

网络预出版:

标题: 改进的基于卷积神经网络的图像超分辨率算法

作者: 肖进胜, 刘恩雨, 朱力, 雷俊锋

收稿日期: 2016-09-27

录用日期: 2016-11-30

DOI: 10.3788/aos201737.0318002

引用格式:

肖进胜, 刘恩雨, 朱力, 雷俊锋. 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(03): 0318002.

网络预出版文章内容与正式出版的有细微差别, 请以正式出版文件为准!

您感兴趣的其他相关论文:

基于深层卷积神经网络的单目红外图像深度估计

许路 赵海涛 孙韶媛

华东理工大学信息科学与工程学院, 上海 200237

光学学报, 2016, 36(7): 0715002

基于深度学习的高分辨率遥感影像分类研究

刘大伟 韩玲 韩晓勇

长安大学地测学院, 陕西 西安 710054

光学学报, 2016, 36(4): 0428001

改进的基于卷积神经网络的图像超分辨率算法

肖进胜^{1,2*} 刘恩雨¹ 朱力¹ 雷俊锋¹

¹ 武汉大学电子信息学院, 湖北 武汉 430072

² 地球空间信息技术协同创新中心, 湖北省武汉市珞喻路 129 号, 430079

摘要 针对现有的基于卷积神经网络的图像超分辨率算法参数较多, 计算量较大, 以及训练时间较长、图像纹理模糊等问题, 本文结合现有的一些优秀的图像分类网络模型以及视觉识别算法, 对基于卷积神经网络的图像超分辨率算法提出了改进。在原有的三层卷积神经网络中, 调整卷积核大小, 减少参数; 加入池化层, 降低维度, 减少计算复杂度; 提高学习率和输入子块的大小, 减少训练消耗的时间; 扩大图像训练库, 使训练库提供的特征更加广泛和全面。实验结果表明, 本文算法生成的网络模型取得了更佳的超分辨率结果, 主观视觉效果明显改善, 处理后图像清晰度和边缘锐度明显得到提高, 细节纹理信息更加丰富, 客观评价指标显著提升。

关键词 图像超分辨率; 深度学习; 卷积神经网络; 卷积核参数;

中图分类号 **文献标识码** A

Improved Image Super-Resolution Based on Convolutional Neural Network

Jinsheng Xiao^{1,2}, Enyu Liu¹, Li Zhu¹, Junfeng Lei¹

¹ *Electronic Information School, Wuhan University, Wuhan, Hubei, China, 430072*

² *Collaborative Innovation Center of Geospatial Technology, 129 Luoyu Road, Wuhan, China, 430079*

Abstract An improved image super-resolution based on the deep learning is proposed to overcome many problems such as a large amount of calculation, more parameters, longer training time and fuzzy texture. Combined with some excellent image classification network models and visual recognition algorithms, proposed algorithm adjusts the convolution kernel size to reduce parameters in the original three layers of convolutional neural network. Pool layers are added to reduce the dimension and decrease computational complexity. The learning rate and the size of the input sub-blocks are improved to reduce the training time. In order to provide more extensive and comprehensive characteristics, the training database is expanded. Experimental results show that the algorithm achieves better super-resolution results. The subjective visual effect and objective evaluation indices are improved obviously. The image resolution and edge sharpness are significantly enhanced.

Key words: image super-resolution; deep learning; convolutional neural network; parameters of convolution kernels

OCIS Codes: 100.6640, 200.4260

收稿日期: 年-月-日; 收到修改稿日期: 年-月-日

基金项目: 国家自然科学基金(61471272), 湖北省自然科学基金(2016CFB499)

作者简介: 肖进胜(1975-), 男, 博士, 副教授, 主要从事视频图像处理, 计算机视觉等方面的研究。E-mail: xiaojsh@whu.edu.cn

1 引言

图像超分辨率 (Super Resolution, SR) 是从一幅或者多幅成像质量较差的低分辨率 (Low Resolution, LR) 图像中获取高分辨率 (High Resolution, HR) 图像, 该技术能够提供视觉效果更好的图像, 提供更多的图像信息^[1-4]。主要分为四种类型: 基于插值^[5]、基于重建^[6]、基于增强^[7]和基于学习^[8]的方法。基于学习的图像超分辨率方法是近些年来发展比较火热的研究领域, 此类方法利用图像数据库或图像本身, 通过学习获得高、低分辨率图像之间的关联, 并利用其作为先验约束条件生成高分辨率图像。基于学习的方法可分为基于自学习和基于外部数据库的方法。

在基于自学习方面, Yang 的方法是利用了图像的自相似性学习^[9], 只使用单一输入的图像, 而不依赖于一个外在的训练图像, 该方法计算简单高效。Zhang^[10]结合框架结构调制和稀疏表示, 解决了依据稀疏性的超分辨率方法中忽略分别在相同和不同尺寸的图像结构特点的问题。Huang^[11]扩大了内部块的搜索空间, 引入仿射变换来适应图像中几何形状的变化, 但在自然风景图像处理效果与其他方法相比并无明显优势。Choi^[12]将低分辨率的图像中提取的块再分成多个子块, 以更好地提取低分辨率与高分辨率图像中的对应关系, 该方法在边缘处的处理效果更好。

在基于外部数据库方面, 2002 年, Freeman 提出基于样本学习 (Example-based) 的图像超分辨率方法, 通过在外部的图像数据库中使用最近邻域搜索的方式获得向量集^[13]。Timofte 结合稀疏学习字典和邻域嵌入的方法, 提出一种固定邻域回归的方法, 在保持图像质量的同时, 提高了算法的处理速度^[14]。Yang 利用稀疏表示, 使计算的复杂度大大降低, 并且对噪声的输入图像处理效果比较好^[15]。同时, 机器学习以及神经网络的应用日趋广泛, 尤其在视觉识别、图像分类以及图像超分辨率等方面。Dong^[16]将卷积神经网络用于图像超分辨率中, 将整个过程视为有三层卷积层的卷积神经网络, 将图像特征从低分辨率空间非线性地映射至高分辨率空间内, 重建质量更优。

本文将深度学习理论应用于图像超分辨率, 研究卷积神经网络用于图像超分辨率重建方法的优化。本文受 Dong^[16]卷积神经网络的图像超分辨率算法的启发, 并借鉴深度卷积网络的大规模视觉识别算法的网络模型, 修改神经网络结构, 加深了卷积网络, 加入池化层, 同时调整卷积核大小以及参数, 并对训练速度进行改善。理论分析和实验数据表明, 本文算法取得了更为优秀的超分辨率效果。

2 基于深度学习的图像超分辨率算法

基于深度学习^[17]的图像超分辨率方法依靠外部库,通过深度学习神经网络获得相应的先验信息,以实现图像的超分辨率,主要分为以下几个步骤:(1)建立训练所需的外部图像库。(2)构建网络模型。(3)对训练集数据进行训练,网络参数进行估计与优化,得到输入数据的特征表达及先验知识。(4)将低分辨率图像输入重建模型,输出高分辨率图像。

卷积神经网络不仅能有效较少网络的训练参数,使得神经网络简单化,而且具有很强的适应性。其中最具代表性的是香港中文大学 Dong 等人的超分辨率算法,简称为 SRCNN(Super-Resolution Convolutional Neural Network)^[16]。SRCNN 是由 3 层卷积层构成,分别为特征提取,非线性映射,高分辨率图像重构。其算法框架如图 1 所示。

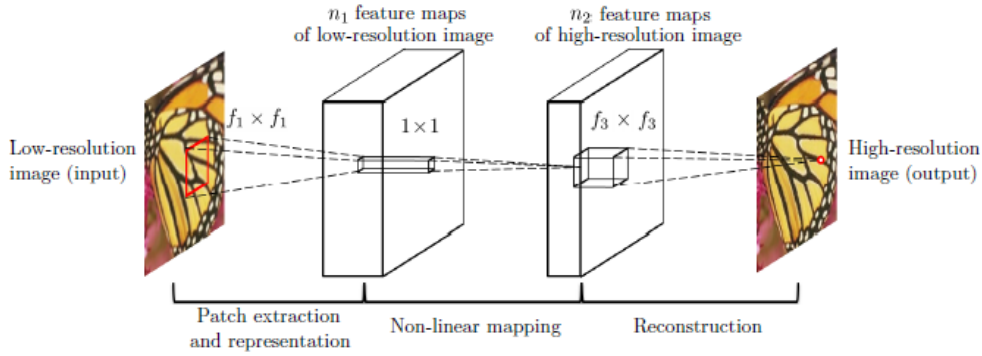


图 1 SRCNN 算法框架

Fig.1 SRCNN algorithm structure

特征提取是从初始图像中提取块,得到输入图像的特征图 $F_1(Y)$, 如公式 (1) 所示:

$$F_1(Y) = \max(0, W_1 * Y + B_1) \quad (1)$$

这里 Y 表示初始高分辨率图像, $*$ 指卷积符号, W_1 为卷积核, B_1 为神经元偏置向量。这里卷积得到的特征图再经过 ReLU 激活函数 $\max(0, x)$ (这里 $x = W_1 * Y + B_1$) 进行了处理^[18]。

设 f_1 为单个滤波器的尺寸, c 是输入图像的通道数 (只对 Y 通道进行处理, 因此 $c=1$), 如果有 n_1 个卷积核, 那么 W_1 的尺寸为 $c \times f_1 \times f_1 \times n_1$ 。

非线性映射将特征向量是从低分辨率空间变换至高分辨空间, 输出特征为 $F_2(Y)$ 。过程如公式 (2) 所示:

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2) \quad (2)$$

同样， W_2 为卷积核， B_2 为神经元偏置向量。倘若有 n_2 个卷积核，卷积后生成 n_2 维的特征图，即 W_2 的尺寸为 $n_1 \times f_2 \times f_2 \times n_2$ 。

高分辨率图像重构利用之前得到的基于块的输出特征图来生成最终的高分辨率图像，输出为 $F(Y)$ 。过程如公式 (3) 所示：

$$F(Y) = W_3 * F_2(Y) + B_3 \quad (3)$$

W_3 包含了 c 个的卷积核，即 W_3 尺寸大小是 $n_2 \times f_3 \times f_3 \times c$ ， B_3 是一个维数为 c 的偏置向量。 W_3 被视为一个均值滤波器，整个重构过程是一个线性操作过程。

SRCNN 的参数可以表示为 $\Theta = \{W_1, W_2, W_3, B_1, B_2, B_3\}$ ，整个网络的训练过程就是对参数的估计和优化。通过最小化 $F(Y; \Theta)$ 与 X 之间的误差来得到参数最优解。给定高分辨率图像集合 $\{X_i\}$ 与其相应低分辨率图像集合 $\{Y_i\}$ ，均方误差 $L(\Theta)$ 为：

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^n \|F(Y_i; \Theta) - X_i\|^2 \quad (4)$$

n 代表训练集的数目， $\{X_i\}$ 集合通过裁剪生成随机排列的子块作为训练阶段的输入数据。为了得到低分辨率图像集 $\{Y_i\}$ ，先对得到的子图像经过高斯滤波模糊之后再下采样，然后再使用双三次插值放大至与原子图像尺寸大小一致。

SRCNN 算法在主观评价以及 PSNR 和 SSIM 等客观指标上与其他算法比较有明显的优势。算法在对图像超分辨率处理的运行速度上比其他算法处理快。对于 14 幅平均像素为 492×446 的图像，SRCNN 算法处理一幅图像的平均时间在 0.39 秒^[16]，而其他对比算法处理一幅图像的平均时间在 1 秒到 100 秒不等。然而 SRCNN 算法在以下几个方面可以进一步改进：调整卷积核大小以及参数，使计算复杂度降低；不增加计算量的基础上，适当增加卷积层数，使提取的特征越准确细致，提升超分辨率图像质量；迭代次数过大，使训练过程比较耗时，尝试修改网络在更小迭代次数时达到更优的训练结果，减少训练时间。

3 改进的基于卷积神经网络的图像超分辨率算法

牛津大学 VGG 小组将卷积神经网络运用在图像分类与视觉识别中，获得了令人满意的成

效，在 ImageNet 大赛中名列前茅。本文在该卷积神经网络的基础之上，借鉴了 VGG 小组提出的基于深度卷积网络的大规模视觉识别算法 VGG_ILSVRC_16_layers 模型、ImageNet 图像分类大赛的优秀网络模型 GoogLeNet，对基于卷积神经网络的图像超分辨率算法进行改进，网络结构框图如 2。

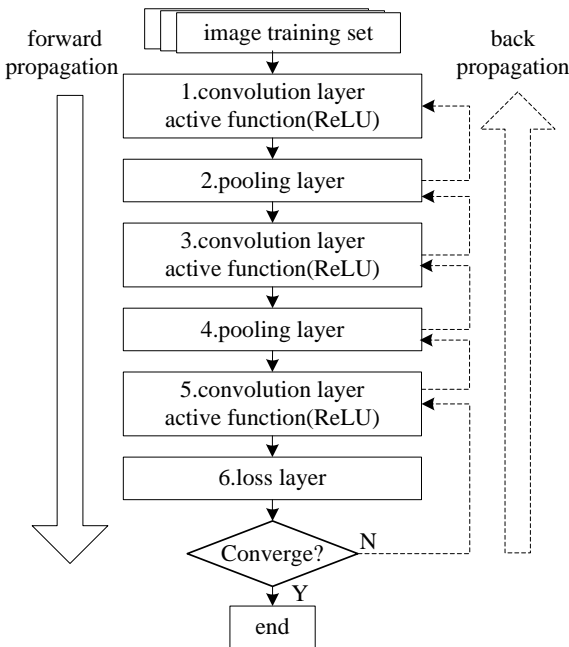


图 2 基于卷积神经网络的图像超分辨率算法框架

Fig.2 Image super resolution algorithm structure based on convolutional neural network

修改后的训练过程为，训练库的图像经过第一层卷积层后得到的特征图经过激活函数，再经过第二层池化层，再对特征进行卷积和激活函数运算，再通过第四层池化层，最后通过第五层卷积层和激活函数预算，计算损失层。训练直到收敛时才能结束。

3.1 卷积层

卷积层主要考虑卷积核尺寸，卷积核数目和学习率对模型效果和速度的影响。

3.1.1 卷积核尺寸

SRCNN 中提出的三层卷积为 9-1-5 模型，本文改为 3-3-3 模型。考虑到要与 SRCNN 算法性能进行比较，本文同样采用三层卷积层，但是与 SRCNN 不同的是，本文每一层卷积层的卷积核大小都是 3×3 。

SRCNN 算法在第一层卷积核尺寸选择 9×9 ，而实际上 3×3 已经能够足够覆盖图像的特征。根据 VGG 小组对深度卷积网络的实验结果和经验，对于 VCC_ILSVRC_16_layers 模型，第一层使用 3×3 大小卷积核，能够有效地减少参数。因此，本文算法第一层卷积核尺寸设为

3×3。

SRCNN 对不同的第二层卷积核尺寸进行测试，发现第二层卷积核尺寸越大，超分辨率效果越好，但同时会增大运算量。本文算法适当增加第二层卷积核尺寸，选择3×3，既提高效率，又能保证计算量不大幅度提升。

同样出于对计算量的考虑，由于第二层增大卷积核尺寸使参数增加，因此在第三层卷积核尺寸选择缩小为3×3，而在 SRCNN 算法中第三层尺寸为5×5。而缩小第三层卷积核尺寸可能会影响超分辨率图像的质量，所以我们分别对 3-3-3 模型和 3-3-5 模型的卷积网络进行训练和测试。迭代次数为 500k，实验表明，在训练阶段 3-3-5 模型仅平均每迭代 1000 次就比 3-3-3 模型多耗时 7s，而测试阶段输出超分辨率图像的平均 PSNR 仅比 3-3-3 模型高 0.13dB。

基于上述经验，出于对处理结果和运算速率两者之间的权衡考虑，我们尝试改为 3-3-3 模型。既增大第二层卷积核尺寸以提升处理效果，同时在该基础上减小第一、三层卷积核尺寸，以控制参数数目，降低运算的复杂度。

3.1.2 卷积核数目

第一层卷积核个数 n_1 与第二层卷积核个数 n_2 与卷积核尺寸共同影响超分辨率效果。在 SRCNN 算法中 9-1-5 模型中 $n_1 = 64$ ， $n_2 = 32$ 。本文算法在 3-3-3 模型下，对不同大小 n_1 和 n_2 进行测试，最终选取 $n_1 = 64$ ， $n_2 = 64$ 。

针对本文算法 3-3-3 模型在相同迭代次数、相同学习率、不同卷积核个数 n_1 和 n_2 对图 4 测试图像进行超分辨率处理，得到超分辨率图像的 PSNR 值，以测试图像 Bird 为例，均采用 50k 次迭代， 10^{-3} 的学习率，最终得到对比数据，如表 1 所示。

表 1 不同卷积核个数测试 PSNR 值

Table 1 Test PSNR values of different convolution kernels' numbers			
convolution kernel's number	$n_1 = 64, n_2 = 32$	$n_1 = 64, n_2 = 64$	$n_1 = 128, n_2 = 64$
PSNR	33.59	33.66	30.14

同时，出于对图像超分辨率过程中信息量和计算量的考虑，SRCNN 选取的子图像尺寸为 33×33 ，本文适当地增大了输入子块的尺寸，经过测试实验，输入子块尺寸适度加大，训练速度提升，时间减少。

3.1.3 学习率

在训练时对权重不断的更新优化中，在梯度项前会乘以一个系数，即学习率 (Learning Rate)。如果学习率太小，则收敛过慢，如果学习速率太大，代价函数会振荡。SRCNN 给出的学习率比较低，第一、二层为 10^{-4} ，第三层为 10^{-5} 。本文的学习率提高到 10^{-3} 。

本文对学习率采取了更进一步的测试。针对本文算法在相同迭代次数不同学习率得到的不同的网络学习模型对图 4 测试图像进行超分辨率处理，得到超分辨率图像的 PSNR 值。以测试图像 Bird 为例，均采用 50k 次迭代，对不同学习率，得到的网络模型进行测试和记录，如表 2 所示。可见，取学习率为 10^{-3} 效果最好。

表 2 不同学习率测试 PSNR 值

Table 2 Test PSNR values of different learning rates							
Learningrate	0.00001	0.00005	0.0001	0.0005	0.001	0.0015	0.002
PSNR	30.99	33.15	33.18	33.35	33.66	13.97	13.78

3.2 池化层

SRCNN 没有使用池化层，本文在第一层和第二层卷积层之后加入池化层。引用池化层之后能够减少卷积层输出的特征向量，降低维度，提高训练速度，并且不容易出现过拟合现象，改善结果^[19]。加上池化层，网络达到了 5 层的“深度”，深层次的网络结构更有利于学习图像数据更本质的表达。池化层可以有效减少参数数目，有利于简化网络，提高参数训练的效率。

池化层具有平移不变性。如果选择图像中的连续范围作为池化区域，并且只是池化相同的隐藏单元产生的特征，那么，这些池化单元就具有平移不变性。这就意味着即使图像经历了一个小的平移之后，依然会产生相同的特征。对于池化层来说，输入特征图的数目不会发生变化，但是每个输出特征图的尺寸会减小，该过程的本质就是下采样。如下式：

$$\mathbf{x}_j^\ell = f\left(\beta_j^\ell \text{down}(\mathbf{x}_j^{\ell-1}) + b_j^\ell\right) \quad (5)$$

\mathbf{x}_j^ℓ 和 $\mathbf{x}_j^{\ell-1}$ 分别代表当前层和前一层的第 j 个特征图， $\text{down}(\cdot)$ 代指下采样， $f(\cdot)$ 表示激活函数， β_j^ℓ 是乘性偏置项， b_j^ℓ 是加性偏置项，实验中令 $\beta_j^\ell = 1$ ， $b_j^\ell = 0$ 。主要是因为输入图像中临近 $n \times n$ 大小的图像块求和产生的影响，所以输出特征图整体缩小。

本文采用了最大池化^[20]，就是将输入图像分割为一系列不重叠的矩阵，然后对每个子区域，输出最大值。

4 实验结果及分析

本文基于卷积神经网络的图像超分辨率算法的实验环境包括硬件设备和软件配置两部分,测试所用计算机配置为 Intel Core i7-5820K CPU @ 3.30GHz x 12, GPU 为 NVIDIA GeForce TITAN X, 内存为 16GB。本文的实验平台搭载的操作系统为 64 位 ubuntu 14.04 LTS; MATLAB R2014a; CUDA Toolkit 7.0; OpenCV 3.0。

本文选择了传统双三次插值(Bicubic Interpolation, BI)方法、稀疏编码算法 ScSR^[15]、基于样本学习的 SR_NE_ANR 算法^[14]以及 SRCNN 算法^[16]等 4 种典型方法进行对比测试。本文算法采用的图像训练库来源于 ImageNet 数据库, 数据库增大到 800 幅。为了保证实验结果的合理客观, 本文选择 6 幅具有代表性、纹理细节丰富的图像进行测试和对比。测试图像如图 3 所示, 图(a)~图(e)来源于网络数据库, 图(f)选取的是手机拍摄的图像。每幅图像的分辨率如表 3 所示。本文所有测试的低分辨率图像放大 9 (3×3) 倍。

为了便于对比和测试, 本文采用的图像退化数学模型为:对原始高分辨率图像高斯模糊后实行下采样, 生成低分辨率图像。经由插值法放大的初始高分辨率图像输入到离线训练好的卷积神经网络模型之中, 经过网络学习和优化后的模型参数的处理, 生成最终的高分辨率图像。实验过程中每种算法选取的缩放因子都为 3, 对于 SRCNN 算法和本文算法, 为了进行客观合理的比较, 最终测试选取相同迭代次数的网络模型的学习结果进行对比测试。由于设备平台等限制, SRCNN 与本文选取的 Max Iterations 均为 500k。

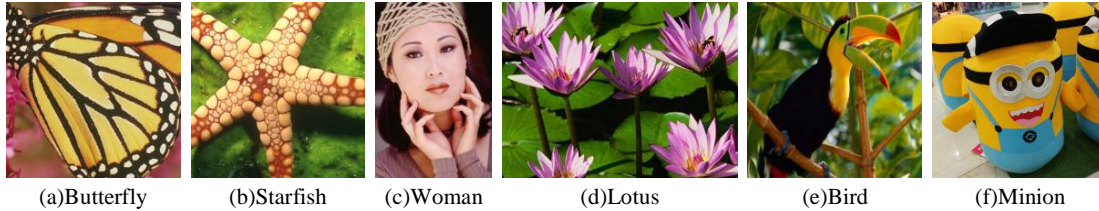


图 3 测试图像

Fig.3 Test images

表 3 测试图像分辨率

Table 3 the resolution of the test images						
	Butterfly	Starfish	Woman	Lotus	Bird	Minion
分辨率	256×256	256×256	228×344	413×309	288×288	330×330

4.1 主观效果

图 4~图 7 给出了包括双三次插值算法(BI)、ScSR 算法、SR_NE_ANR 算法、SRCNN 算法以及本文算法在内五种算法对 4 幅测试图像分别处理的结果。

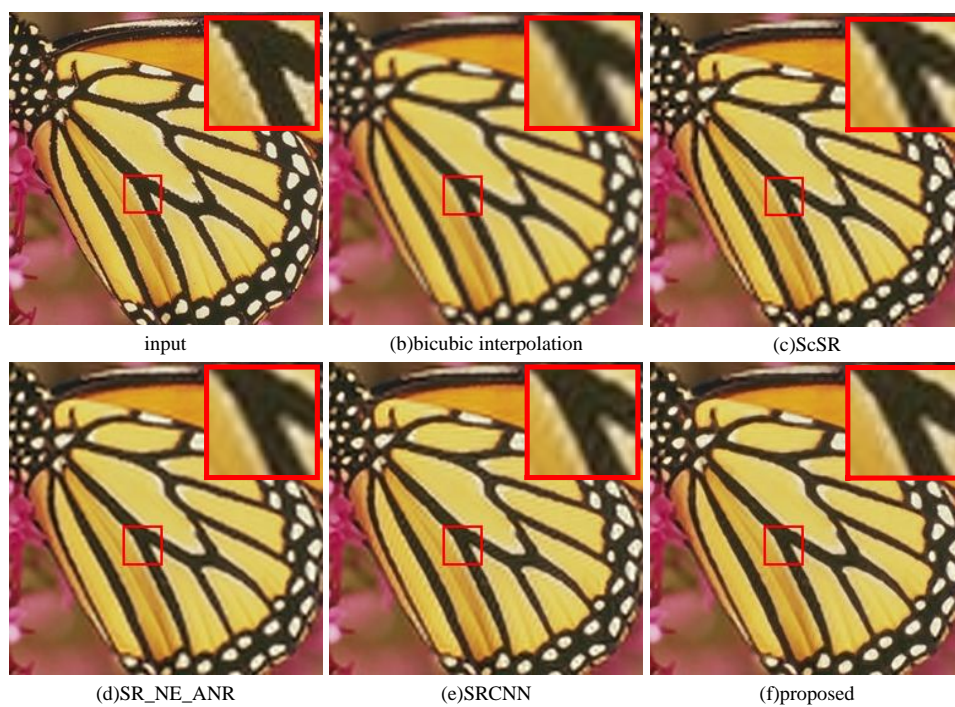


图 4 Butterfly 效果对比
Fig.4Thecomparison of 'Butterfly'

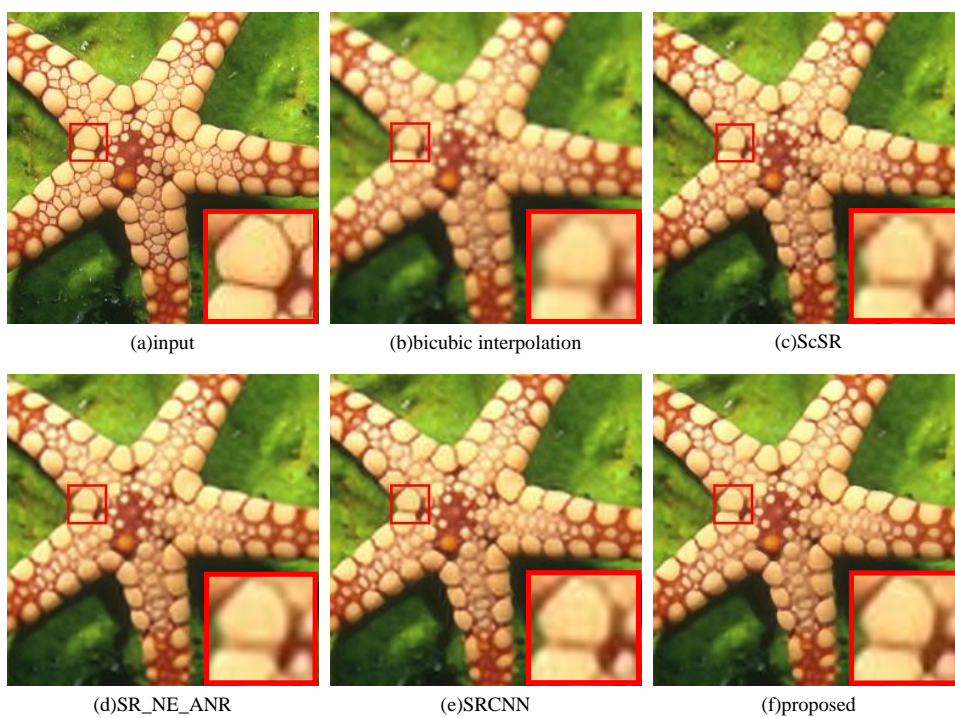


图 5 Starfish 效果对比
Fig.5Thecomparison of 'Starfish'

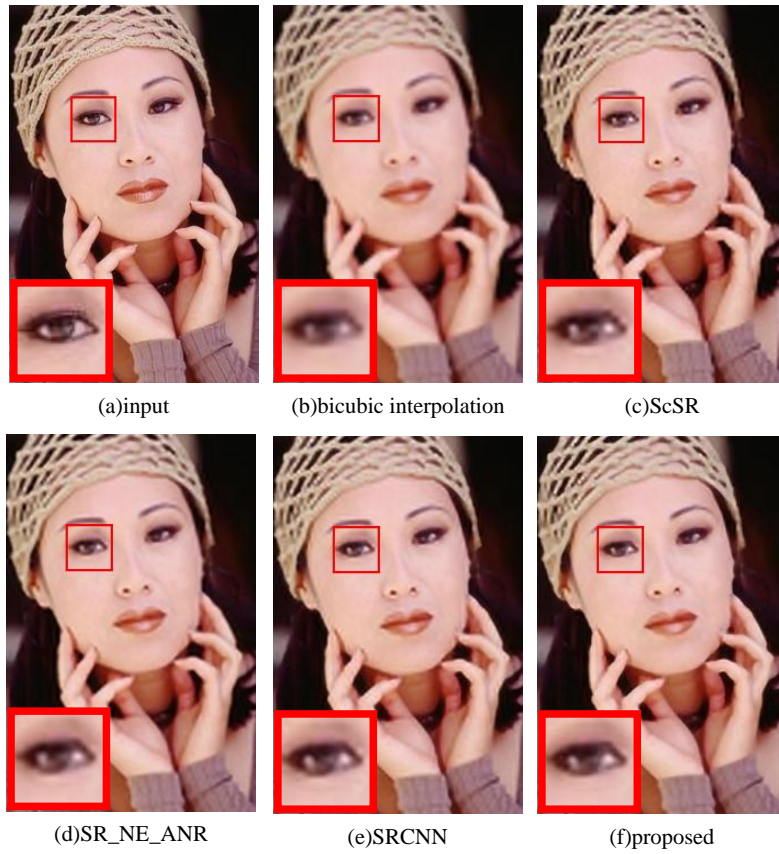


图 6 Woman 效果对比

Fig.6Thecomparison of'Woman'

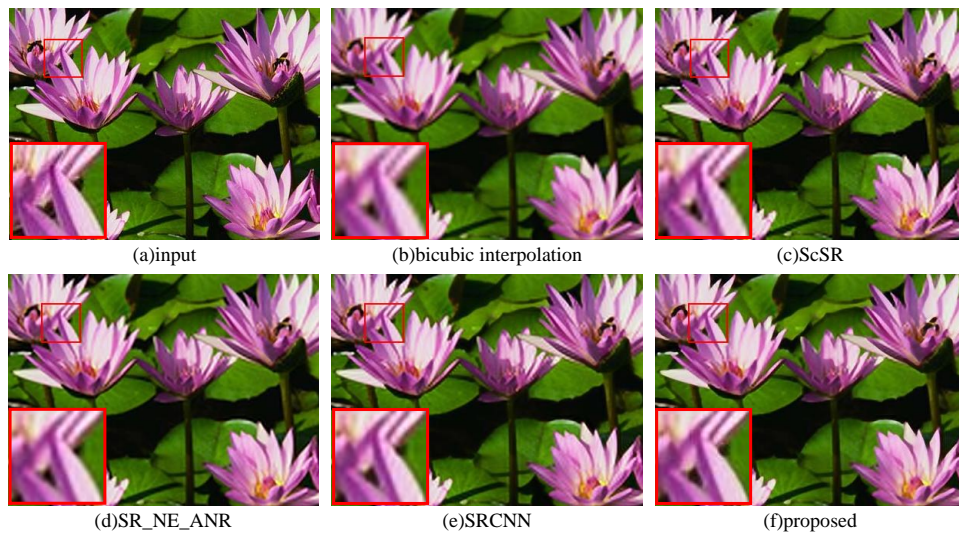


图 7 Lotus 效果对比

Fig.7Thecomparison of'Lotus'

从图 4~图 7 给出的处理结果明显可以看到，ScSR 算法、SR_NE_ANR 算法、SRCNN 算法和本文算法这四个基于学习的图像超分辨率算法处理结果要明显优于传统的双三次插值算法。主观上来看，ScSR 算法、SR_NE_ANR 算法和 SRCNN 算法都得到了较好的超分辨率图像。但是对某些测试图像，图像的边缘锐度相比其他几种算法，本文算法在边缘保持方面的结果要好。

图 4、图 5 和图 6 中分别对蝴蝶翅膀颜色变化的部分、海星的花纹和人眼进行局部放大，双三次插值、ScSR 算法得到的结果十分模糊，边缘不清晰，尤其在图 4 的蝴蝶翅膀黄黑色的边缘、图 5 中海星的白色和棕色的边缘以及图 6 中人眼和皮肤的边缘；SR_NE_ANR 算法、SRCNN 算法的实验结果略微能看清边缘，但是仍与原图差距过大；本文算法边缘处理比以上四种算法更加清晰，边缘更加完整，边缘保持方面更好，锐度更强，与原图效果最接近。图 7 中对莲花的花和叶的边缘进行部分放大，可以看出在双三次插值、ScSR 算法、SR_NE_ANR 算法放大后效果十分模糊，而 SRCNN 算法和本文算法处理后边缘较清晰。

4.2 客观效果

本文采用的客观评价指标主要是峰值信噪比(Peak Signal to Noise Ratio, PSNR)和图像的结构相似性(Structural Similarity, SSIM)^[7]。PSNR 是对经过处理的结果跟原图相比之下的误差进行定量计算，PSNR 愈高说明失真愈小。SSIM 愈发逼近 1，说明经过处理的结构与原图结构极为近似，也就意味着生成的结果图更好。

本文用以上两个指标对五种算法的处理结果做了客观评价。由表 4 客观测试结果可知，除了 SR_NE_ANR 算法对 Bird 测试图像的客观指标略高于本文算法之外，其余测试图像的测试结果表明，本文算法测试获取的 PSNR 和 SSIM 两者均高于双三次插值(BI)方法、ScSR 算法和 SRCNN 算法。说明本文算法产生的结果与原图特别接近，处理效果更好。

表 4 客观指标对比

Table 4 Objective index contrast

image	indices	BI	ScSR	SR_NE_ANR	SRCNN	proposed
Butterfly	PSNR	24.04	25.44	25.90	26.52	27.80
	SSIM	0.8216	0.8557	0.8717	0.8726	0.9034
Starfish	PSNR	27.00	28.04	28.05	28.42	28.77
	SSIM	0.8121	0.8485	0.8475	0.8519	0.8641
Woman	PSNR	28.56	29.82	30.33	30.44	30.46
	SSIM	0.8896	0.9095	0.9169	0.9144	0.9171
Lotus	PSNR	27.11	28.34	28.74	28.72	29.30
	SSIM	0.8421	0.8750	0.8839	0.8774	0.8895
Bird	PSNR	32.58	33.89	34.60	34.19	34.49
	SSIM	0.9256	0.9424	0.9488	0.9426	0.9462
Minion	PSNR	28.32	29.26	29.80	29.93	30.59
	SSIM	0.8656	0.8819	0.8921	0.8859	0.8952

除上述实验之外本文采取了更进一步的测试。为了突出本文算法相对 SRCNN 算法的有效提高，本文还进行了一系列的训练和学习过程，得到了很多种网络学习模型结果。本文针对 SRCNN 算法和本文算法在相同迭代次数下得到的两种不同的网络学习模型对上述测试图像进行超分辨率处理，得到超分辨率图像的 PSNR 值，以测试图像 Butterfly 为例，采用不同

迭代次数得到的网络模型进行测试和记录，最终绘制出其 PSNR 变化趋势图，如图 10 所示。其 SSIM 值变化趋势如图 11 所示。为了便于对比，图中加入双三次插值、ScSR、SR_NE_ANR 算法的测试结果，而且需要说明的是：双三次插值测试结果与迭代次数无关；ScSR、SR_NE_ANR 算法仅在 10 到 100 次迭代下对测试结果有影响，在 50k 次迭代以上时对测试结果没有影响。

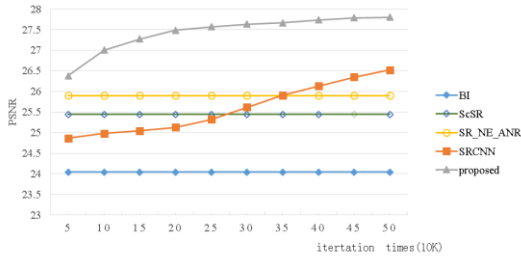


图 8 不同算法 PSNR

Fig.8 The PSNR of different algorithms

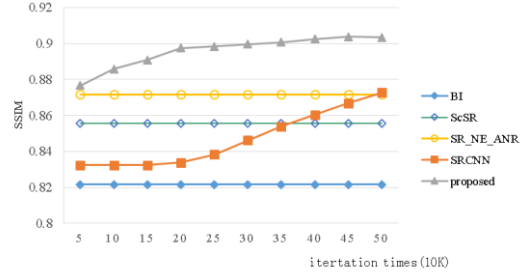


图 9 不同算法 SSIM

Fig.9 The SSIM of different algorithms

由图 8 和图 9 可以看出，无论是 PSNR 值还是 SSIM 值，本文算法测试得到的结果要整体优于 SRCNN 算法和双三次插值算法。并且相较于 SRCNN 算法，在训练过程中，本文算法达到最优效果的收敛速度比较快，如图 9 所示，SRCNN 算法需要 500k 次迭代才能达到的效果，本文算法只需要大约 60k 次迭代，提升幅度达到 88%。以上测试结果从客观方面说明了本文算法的超分辨率效果更好。

4.3 效率对比

该卷积网络的复杂度为 $O\{f_1^2 n_1 + n_1 f_2^2 n_2 + n_2 f_3^2\}$ 。对于 SRCNN 网络，其一次迭代的计算复杂度为 8032。对于本文改进的卷积神经网络结构，一次迭代的计算复杂度为 38016。因此本文每次迭代训练时参数计算量上比 SRCNN 要大，也即单次迭代训练时消耗更多时间。但值得注意的是，SRCNN 算法需要 500k 次迭代才能达到的效果，本文算法只需要 60k 次迭代，就能达到同样训练效果。平均耗时和总耗时情况如表 5 所示。

表 5 训练总耗时对比

Table5 Comparison of training time

	每 1k 次迭代用时 (s)	60k 次迭代用时 (s)	500k 次迭代用时 (s)
SRCNN	62.98	3778.8	31490
proposed	88.78	5326.8	

所以，虽然本文算法计算量比 SRCNN 算法略大，但从训练过程的整体来看，本文算法在低于 SRCNN 迭代次数的一个数量级的条件下，既保证超分辨率效果比 SRCNN 更好，同时训练总耗时大幅减少，效率高于 SRCNN 算法。

5 总结

本文阐述了深度学习应用于图像超分辨率的整体思路,从理论上剖析了卷积神经网络的学习过程。本文对基于卷积神经网络的图像超分辨率算法提出了一系列的改进方法。最后针对本文算法进行了一系列的实验测试工作,并和传统的双三次插值算法以及 SRCNN 算法做了对比。经过测试,在本文算法的超分辨率结果较好,图像边缘保持的比较好,边缘锐度更高,整体看起来更清晰。本文采用 PSNR 和 SSIM 两种图像评价标准对包括本文算法在内的各种算法进行了定量的测试,结果表明本文算法能够产生更佳的超分辨率效果。且在训练过程中,与 SRCNN 算法相比,在迭代次数较少的情况下能达到更优的结果,因此能够大幅减少训练时间。但在相同迭代次数下,本文算法的计算复杂度比 SRCNN 略高。因此,下一步研究内容,应该考虑进一步减少计算复杂度以提高运行速度。

参考文献

- [1]Juefei-Xu F, Savvides M. Single face image super-resolution via solo dictionary learning[C]//Image Processing (ICIP), 2015 IEEE International Conference on. IEEE, 2015: 2239-2243.
- [2]Xiao J, Pang G, Zhang Y, et al. Adaptive shock filter for image super-resolution and enhancement[J]. Journal of Visual Communication and Image Representation, 2016, 40: 168-177.
- [3]Li Min, Cheng Jian, Le xiang, et al.. Super-resolution based on sparse dictionary coding[J]. Journal of Software, 2012,05:1315-1324.
李民,程建,乐翔,等. 稀疏字典编码的超分辨率重建[J]. 软件学报,2012,05:1315-1324.
- [4]Zhang K, Tao D, Gao X, et al. Learning multiple linear mappings for efficient single image super-resolution[J]. IEEE Transactions on Image Processing, 2015, 24(3): 846-861.
- [5]Bäz M, Eichenseer A, Seiler J, et al. Hybrid super-resolution combining example-based single-image and interpolation-based multi-image reconstruction approaches[C]//Image Processing (ICIP), 2015 IEEE International Conference on. IEEE, 2015: 58-62.
- [6]Zhou Jinghong, Zhou Cui, Zhu Jianjun, et al. A method of super-resolution reconstruction for remote sensing image based on non-subsampled contourlet transform[J]. Acta Optica Sinica,2015,01:106-114.
周靖鸿,周瑾,朱建军,等. 基于非下采样轮廓波变换遥感影像超分辨重建方法[J]. 光学学报,2015,01:106-114.
- [7]Xiao Jinsheng, Gao wei, Peng hong, et al. Detail Enhancement for Image Super-Resolution Algorithm Based on SVD and Local Self-Similarity [J]. Chinese Journal of Computers, 2016, 07:1393-1406.
肖进胜, 高威,彭红, 等,基于局部自相似性和奇异值分解的超采样图像细节增强[J],计算机学报, 2016,07:1393-1406.
- [8]LianQiusheng, Zhang Wei. Image super-resolution algorithms based on sparse representation of classified image patches[J]. Acta Electronica Sinica, 2012, 05:920-925.
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, 2012,05:920-925.
- [9]Yang C Y, Huang J B, Yang M H. Exploiting self-similarities for single frame super-resolution[C]//Asian Conference on Computer Vision. Springer Berlin Heidelberg, 2010: 497-510.
- [10]Zhang Y, Liu J, Yang W, et al. Image super-resolution based on structure-modulated sparse representation[J]. IEEE Transactions on Image Processing, 2015, 24(9): 2797-2810.
- [11]Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2015: 5197-5206.
- [12]Choi J S, Bae S H, Kim M. Single image super-resolution based on self-examples using context-dependent subpatches[C]//Image Processing (ICIP), 2015 IEEE International Conference on. IEEE, 2015: 2835-2839.
- [13]Freeman W T, Jones T R, Pasztor E C. Example-based super-resolution[J]. IEEE Computer graphics and Applications, 2002, 22(2): 56-65.
- [14]Timofte R, De Smet V, Van Gool L. Anchored neighborhood regression for fast example-based super-resolution[C]//Proceedings of the IEEE International Conference on Computer Vision. 2013: 1920-1927.

- [15]Yang J, Wright J, Huang T S, et al. Image super-resolution via sparse representation[J]. IEEE transactions on image processing, 2010, 19(11): 2861-2873.
- [16]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.
- [17]LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [18]Nair V, Hinton G E. Rectified linear units improve restricted boltzmann machines[C]//Proceedings of the 27th International Conference on Machine Learning (ICML-10). 2010: 807-814.
- [19]He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9): 1904-1916.
- [20]Scherer D, Müller A, Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition[C]//International Conference on Artificial Neural Networks. Springer Berlin Heidelberg, 2010: 92-101.