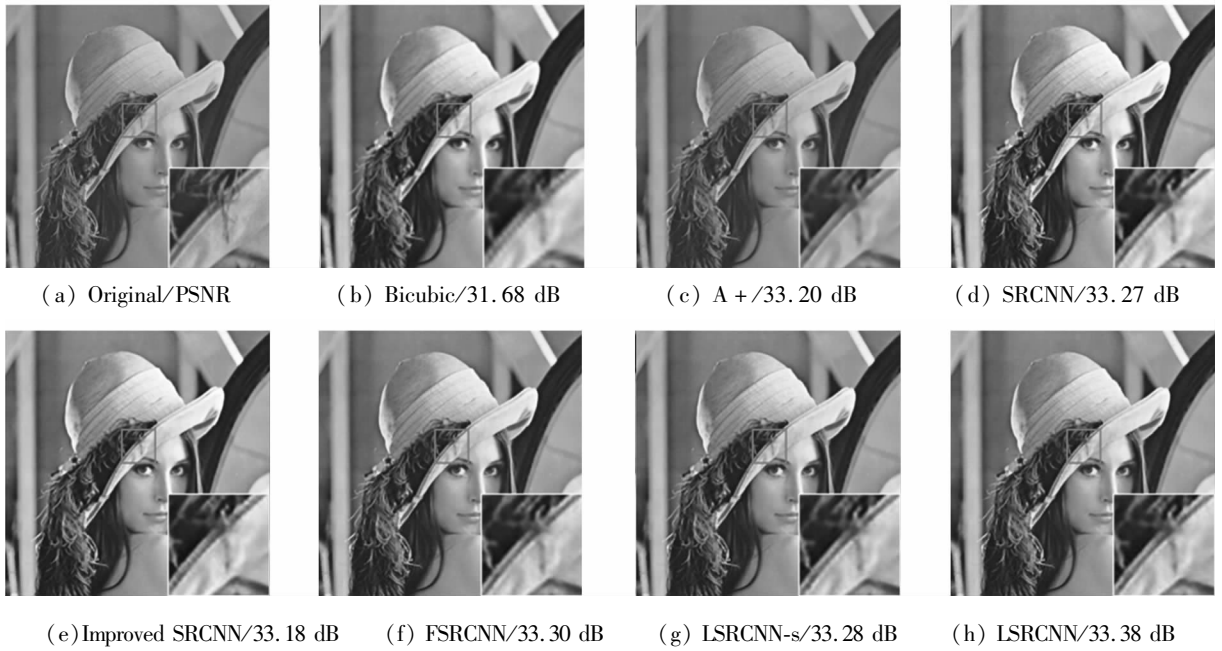


图 5 测试集 Set14 的 lena 图像的在放大因子为 3 条件下的重建效果

Fig. 5 Reconstruction effect of lena image from the Set14 dataset with an upscaling factor 3

## 参考文献 (References)

- [1] 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.  
LI Sumei, LEI Guoqing, FAN Ru. Depth map super-resolution based on deep convolutional neural networks[J]. Acta Optica Sinica, 2017, 37(12): 1210002. (in Chinese)
- [2] 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.
- [3] Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1637–1645.
- [4] Shi W Z, Caballero J, Huszar F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1874–1883.
- [5] Ledig C, Theis L, Huszar F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]. arXiv: 1609.04802v2, 2016.
- [6] Dong C, Loy C C, Tang X O. Accelerating the super-resolution convolutional neural network [C]. arXiv: 1608.00367v1, 2016.
- [7] 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): 0318011–1–0318011–9.  
XIAO Jinsheng, LIU Enyu, ZHU Li, et al. Improved image super-resolution based on convolutional neural network [J]. Acta Optica Sinica, 2017, 37(3): 0318011–1–0318011–9. (in Chinese)
- [8] Chollet F. Xception: deep learning with depthwise separable convolutions[C]. arXiv: 1610.02357v1[cs], 2016.
- [9] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [J]. arXiv: 1502.03167v3, 2015.
- [10] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2015.
- [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] Bevilacqua M, Roumy A, Guillemot C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding [C]//Proceedings of the 23rd British Machine Vision Conference (BMVC), 2012.
- [13] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations [C]//Proceedings of International Conference on Curves and Surfaces, 2010: 711–730.
- [14] Kingma D P, Ba J. Adam: a method for stochastic optimization[C]//Proceedings of 3rd International Conference for Learning Representations, 2015.
- [15] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks [J]. Journal of Machine Learning Research, 2010, 9: 249–256.
- [16] Chollet F. Deep learning library for python [M]. Shelter Island, Manning Publications Co., 2017.
- [17] Timofte R, de Smet V, van Gool L. A + : adjusted anchored neighborhood regression for fast super-resolution [C]//Proceedings of Asian Conference on Computer Vision, 2014: 111–126.