



DOI: 10.12086/oe.2019.180419

## 基于多尺度特征损失函数的 图像超分辨率重建

徐亮, 符冉迪\*, 金炜, 唐彪, 王尚丽

宁波大学信息科学与工程学院, 浙江 宁波 315211



**摘要:** 在图像超分辨率重建问题中,许多基于深度学习的方法大多采用传统的均方误差(MSE)作为损失函数,重建后的图像容易出现细节模糊和过于平滑的问题。针对这一问题,本文对传统的均方误差损失函数进行改进,提出一种基于多尺度特征损失函数的图像超分辨率重建方法。整个网络模型由基于 DenseNet 的重建模型和一个用来优化多尺度特征损失函数的卷积神经网络串联构成。将重建后得到的图像和对应的原始高清图像作为串联的卷积神经网络的输入,计算重建图像卷积得到的不同尺度特征图与对应的原始高清图像卷积得到的不同尺度特征图的均方误差。实验结果表明,本文提出的方法在主观视觉效果和 PSRN、SSIM 上均有所提升。

**关键词:** 图像超分辨率重建; 稠密卷积神经网络; 多尺度特征损失函数; 深度学习

**中图分类号:** TP391.41

**文献标志码:** A

**引用格式:** 徐亮,符冉迪,金炜,等. 基于多尺度特征损失函数的图像超分辨率重建[J]. 光电工程, 2019, 46(11): 180419

## Image super-resolution reconstruction based on multi-scale feature loss function

Xu Liang, Fu Randi\*, Jin Wei, Tang Biao, Wang Shangli

School of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China

**Abstract:** In the image super-resolution reconstruction, many methods based on deep learning mostly adopt the traditional mean squared error (MSE) as the loss function, and the reconstructed image is prone to the problem of fuzzy details and too smooth. In order to solve this problem, this paper improves the traditional mean square error loss function and proposes an image super-resolution reconstruction method based on multi-scale feature loss function. The whole network model consists of a DenseNet-based reconstruction model and a convolutional neural network which is used to optimize the multi-scale feature loss function. Taking the reconstructed image and the corresponding original HD image as the input of the convolved neural network in series, the mean square error of the different scale feature images obtained by convolution of the reconstructed image with the corresponding original HD image was calculated. Experimental results show that the method in this paper is improved in subjective vision, PSRN and SSIM.

**Keywords:** image super-resolution reconstruction; densely connected convolutional neural networks; multi-scale feature loss function; deep learning

**Citation:** Xu L, Fu R D, Jin W, et al. Image super-resolution reconstruction based on multi-scale feature loss function[J]. *Opto-Electronic Engineering*, 2019, 46(11): 180419

收稿日期: 2018-08-04; 收到修改稿日期: 2019-01-04

基金项目: 国家自然科学基金资助项目(61471212); 浙江省自然科学基金资助项目(LY16F010001)

作者简介: 徐亮(1994-),男,硕士研究生,主要从事数字图像处理与模式识别的研究。E-mail: 692118188@qq.com

通信作者: 符冉迪(1971-),男,副教授,硕士生导师,主要从事压缩感知和数字图像处理等的研究。E-mail: 17346807657@163.com

## 1 引言

图像超分辨率重建技术利用相关网络模型来进行特征的提取,恢复出相关的高频信息,在一幅或一组低分辨率(low resolution, LR)图像中通过增加高频信息,以此来重建出高分辨率(high resolution, HR)的图像,是计算机视觉领域中最经典的问题之一<sup>[1]</sup>。重建后的图像具有较高的像素密度,细节信息更加丰富,在医学影像处理、卫星遥感、视频监控、模式识别等领域有着广泛的应用<sup>[2-5]</sup>。

目前,图像超分辨率重建技术主要可以分成三类:基于插值的方法<sup>[6-7]</sup>、基于重建的方法<sup>[8-9]</sup>以及基于学习的方法<sup>[10-13]</sup>。其中,基于插值的方法有最近邻插值<sup>[14]</sup>和双三次插值<sup>[15]</sup>。最近邻插值法规律较为简单,重建效果较差,容易出现方块效应和边缘锯齿效应。双三次插值法相比最近邻插值方法在图像边缘的平滑度上有所提升,不过总体重建效果仍不理想。基于插值的方法虽然简单易于实现,可是细节信息丢失较为严重,造成图像边缘轮廓的复原不理想。基于重建的方法主要分为频域法和空域法两大类<sup>[16]</sup>。频域法可用的先验知识十分有限,一般通过消除频谱混叠的现象从而达到提升图像分辨率的效果。空域法主要有凸集投影法(projection onto convex set, POCS)、迭代反投影法(iterative back projection, IBP)、最大后验概率估计法(maximum a posterior, MAP)、混合 MAP/POCS 法等。基于重建的方法十分依赖高分辨率图像的先验知识,若图像较为复杂,则重建的效果会受到严重影响。基于学习的图像超分辨率重建技术通过提取丰富的高分辨率图像和低分辨率图像的特征来学习 LR 与 HR 图像块之间的映射关系。其中较传统的方法有基于邻居嵌入的方法<sup>[17]</sup>和稀疏表示的方法<sup>[18-19]</sup>。基于邻居嵌入的单幅图像超分辨率重建方法(single image super-resolution, SISR)利用 LR 图像块和 HR 图像块之间的几何相似度来学习它们之间的回归或映射关系。在稀疏表示的方法中,将 LR 图像块和 HR 图像块利用稀疏编码理论建立字典之间的映射关系,通过学习映射关系重建 HR 图像。

近几年来随着深度学习理论的大火大热,其在图像处理领域取得了许多突出性的成果。和传统的浅层学习只能提取简单的特征进行函数拟合相比,深度学习能通过有效地学习得到层次化的特征表示。2014年,香港中文大学 Dong 等<sup>[13]</sup>提出了基于卷积神经网络的图像超分辨率方法(learning a deep convolutional

network for image SR, SRCNN)。SRCNN 由 3 层卷积神经网络组成,训练过程可以用 GPU 加速,直接由 LR 图像预测 HR 图像。SRCNN 的提出很好的证明了利用深度学习直接学习 LR-HR 之间端到端的映射是可行的,为今后的深度学习实现图像超分辨率重建打下了重要的基础。但是 SRCNN 仍有不足之处,网络深度不够,训练收敛较慢,特征提取不充分。后续 Dong 等<sup>[20]</sup>对 SRCNN 的网络结构进行改进,提出了 FSRCNN,使用更小的卷积核和更多的映射层,在最后使用了反卷积层来放大尺寸,不需要先通过双三次插值的方法对输入图像进行放大。由于低分辨率图像携带的低频信息与高分辨率图像携带的低频信息在一定程度上是相近的,2016年, Kim 等<sup>[21]</sup>结合残差网络提出了 VDSR(accurate image super-resolution using very deep convolutional networks),该方法通过残差连接学习高分辨率图像和低分辨率图像之间的高频部分残差。2017年, Tong 等<sup>[22]</sup>以及 Yamanaka 等<sup>[23]</sup>结合稠密卷积神经网络分别提出了 SRDenseNet(image super-resolution using dense skip connections)和 DCSCN(fast and accurate image super resolution by deep CNN with skip connection and network in network),通过稠密连接的方式减轻了梯度消失问题,加强了特征的传播,实现特征的复用,进一步提升了重建效果。

传统的基于 CNN 的图像超分辨重建方法大多采用均方误差作为损失函数,可获得较好的 PSNR,例如 SRCNN、FSRCNN、SRDenseNet,但是重建后的图像容易出现边缘模糊和过于平滑的问题。针对这一问题本文对均方误差损失函数进行改进,提出了基于多尺度特征损失的图像超分辨率重建方法。在基于 DenseNet 的重建模型后,串联一个 3 层的卷积神经网络用于对重建图像和其对应的原始高清图像进行多尺度的特征提取,然后计算两幅图像从底层特征到高层特征对应层次特征的均方误差。

## 2 基于深度学习的图像超分辨率重建方法

### 2.1 SRCNN

在图像的超分辨率重建问题中,LR 图像  $R_{LR}$  是由 HR 图像  $R_{HR}$  经过降质变换所产生的,其目的是通过对 LR 图像的学习与分析恢复出对应的 HR 图像,其观测模型可以表示成:

$$R_{LR} = D \cdot B \cdot M \cdot R_{HR} + N, \quad (1)$$

其中： $D$  为降采样算子， $B$  为光学模糊算子， $M$  为运动模糊算子， $N$  为加性噪声。

目前，基于深度学习的图像超分辨率重建方法取得了较好的成果。其中，SRCNN 是一种经典的基于深度学习的单幅图像超分辨率重建方法，该网络模型结构简单，如图 1 所示，由三层卷积神经网络组成，分别为图像特征提取模块、非线性映射模块和重建模块，其尺寸依次为  $9 \times 9 \times 64$ ， $1 \times 1 \times 32$  和  $5 \times 5 \times 1$ 。SRCNN 在训练时采用 91 张自然图像作为训练的数据集，并进行双三次差值的预处理。SRCNN 第一部分的特征提取层利用  $9 \times 9$  卷积核对输入图像进行特征提取，表示成 64 幅不同的特征图。第二部分的非线性映射层通过组合特征进一步将其映射成 32 幅特征图。第三层对第二层的特征图进行组合表示，从而得到 HR 图像。SRCNN 的重建效果相对于一些基于浅层学习的图像超分辨率重建方法有较大的改善，但是 SRCNN 仍然有几个方面可以改进：首先 SRCNN 采用的卷积核较大，会导致计算复杂度增加且特征提取过程中细节容易丢失。其次，适当地加深网络有利于组建更丰富的特征。利用双三次插值进行上采样，该方法其本身具有一定的不确定性，运行起来速度较慢，导致图像细节的丢失，不利于提取特征。

## 2.2 FSRCNN

FSRCNN 对 SRCNN 进行了改进，结构如图 2 所

示。FSRCNN 不再对输入图像进行双三次插值的处理，而是直接将 LR 图像作为输入，最后通过反卷积层进行放大，并且使用了更小的卷积核和更多的非线性映射层来加深网络深度。FSRCNN 的整个模型可分为五个部分。第一部分为特征提取层，采用  $5 \times 5$  的卷积核实现。第二部分为降采样层，利用  $1 \times 1$  卷积核进行降维，从而减少网络参数，降低计算复杂度。第三部分为非线性映射层，相比于 SRCNN 采用  $5 \times 5$  的卷积核，FSRCNN 采用多个  $3 \times 3$  卷积核串联的方式来减少计算量。第四部分为扩张层，利用  $1 \times 1$  卷积核进行升维操作。第五部分为反卷积层，实现对图像的放大。FSRCNN 增加了非线性映射层的数量，但是重建图像依然来源于非线性映射模块的单一映射，特征的利用率较低。

## 2.3 基于 DenseNet 的图像超分辨率重建

近几年来随着卷积神经网络快速发展，各种优秀的模型层出不穷，比如 GoogleNet、VGGNet、ResNet 等。但是随着卷积神经网络的加深，训练过程中的前向传播信号和梯度信号在经过很多层之后可能会逐渐消失，该问题在 Highway Network 和 ResNet 中得到了较好的解决。稠密连接的卷积神经网络(densely connected convolutional networks, DenseNet)同样可以解决上述问题，并且实现了特征的复用，极大程度上提高了特征的利用率。

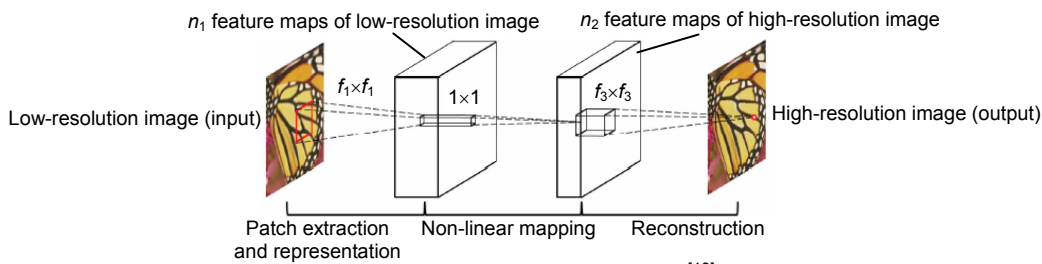


图 1 SRCNN 结构图<sup>[13]</sup>

Fig. 1 Structure of SRCNN<sup>[13]</sup>

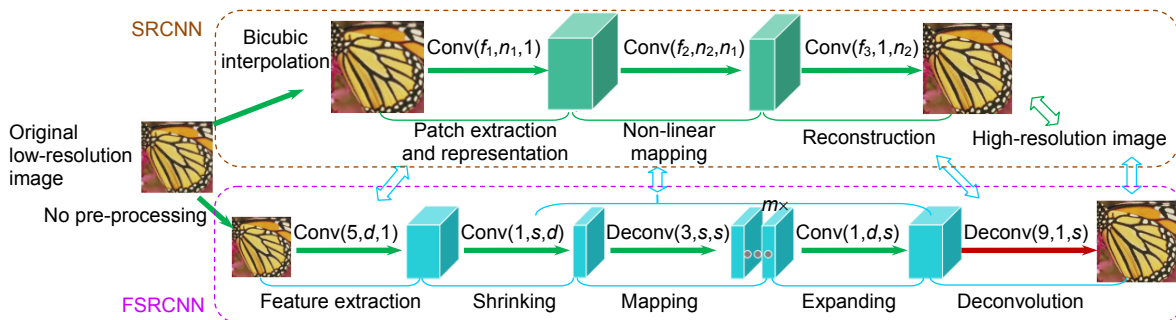


图 2 FSRCNN 结构图<sup>[20]</sup>

Fig. 2 Structure of FSRCNN<sup>[20]</sup>

Tong 等<sup>[22]</sup>将稠密连接结构应用到了图像超分辨率重建问题中,提出了 SRDenseNet,原文中设计了三种结构并做了比较,如图 3 所示。图 3(a)模型将最顶层稠密块的输出作为反卷积层的输入。图 3(b)模型将底层卷积神经网络提取的特征和最顶层稠密块的输出作为反卷积层的输入。图 3(c)模型将最底层卷积神经网络提取的特征和每一个稠密块的输出作为后续每一个稠密块的输入。由于把所有稠密块的输出都进行了串联,增加了计算量,因此在反卷积层前添加了一个卷积核为  $1 \times 1$  的瓶颈层来进行降维。

### 3 基于多尺度特征损失的图像超分辨率重建方法

#### 3.1 本文模型

在图像超分辨率重建问题中采用传统的均方误差作为损失函数能获得较好的 PSNR,在 SRCNN、FSRCNN、SRDenseNet 以及一些其它的网络模型中得到了广泛的应用。但是重建后的图像容易出现边缘模

糊或者过于平滑的问题,本文提出的多尺度特征损失函数在一定程度上能改善这一问题。通过对 SRCNN、FSRCNN 以及 SRDenseNet 等方法的分析,以 DenseNet 模型为基本框架搭建重建模型,并在重建模型后串联了一个 3 层的卷积神经网络用以计算多尺度特征损失函数,模型结构如图 4 所示。重建模型由稠密连接块、降维层、反卷积层以及重建层四个部分组成。每个稠密连接块由 4 个卷积层组成,均采用  $3 \times 3$  的卷积核,输出的特征图数量均为 256。由于将所有稠密连接块的输出都进行了串联,为减少计算负担,在降维层通过  $1 \times 1$  卷积核将特征图减少到 256。在 SRCNN 中,将输入图像进行双三次插值的预处理,导致图像的细节丢失,不利于特征的提取,本文模型直接使用原始图像作为输入,然后利用反卷积层进行上采样,恢复出特定尺寸大小的图像。在反卷积层之后通过  $3 \times 3$  卷积核重建出单通道高清图。最后对重建后的图像和对应的原始高清图通过串联的 3 层卷积神经网络进行特征提取,计算重建图像卷积得到的特征图与原始高

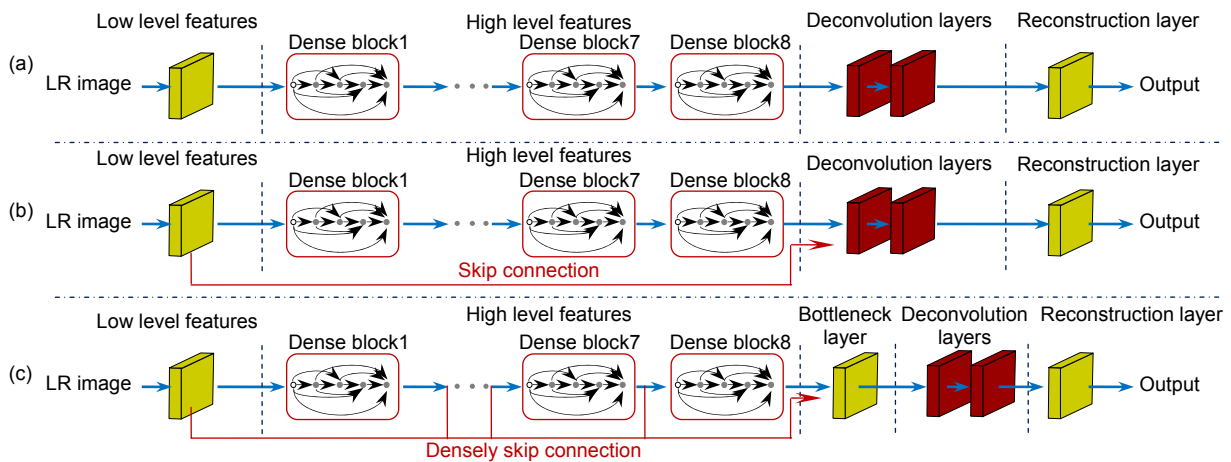


图 3 SRDenseNet 结构图<sup>[22]</sup>

Fig. 3 Structure of SRDenseNet<sup>[22]</sup>

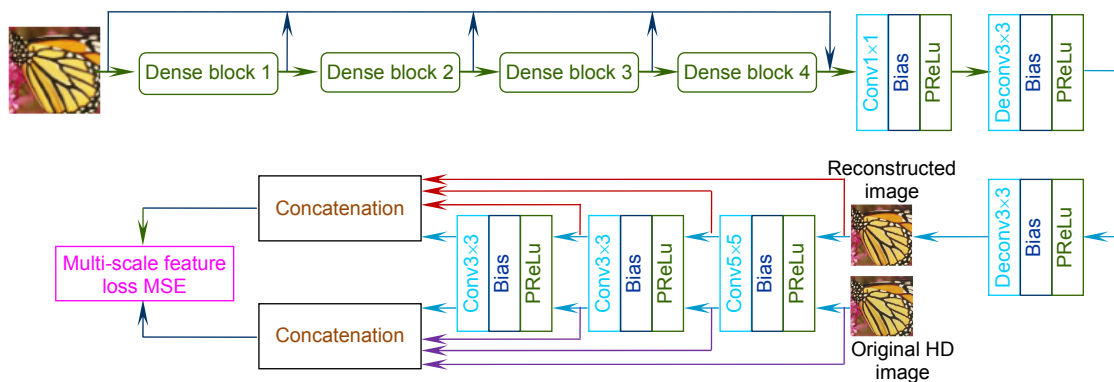


图 4 基于多尺度特征损失函数的图像超分辨率重建模型

Fig. 4 Image super-resolution reconstruction model based on multi-scale feature loss function



清图像卷积得到的特征图之间的均方误差,多尺度特征损失函数的计算过程可定义为

$$L_{MSE}(F(x), F(x')) = \frac{1}{L} \sum_{i=1}^L \|F^i(x) - F^i(x')\|^2, \quad (2)$$

其中:  $x$  为重建后的图像,  $x'$  为对应的原始高清图像,  $F^i(x)$  为重建图像经过第  $i$  层卷积层后提取的特征,同理  $F^i(x')$  为对应的原始高清图像经过第  $i$  层卷积层后提取的特征,  $L$  为串联的卷积神经网络的卷积层数量,本文中  $L=3$ 。

为了达到更好的重建质量,本文的网络模型并没有池化层。池化层经常被用于图像分类的神经网络中,对卷积后的结果进行下采样,以此来减少参数,从而得到旋转、伸缩、平移等不变性质。在超分辨率重建技术中希望能够增加图像的细节信息,使其具有更加丰富的高频信息,池化层则会使得卷积后的细节信息丢失,不利于重建。

### 3.2 激活函数 PReLU 和 ReLu

SRCNN 采用修正线性单元 ReLu(rectified linear unit, ReLu)来作为激活函数(用  $f_{ReLu}$  表示):

$$f_{ReLu}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (3)$$

ReLu 将 0 作为阈值,相比于 Sigmoid 和 tanh 激活函数,ReLu 激活函数在梯度下降上有更快的收敛速度,并且在一定程度上能抑制梯度消失现象。但是不足之处在于 ReLu 强行将负数置零的性质可能会导致相关特征丢失。

针对上述问题,本文采用 PReLU(parametric rectified linear unit, 用  $f_{PReLU}$  表示)作为激活函数,PReLU 激活函数由 He 等<sup>[24]</sup>提出,在 ReLu 上进行了改进,是一种带参数的修正线性单元:

$$f_{PReLU}(x_i) = \begin{cases} x_i, & x_i > 0 \\ a_i x_i, & x_i \leq 0 \end{cases} \quad (4)$$

从式(3)和式(4)可以发现,PReLU 比 ReLu 多了一个参数  $a_i$ 。一般  $a_i$  是一个非常小的固定值(比如  $a_i=0.01$ ) 如果  $a_i=0$  那么 PReLU 就是 ReLu。采用 PReLU 作为激活函数不但可以保留 ReLu 加速收敛的优势,同时也能防止特征的丢失。

### 3.3 图像质量评价指标 PSNR 和 SSIM

本文利用峰值信噪比(peak signal to noise ratio, PSNR, 用  $R_{PSNR}$  表示)和结构相似性(structural similarity, SSIM, 用  $S_{SSIM}$  表示)这两个客观评价指标来对重建图像进行质量评价。PSNR 用于对处理后的结果和

原图相比的误差进行计算,PSNR 值越大,说明失真越小,单位为 dB,其计算式:

$$R_{PSNR} = 10 \lg \left( \frac{(2^n - 1)^2}{e_{MSE}} \right), \quad (5)$$

其中:  $n$  为每像素的比特数,一般的灰度图像取 8,即像素灰阶数为 256, MSE 为当前图像  $X$  和参考图像  $Y$  的均方误差(用  $e_{MSE}$  表示):

$$e_{MSE} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2 \quad (6)$$

结构相似性的取值范围在 [0,1], 其值越大,表示图像失真越小。它分别从亮度、对比度、结构三方面度量图像结构相似性:

$$l(X, Y) = \frac{2\mu_X \mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}, \quad (7)$$

$$c(X, Y) = \frac{2\sigma_X \sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2}, \quad (8)$$

$$s(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X \sigma_Y + C_3}, \quad (9)$$

其中:  $C_1$ 、 $C_2$ 、 $C_3$  设置为常数,目的是为了避开分母为 0 的情况,一般取  $C_1 = (K_1 \times L)^2$ ,  $C_2 = (K_2 \times L)^2$ ,  $C_3 = C_2 / 2$ , 通常设置常数  $K_1=0.01$ ,  $K_2=0.03$ ,  $L=255$ 。 $\mu_X$  和  $\mu_Y$  分别表示图像  $X$  和  $Y$  的均值:

$$\mu_X = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X(i, j), \quad (10)$$

$$\mu_Y = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W Y(i, j) \quad (11)$$

$\sigma_X$  和  $\sigma_Y$  分别表示图像  $X$  和  $Y$  的方差:

$$\sigma_X^2 = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - \mu_X)^2, \quad (12)$$

$$\sigma_Y^2 = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W (Y(i, j) - \mu_Y)^2 \quad (13)$$

$\sigma_{XY}$  则表示图像  $X$  和  $Y$  的协方差:

$$\sigma_{XY} = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W ((X(i, j) - \mu_X)(Y(i, j) - \mu_Y)) \quad (14)$$

根据式(7)至式(9)可以写出 SSIM 的公式:

$$S_{SSIM}(X, Y) = l(X, Y) \times c(X, Y) \times s(X, Y) \quad (15)$$

## 4 实验及结果分析

### 4.1 实验说明

本文实验采用操作系统为 Windows10(64 位), 64 GB 运行内存, NVIDIA GeForce 1080Ti GPU。使用的深度学习框架为 Tensorflow, 使用的编程语言为 Python3.5。设置 dropout=0.5, 选择优化器为 Adam, 采用初始的学习率为 0.0002。

本文实现步骤主要分为五个部分：1) 训练所需的外部图像库的建立；2) 搭建网络模型；3) 利用建立的外部图像库进行训练，不断优化网络参数，得到输入图像的特征表达和先验信息；4) 将 LR 图像重建得到 HR 图像；5) 将重建图像以及对应的原始高清图像输入串联的卷积神经网络进行不同尺度特征的提取，计算重建图像卷积得到的特征与原始高清图像卷积得到的特征的均方误差。

训练数据集方面，本文实验采用著名的 Yang 用过的 91 张和 Dong 用过的 200 张 Berkeley Segmentation Dataset 作为网络模型的训练样本，合计 291 张图像。考虑到卷积神经网络的训练依赖于大量的数据样本，为了保证样本的充足，需要对原 291 张训练样本进行适当的扩充。首先对原样本集进行左右和上下翻转，得到比原先多 4 倍的训练样本集，即  $291+(291\times 4)=1455$  张训练样本。然后再对原样本进行尺寸变化，分别放大 2 倍、3 倍、4 倍，并且对放大后的图像进行 180° 镜像变换，即  $291\times 2\times 3=1746$ ，总共

合计  $1455+1746=3201$  张训练样本。测试集选用超分辨率研究领域标准评测数据集 Set5、Set14、BSD100，一共 119 张测试样本，并利用 PSNR 和 SSIM 进行客观指标的评价，有效地展示算法的实验效果。

#### 4.2 结果分析

利用扩充后的数据集 Yang91 和 BSD200 训练本文网络模型，表 1 列出了基于插值的 Bicubic 方法，经典的深度学习算法 SRCNN，结合残差学习的 DnCNN-3 算法<sup>[25]</sup>，使用本文基于 DenseNet 重建模型(不包括串联的卷积神经网络)的算法以及基于多尺度特征损失函数的方法，分别在测试集 Set5、Set14、BSD100 上的平均 PSNR 和 SSIM 数值。

本文提出的模型主要由基于 DenseNet 的重建模型以及用来优化多尺度特征损失函数的卷积神经网络两个部分组成。图 5 展示了使用本文提出的基于 DenseNet 重建模型(不包括串联的卷积神经网络)和传统的均方误差作为损失函数重建得到的效果图与使用

表 1 不同超分辨率算法在数据集 Set5、Set14、BSD100 上的平均 PSNR 以及 SSIM

Table 1 Average PSNR and SSIM of different super-resolution algorithms on datasets Set5, Set14 and BSD100

Dataset	Scale	Bicubic	SRCNN	DnCNN-3	Based on DenseNet	Our method
Set5	2×	33.66/0.9299	36.66/0.9542	37.58/0.9590	37.79/0.9589	<b>37.98/0.9632</b>
	3×	30.39/0.8682	32.75/0.9090	33.75/0.9222	33.99/0.9256	<b>34.36/0.9264</b>
	4×	28.42/0.8104	30.48/0.8628	<b>31.40/0.8845</b>	30.93/0.8754	31.32/0.8819
Set14	2×	30.24/0.8688	32.42/0.9063	33.03/0.9128	33.48/0.9145	<b>33.69/0.9175</b>
	3×	27.55/0.7742	29.28/0.8209	<b>29.81/0.8321</b>	29.60/0.8281	29.69/0.8296
	4×	26.00/0.7072	27.49/0.7503	28.04/0.7672	28.29/0.7689	<b>28.56/0.7714</b>
BSD100	2×	29.56/0.8431	31.36/0.8879	31.90/0.8961	32.19/0.8988	<b>32.73/0.9023</b>
	3×	27.21/0.7385	28.41/0.7863	28.85/0.7981	29.18/0.8001	<b>29.43/0.8103</b>
	4×	25.96/0.6675	26.90/0.7101	<b>27.29/0.7253</b>	27.03/0.7203	27.24/0.7232

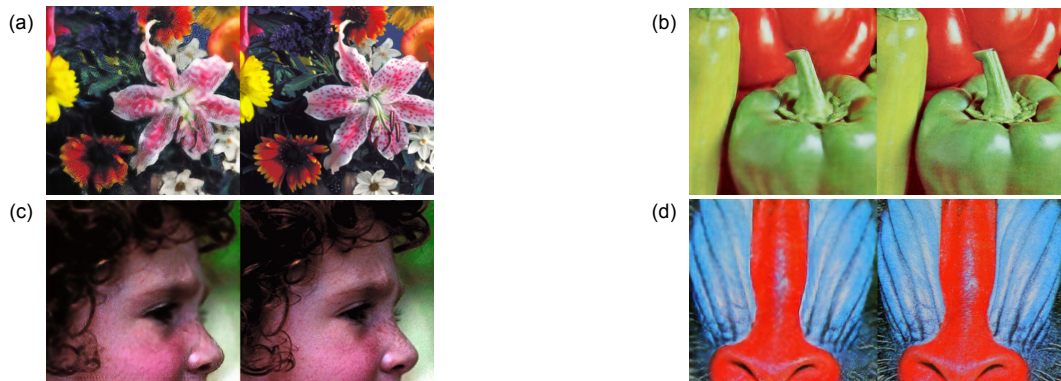


图 5 基于 DenseNet 的重建模型以及传统均方误差作为损失函数的重建效果与本文方法的对比

Fig. 5 The reconstructed results based on the DenseNet reconstruction model and the traditional mean square error as the loss function are compared with the method in this paper

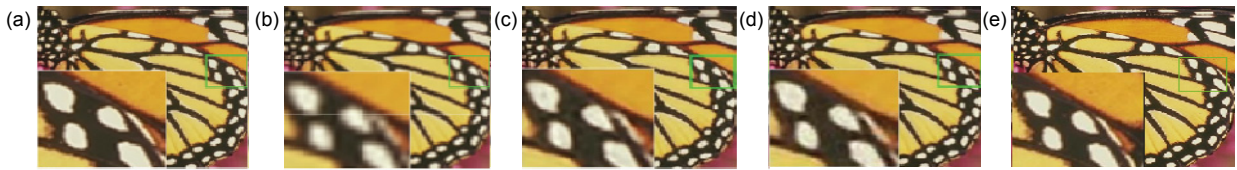


图 6 Butterfly 放大 2 倍时各算法重建效果。(a) 原图; (b) Bicubic; (c) SRCNN; (d) DnCNN-3; (e) 本文方法  
Fig. 6 Reconstruction results of each algorithm for butterfly amplification by 2 times. (a) Original; (b) Bicubic; (c) SRCNN; (d) DnCNN-3; (e) Our method

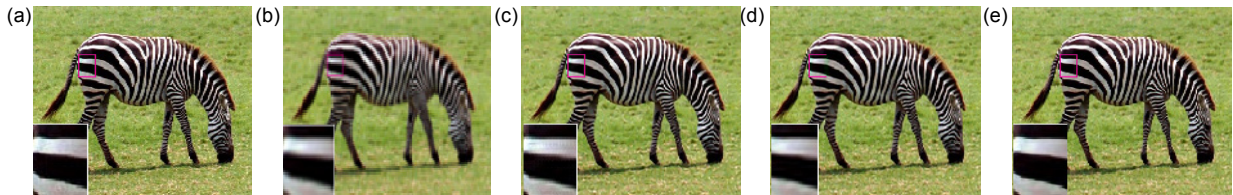


图 7 Zebra 放大 3 倍时各算法重建效果。(a) 原图; (b) Bicubic; (c) SRCNN; (d) DnCNN-3; (e) 本文方法  
Fig. 7 Reconstruction results of each algorithm for zebra amplification by 3 times. (a) Original; (b) Bicubic; (c) SRCNN; (d) DnCNN-3; (e) Our method

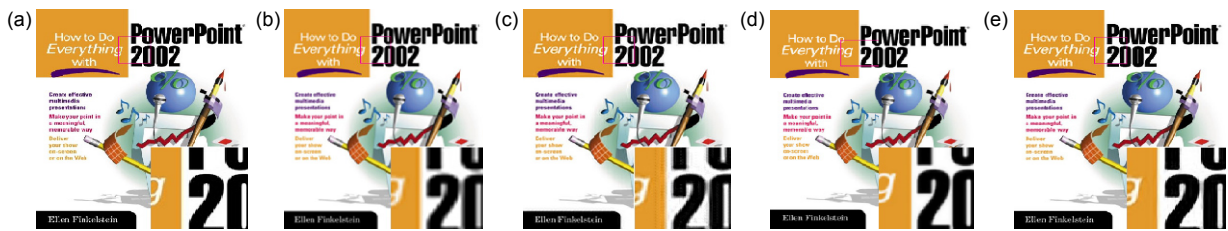


图 8 PPT 放大 4 倍时各算法重建效果。(a) 原图; (b) Bicubic; (c) SRCNN; (d) DnCNN-3; (e) 本文方法  
Fig. 8 Reconstruction results of each algorithm for PPT amplification by 4 times. (a) Original; (b) Bicubic; (c) SRCNN; (d) DnCNN-3; (e) Our method

多尺度特征损失函数得到的效果图对比。通过观察图 5 可知，采用本文提出的多尺度特征损失函数在一定程度上能够改善重建图像的模糊问题，提升图像在细节方面的重建效果。图 6 至图 8 分别展示了本文方法和 Bicubic、SRCNN、DnCNN-3 以及原始高清图像的对比。

图 9 给出了本文方法在训练集和测试集上的损失函数变化曲线。从图中可知，训练集的损失函数变化曲线存在一定程度的震荡，但是随着训练的进行，损失曲线逐渐趋于收敛，模型得到优化，震荡现象也得到改善，而测试集中损失函数曲线已十分平滑。并且测试集的损失函数曲线略微小于训练集的损失函数曲线，说明模型能够较好地预防过拟合问题，即不存在模型在训练集上表现过于优越而在测试集上表现不佳的问题。

## 5 结论

本文在基于 DenseNet 的图像超分辨率重建方法

的基础上加以扩展，对传统的均方差损失函数进行改性，在重建模型后额外串联一个卷积神经网络用来计算重建图像与对应的原始高清图像在卷积后得到不同尺度特征图的均方差，以此来优化整个网络从底层特征到高层特征的学习。最后通过实验证明了本文的方法在主观视觉效果和客观图像质量评价指标上均

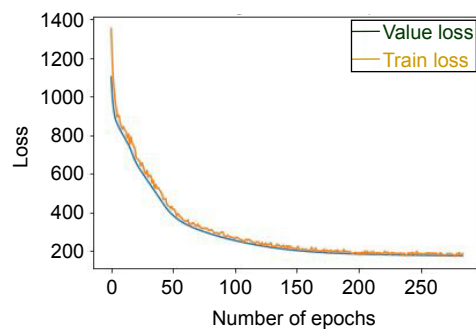


图 9 本文方法在训练集和测试集上的损失函数曲线

Fig. 9 The loss function curves of our method in training set and test set



有所改善和提高。由于本文采用上采样的方式来放大图像尺寸,在放大倍数较大时,本文的方法略显不足,在今后的研究中需要解决。

## 参考文献

- [1] Glasner D, Bagon S, Irani M. Super-resolution from a single image[C]//*Proceedings of the IEEE 12th International Conference on Computer Vision*, Kyoto, Japan, 2009: 349–356.
- [2] Peled S, Yeshurun Y. Superresolution in MRI: application to human white matter fiber tract visualization by diffusion tensor imaging[J]. *Magnetic Resonance in Medicine*, 2001, **45**(1): 29–35.
- [3] Shi W Z, Caballero J, Ledig C, et al. Cardiac image super-resolution with global correspondence using multi-atlas patchmatch[C]//*Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention*, Nagoya, Japan, 2013: 9–16.
- [4] Gunturk B K, Batur A U, Altunbasak Y, et al. Eigenface-domain super-resolution for face recognition[J]. *IEEE Transactions on Image Processing*, 2003, **12**(5): 597–606.
- [5] Zhang L P, Zhang H Y, Shen H F, et al. A super-resolution reconstruction algorithm for surveillance images[J]. *Signal Processing*, 2010, **90**(3): 848–859.
- [6] Zhou F, Yang W M, Liao Q M. Interpolation-based image super-resolution using multisurface fitting[J]. *IEEE Transactions on Image Processing*, 2012, **21**(7): 3312–3318.
- [7] 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.
- [8] Lin Z C, Shum H Y. Fundamental limits of reconstruction-based superresolution algorithms under local translation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, **26**(1): 83–97.
- [9] Rasti P, Demirel H, Anbarjafari G. Image resolution enhancement by using interpolation followed by iterative back projection[C]//*Proceedings of the 21st Signal Processing and Communications Applications Conference*, Haspolat, Turkey, 2013: 1–4.
- [10] Zhou J H, Zhou C, Zhu J J, et al. A method of super-resolution reconstruction for remote sensing image based on non-subsampled contourlet transform[J]. *Acta Optica Sinica*, 2015, **35**(1): 0110001.  
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, **35**(1): 0110001.
- [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] Timofte R, De V, van Gool L. Anchored neighborhood regression for fast example-based super-resolution[C]//*Proceedings of 2013 IEEE International Conference on Computer Vision*, Sydney, NSW, Australia, 2013: 1920–1927.
- [13] 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.
- [14] Su H, Zhou J, Zhang Z H. Survey of super-resolution image reconstruction methods[J]. *Acta Automatica Sinica*, 2013, **39**(8): 1202–1213.  
苏衡, 周杰, 张志浩. 超分辨率图像重建方法综述[J]. *自动化学报*, 2013, **39**(8): 1202–1213.
- [15] Keys R. Cubic convolution interpolation for digital image processing[J]. *IEEE Transactions on Acoustics Speech & Signal Processing*, 1981, **29**(6): 1153–1160.
- [16] Yuan Q, Jing S X. Improved super resolution reconstruction method for video sequence[J]. *Journal of Computer Applications*, 2009, **29**(12): 3310–3313.  
袁琪, 荆树旭. 改进的序列图像超分辨率重建方法[J]. *计算机应用*, 2009, **29**(12): 3310–3313.
- [17] Chang H, Yeung D Y, Xiong Y. Super-resolution through neighbor embedding[C]//*Proceedings of the 2004 Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2004, **1**: I–I.
- [18] 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.
- [19] Yang J, Wright J, Huang T, et al. Image super-resolution as sparse representation of raw image patches[C]//*Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2008: 1–8.
- [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*, Amsterdam, The Netherlands, 2016.
- [21] Kim J W, 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*, Las Vegas, NV, USA, 2016: 1646–1654.
- [22] Tong T, Li G, Liu X J, et al. Image super-resolution using dense skip connections[C]//*Proceedings of 2017 IEEE International Conference on Computer Vision*, Venice, Italy, 2017.
- [23] Yamanaka J, Kuwashima S, Kurita T. Fast and accurate image super resolution by deep CNN with skip connection and network in network[C]//*Proceedings of the 24th International Conference on Neural Information Processing*, Guangzhou, China, 2017.
- [24] 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*, Santiago, Chile, 2015: 1026–1034.
- [25] 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*, Honolulu, Hawaii, 2017.



# Image super-resolution reconstruction based on multi-scale feature loss function

Xu Liang, Fu Randi\*, Jin Wei, Tang Biao, Wang Shangli

School of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China

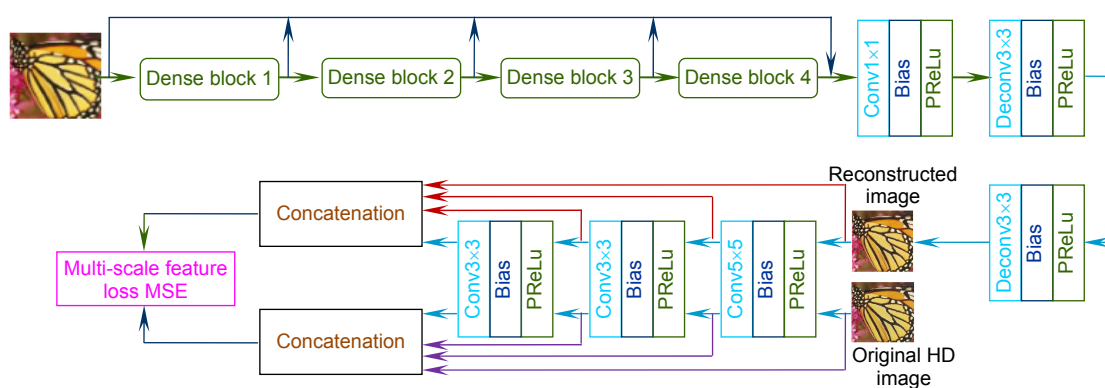


Image super-resolution reconstruction model based on multi-scale feature loss function

**Overview:** In recent years, with the research and development of deep learning, it has been widely used in image processing. Compared with the traditional shallow learning, which can only extract the features of images simply, deep learning can learn the deeper feature representation, so as to have better performance in image processing. The traditional mean squared error (MSE) as the loss function is mostly adopted in the image super-resolution based on deep learning to obtain better PSNR, such as SRCNN, FSRCNN, and SRDenseNet. However, the reconstructed image is prone to edge blur and may be too smooth. The multi-scale loss function proposed in this paper can improve this problem. Based on the analysis of SRCNN, FSRCNN, SRDenseNet, and other methods, the reconstruction model was built with the DenseNet model as the basic framework, and a three-layer convolutional neural network was connected after the reconstruction model to calculate the multi-scale feature loss function. The reconstruction model consists of four parts: dense connection block, dimension reduction layer, deconvolution layer, and reconstruction layer. Each dense connection block is composed of 4 convolution layers, and  $3 \times 3$  convolution kernel is adopted. The number of feature maps output by each dense connection block is 256. Since the output of all dense connection blocks is concatenated, the feature map is reduced to 256 by means of  $1 \times 1$  convolution kernel in the dimension reduction layer to reduce the computational burden. After the deconvolution layer, a single channel image is reconstructed by  $3 \times 3$  convolution kernel. At last, the reconstructed image and the corresponding original HD image were extracted by the three-layer convolution neural network in series, and the difference between the reconstructed image and the original HD image was compared by calculating the mean square error. This article uses Yang91 and BSD200 dataset that consists of 291 images. Considering that the training of convolution neural network depends on a large number of data samples, the original 291 data sets are extended to ensure sufficient samples. First, the original sample set was flipped from left to right and from top to bottom, and the training sample set was 4 times more than the original one, obtaining  $291 + (291 \times 4) = 1455$  training samples. Then, the original sample size is enlarged by 2, 3, and 4 times, respectively, with further  $180^\circ$  mirror transformation. After that,  $291 \times 2 \times 3 = 1746$  training samples were obtained, with total samples  $1455 + 1746 = 3201$ . Set5, Set14 and BSD100 were selected as the standard evaluation dataset in the field of super-resolution research for the test samples, and objective indicators were evaluated using peak signal to noise ratio (PSNR) and structural similarity (SSIM). The experimental results show that the details of the reconstructed images become richer and the edge blur is improved.

**Citation:** Xu L, Fu R D, Jin W, *et al.* Image super-resolution reconstruction based on multi-scale feature loss function[J]. *Opto-Electronic Engineering*, 2019, 46(11): 180419

Supported by National Natural Science Foundation of China (61471212) and Zhejiang Province Natural Science Fund (LY16F010001)

\* E-mail: 17346807657@163.com