计算机系统应用 ISSN 1003-3254, CODEN CSAOBN
Computer Systems & Applications,2022,31(7):217−223 [doi: 10.15888/j.cnki.csa.008606]
©中国科学院软件研究所版权所有.
E-mail: csa@iscas.ac.cn
http://www.c-s-a.org.cn
Tel: +86-10-62661041

# 基于多尺度融合 CNN 的图像超分辨率重建①

祝旭阳[1], 于俊洋[1], 郝艳艳[2]

[1](河南大学 软件学院, 开封 475004)
[2](河南工业贸易职业学院 信息工程学院, 郑州 451191)
通信作者: 于俊洋, E-mail: jyyu@henu.edu.cn

**摘　要**: 针对 FSRCNN 模型中存在的特征提取不充分和反卷积带来的人工冗余信息的问题, 本文提出了一种基于多尺度融合卷积神经网络的图像超分辨率重建算法. 首先设计了一种多尺度融合的特征提取通道, 解决对图像不同尺寸信息利用不充分问题; 其次在图像重建部分, 采用子像素卷积进行上采样, 抑制反卷积层带来的人工冗余信息. 与 FSRCNN 模型相比, 在 Set5 和 Set14 数据集中, 2 倍放大因子下的 *PSNR* 值和 *SSIM* 值平均提高了 0.14 dB、0.001 0, 在 3 倍放大因子下平均提高 0.48 dB、0.009 1. 实验结果表明, 本文算法可以更大程度的保留图像纹理细节, 提升图像整体重建效果.

**关键词**: 图像超分辨率重建; 卷积神经网络 (CNN); 多尺度特征融合; 深度学习

引用格式: 祝旭阳,于俊洋,郝艳艳.基于多尺度融合 CNN 的图像超分辨率重建.计算机系统应用,2022,31(7):217–223. http://www.c-s-a.org.cn/1003-3254/8606.html

## Image Super-resolution Reconstruction Based on Multi-scale Fusion CNN

ZHU Xu-Yang[1], YU Jun-Yang[1], HAO Yan-Yan[2]

[1](School of Software, Henan University, Kaifeng 475004, China)
[2](School of Information Engineering, Henan Industry and Trade Vocational College, Zhengzhou 451191, China)

**Abstract**: To address the insufficient feature extraction of the FSRCNN model and its artificial redundant information caused by deconvolution, this study proposes an image super-resolution reconstruction algorithm based on a multi-scale fusion convolutional neural network. Specifically, a multi-scale fusion feature extraction channel is designed to cope with the insufficient utilization of image information of different sizes. Then, sub-pixel convolution is used for up-sampling in the image reconstruction part to suppress the artificial redundant information caused by deconvolution. Compared with the FSRCNN model, the algorithm in this study respectively increases the *PSNR* and *SSIM* by 0.14 dB and 0.001 0 on average at an amplification factor of 2 and by 0.48 dB and 0.009 1 on average at an amplification factor of 3 on Set5 and Set14 data sets. Experimental results show that the proposed algorithm can better retain the texture details of the image and improve the overall image reconstruction effect.

**Key words**: image super-resolution reconstruction; convolutional neural network (CNN); multi-scale feature fusion; deep learning

## 1　引言

超分辨率重建 (super-resolution, SR) 技术旨在通过特定的约束函数把低分辨率图像重建为高分辨率图像. 由于图像生成容易受到设备, 天气, 光照和环境等因素的干扰只能获得低分辨率图像, 然而高分辨率图像往往含有更丰富的信息, 具有极大价值. 提高图像分辨率的最直接手段就是改善成像设备, 但是这种方式会加重个人和中小企业的成本负担. 而图像超分辨率

① 基金项目: 河南省科技研发项目 (212102210078, 212102310548)
收稿时间: 2021-10-22; 修改时间: 2021-12-02; 采用时间: 2021-12-13; csa 在线出版时间: 2022-05-31



© 中国科学院软件研究所 http://www.c-s-a.org.cn

重建技术是通过编程算法对低分辨率图像进行处理得到高分辨率图像, 这种方法成本极低, 因此图像超分辨率技术被广泛应用于智能监控、国防军事、医学图像和遥感卫星等领域中.

图像超分辨率重建技术方法大致可以分为 3 类: 基于插值[1], 基于重建[2] 和基于学习[3]. 基于插值的方法通过子像素和周围像素的空间位置关系利用数学公式推导插入新像素点, 基于插值的方法有最近邻插值, 双线性插值和双三次插值 (bicubic), 这种插值方法实现简单快捷. 基于重建的方法是对图像的整体信息通过数学方式进行结合然后重建出高分辨率图像, 常用的方法有迭代反投影法 (IBP)[4], 凸集投影法 (POCS)[5] 和最大后验概率估计法 (MAP)[6]. 基于学习[7–9] 的图像超分辨率重建方法通过学习低分辨率图像和高分辨率图像之间的映射关系, 得到大量先验知识来指导图像重建. 上述都属于基于传统方法, 算法简单, 处理速度快, 但是存在着边缘纹理模糊, 图像效果不好的问题.

近年来, 通过结合卷积神经网络, 学习低分辨率图像和高分辨率图像之间的复杂关系, 达到了更好的重建效果. Dong 等人[10] 提出了深度卷积网络的图像超分辨率 (super-resolution convolutional neural network, SRCNN) 重建模型, 结合卷积神经网络通过端到端的训练方式, 大大提高了重建图像质量, 但是由于 SRCNN 在网络模型中使用双三次插值的方法实现对图像的放大操作, 会破坏图像原有的细节信息, 导致重建效果还不够理想. 因此在 SRCNN 的基础上, Dong 等人[11] 提出了加速图像超分辨率重建的卷积神经网络 (fast super-resolution convolutional neural networks, FSRCNN) 模型, 在网络层最后引入反卷积层代替双三次插值对输入图像进行尺寸放大, 并且使用更小的卷积核和更多的卷积层达到加快模型速度的目的. Shi 等人[12] 提出了一种基于亚像素卷积网络的 ESPCN, 将亚像素卷积引入到图像上采样工作中, 提高了图像重建效果.

本文针对 FSRCNN 模型中存在的图像特征信息提取不充分和反卷积上采样带来的人工冗余信息问题, 提出了一种多尺度融合卷积神经网络, 首先在浅层特征提取阶段设计了一种多尺度融合通道, 增强对图像中的高频有用信息特征的提取能力; 其次在图像重建部分使用亚像素卷积代替反卷积层进行上采样, 避免反卷积层的人工冗余信息引入. 实验结果表明, 与 FSRCNN 模型相比, 在 Set5 和 Set14 数据集中, 2 倍放大因子下的 *PSNR* 值和 *SSIM* 值平均提高了 0.14 dB、0.001 0, 在 3 倍放大因子下平均提高 0.48 dB、0.009 1.

## 2 相关工作

### 2.1 FSRCNN 模型

图 1 为 FSRCNN 的网络结构, 其中, Conv($x$,$y$,$z$) 中, $x$ 代表卷积核大小, $y$ 代表卷积层输出通道, $z$ 代表卷积层输入通道. 分为 5 层: 第 1 层对低分辨率图像进行特征提取, 卷积核大小是 5, 输入通道为 1, 输出通道是 56. 第 2 层是缩减层, 减少模型的参数量, 提高算法速度. 第 3 层是学习映射层, 对图像信息进行非线性学习. 第 4 层是扩大层, 增加图像特征图数量. 第 5 层是反卷积层, 对图像上采样得到重建图像.

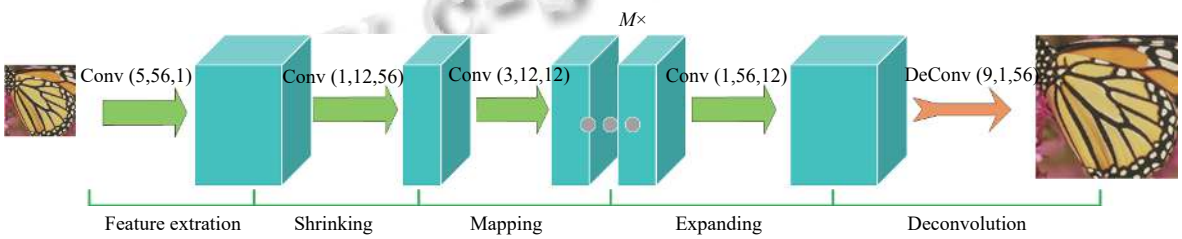



图 1 FSRCNN 的网络结构

### 2.2 子像素卷积

子像素卷积又称像素重排序 (pixel shuffle), 像素重排是将不同通道的特征图排序到同一个通道上, 以此达到对图像的像素增多的目的. 如图 2 为子像素卷积对图像重建过程, 假设低分辨率图像大小为 3×3 大小, 若想得到一个 9×9 大小的图像, 即需要对原本的图像放大 3 倍, 首先使用 1×1 大小的卷积核对原低分辨率图像进行卷积操作, 生成$3^2$=9 个 3×3 大小的特征图, 然后将这 9 个相同尺寸特征图进行像素的重新排列, 拼成一个放大 3 倍的高分辨率图像.



© 中国科学院软件研究所 http://www.c-s-a.org.cn

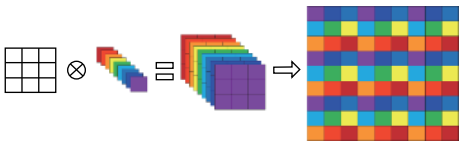


图 2 子像素卷积

## 3 基于多尺度融合卷积神经网络的图像超分辨率重建算法

### 3.1 网络结构

如图 3 所示为本文网络结构, 分为 5 层: 第 1 层为本文算法设计的一种多尺度特征融合提取通道, 引用不同大小尺寸的卷积核对低像素图像进行特征提取, 提高对图像中不同尺度的高频有用信息的提取能力. 第 5 层为子像素卷积重建层, 引入了子像素卷积的方法对图像进行重建, 抑制反卷积层带来的人工信息冗余问题, 提高图像的重建质量.

为实现图像超分辨率重建, 需要对算法卷积神经网络模型进行训练. 如图 4 所示, 训练过程首先将高分辨率图像下采样得到低分辨率图像, 将高分辨率图像和低分辨率图像分别切割, 获得若干成对对应训练样本, 将高分辨率图像和低分辨率图像成对输入卷积神经网络中进行训练, 通过损失函数指导卷积神经网络更新权重, 经过多轮迭代, 最终得到图像超分辨率重建模型. 在测试阶段, 将低分辨率图像送入训练完成的模型, 即可重建出高分辨率图像.

综上, 本文在 FSRCNN 模型的基础上设计了一种了多分支融合特征提取通道, 提高算法对图像有用信息的提取能力; 在图像重建部分, 使用子像素卷积代替反卷积, 实现对图像的重建, 抑制人工冗余信息的引入.

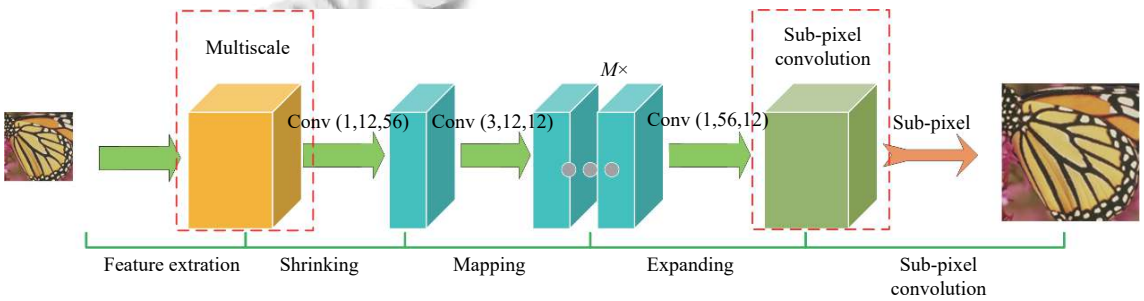



图 3 本文结构

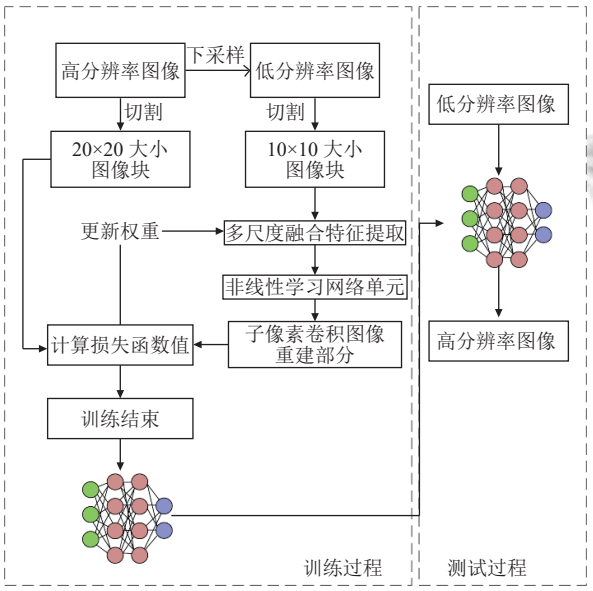



图 4 模型构建和训练过程

### 3.2 多尺度特征融合提取通道

在 FSRCNN 模型使用了一层 56 个 5×5 大小卷积核的卷积层进行特征提取工作, 本文设计了一种多尺度融合特征提取通道, 结合跳跃连接和特征融合操作, 使用 3 个不同大小的卷积核分别对输入图像进行特征提取. 针对图像的易模糊边缘轮廓和细节特征的提取难度问题, 如图 5 所示为多尺度特征提取通道的结构图. 第一个通道为 56 个 1×1 大小的卷积核, 获得图像的细节特征, 具体实现公式为:

$$F_{11}(Y) = W_{11} \times I_{LR} + B_{11} \tag{1}$$

其中, $F_{11}(Y)$ 表示第一通道卷积层输出的特征图, $W_{11}$ 表示第一通道卷积层的权重, $B_{11}$ 表示偏置项.

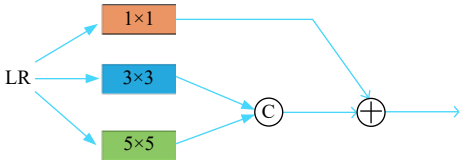



图 5 多尺度特征提取通道结构

第 2 通道和第 3 通道分别为 28 个 3×3 和 5×5 大小卷积核进行特征提取并拼接到一起, 第 2 通道的具



© 中国科学院软件研究所 http://www.c-s-a.org.cn

体实现公式为:

$$F_{12}(Y) = W_{12} \times I_{LR} + B_{12} \tag{2}$$

其中, $F_{12}(Y)$表示第 2 通道卷积层输出的特征图, $W_{12}$表示第 2 通道卷积层的权重, $B_{12}$表示偏置项. 第 3 通道的具体实现公式为:

$$F_{13}(Y) = W_{13} \times I_{LR} + B_{13} \tag{3}$$

其中, $F_{13}(Y)$表示第 3 通道卷积层输出的特征图, $W_{13}$表示第 3 通道卷积层的权重, $B_{13}$表示偏置项.

由于第 2 通道和第 3 通道分别采用了 3×3 和 5×5 的卷积核进行特征提取, 不同尺寸的卷积核可以获得不同尺度的特征信息, 通过 Contact 操作将不同的特征信息进行拼接, 具体实现公式为:

$$F_{23}(Y) = F_{12}(Y) + F_{13}(Y) \tag{4}$$

之后将不同尺度的特征信息通过长跳跃连接 (LSC) 操作进行特征融合从而保证图像信息的流通性和完整性. 具体实现公式为:

$$F_1(Y) = W_{14}(F_{12}(Y) + F_{13}(Y)) + B_1 \tag{5}$$

其中, $F_1(Y)$表示特征融合后得到的特征图, $W_{14}$表示特征融合层的权重, $B_1$表示该层的偏置项. 相比 FSR-CNN 模型[11] 仅有一个 5×5 大小的卷积核作为特征提取通道, 多尺度特征融合特征提取通道对图像的不同尺度的信息提取更充分, 加强了算法网络模型的特征提取能力.

### 3.3 子像素卷积上采样

在算法的图像重建部分, 反卷积操作实现了对图像的放大, 在重建过程中使用反卷积填充图像内容, 使得图像内容变得丰富, 然而反卷积存在过多人工冗余信息引入的问题, 为了解决这个问题, 我们在图像重建部分引用子像素卷积替代反卷积实现对图像的上采样过程.

如图 6 所示, 首先使用 1×1 的卷积对扩大图像输入通道, 然后通过像素重排扩大图像的分辨率. 若子像素卷积重建之前有$r \times r$个通道特征图, 则通过使用子像素卷积重建会将$r \times r$个特征图通过特定的排序规则重新生成一副$w \times r, h \times r$的图像.

图 6 上采样路径示意图

### 3.4 损失函数和激活函数

本文算法采用 *MSE* 损失函数, 即均方损失函数. 经过超分辨率重建神经网络重建得到的图像$f(x_i)$和真值图像$y_i$, 通过对网络模型进行训练使损失函数的值减小从而指导低像素图像重建出更好效果的高分辨率图像. *MSE* 损失函数如式 (6) 所示:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(y_i - f(x_i))^2 \tag{6}$$

使用的激活函数为 PReLU[13], 即带参数的 ReLU 函数, 避免 ReLU 函数中零梯度坏死特性, 并且还可以实现更高的准确率, PReLU 激活函数表达式如式 (7):

$$\rho(Y_i) = \begin{cases} Y_i, & x > 0 \\ \lambda Y_i, & x \leqslant 0, \lambda \in [0,1] \end{cases} \tag{7}$$

其中, $Y_i$ 是第 $i$ 通道的输出, $\lambda$ 为 PReLU 的参数, 且会随着训练次数的增加而不断更新.

如图 7, 图 8 所示为本文算法训练阶段的损失函数收敛曲线图. 放大因子为×2 时, 在迭代到 17 时, 损失函数收敛. 当放大因子为×3 时, 训练迭代到 18 代时, 损失函数收敛.

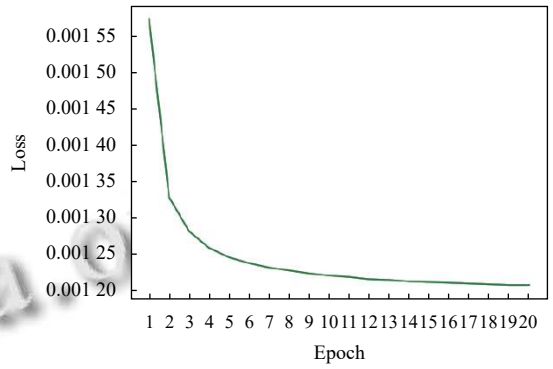



图 7 放大因子为×2 时损失函数收敛图

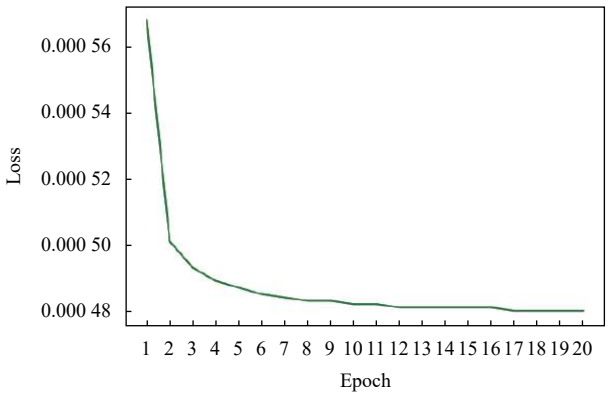



图 8 放大因子为×3 时损失函数收敛图



© 中国科学院软件研究所 http://www.c-s-a.org.cn

## 4 实验

### 4.1 实验数据以及设置

实验用 91-image 数据集对模型进行训练, 首先对输入的图片进行数据增强, 将图像旋转 90°、180°、270°, 并缩小 0.9、0.8、0.7、0.6 倍, 实现对数据集扩充的目的.

测试集: 选用 Set5[14] 和 Set14[15] 公开数据集作为测试集进行效果度量.

本文算法采用 Adam (Adam optimization algorithm) 来优化损失函数, 学习率设置为 0.000 1, 训练的批处理大小为 16, 训练迭代次数设置为 20 次. 深度学习框架为 PyTorch 1.2, 实验的操作系统是 Ubuntu 16.04, 服务器是内存 8 GB 的 Tesla P4.

### 4.2 评价指标

本文采用峰值信噪比 (*PSNR*)[16] 和结构相似度 (*SSIM*)[17] 作为重建效果的评价指标, *PSNR* 值越大证明重建图像比真值图像失真越少. *PSNR* 的公式如式 (9):

$$MSE = \frac{1}{HW}\sum_{i=1}^{H}\sum_{j=1}^{W}[X(i,j)-Y(i,j)]^2 \tag{8}$$

$$PSNR = 10\log_{10}\left(\frac{(2^n-1)^2}{MSE}\right) \tag{9}$$

其中, *MSE* 表示均方误差, $H$和$W$分别表示图像的长度和宽度, $X(i,j)$表示真值图像, $Y(i,j)$表示重建后的图像. 峰值信噪比的单位为 dB, 值越大表示重建效果越好.

*SSIM* 也常用来评价图像重建效果, *SSIM* 用来评价重建图像与真值图像的结构相似性, 值越大代表相似性越高, 重建效果越好. *SSIM* 的计算公式为:

$$SSIM(H,L) = \frac{(2\mu_H\mu_L + c_1)(2\sigma_{HL} + c_2)}{(\mu_H^2 + \mu_L^2 + c_1)(\sigma_H^2 + \sigma_L^2 + c_2)} \tag{10}$$

其中, $\mu_H$、$\mu_L$分别表示图像$H$和$L$的均值, $\sigma_H$、$\sigma_L$分别表示图像$H$和$L$的标准差, $\sigma_H^2$、$\sigma_L^2$分别表示图像$H$和$L$的方差, $\sigma_{HL}$表示图像$H$和$L$协方差. $c_1$, $c_2$和$c_3$为常数.

### 4.3 与其它方法比较

本文在 Set5 和 Set14 数据集上, 对比了 Bicubic、SRCNN[10]、FSRCNN[11]. 由表 1 和表 2 可以看出, 从客观评价标准出发, 对比不同算法, 本文算法的重建质量优于其他几种算法. 相较于 FSRCNN 算法, 在 Set5 测试集上, 当放大因子为×2 和×3 时, *PSNR* 值分别提高了 0.15 dB 和 0.57 dB. 在 Set14 测试集上分别提高了 0.13 dB、0.39 dB. 在 Set5 测试集上, 当放大因子为×2 和×3 时, *SSIM* 值分别提高了 0.001 9 和 0.010 8. 在 Set14 测试集上, *SSIM* 值分别提高了 0.000 2 和 0.007 5.

表 1 不同重建算法的 *PSNR* 结果 (dB)

| 数据集 | 放大因子 | Bicubic | SRCNN | FSRCNN | 本文算法 |
|---|---|---|---|---|---|
| Set5 | ×2 | 33.66 | 36.33 | 36.94 | **37.09** |
| | ×3 | 30.29 | 32.45 | 33.06 | **33.63** |
| Set14 | ×2 | 30.23 | 32.15 | 32.54 | **32.67** |
| | ×3 | 27.54 | 29.01 | 29.37 | **29.76** |

表 2 不同重建算法的 *SSIM* 结果

| 数据集 | 放大因子 | Bicubic | SRCNN | FSRCNN | 本文算法 |
|---|---|---|---|---|---|
| Set5 | ×2 | 0.929 9 | 0.952 1 | 0.953 2 | **0.955 1** |
| | ×3 | 0.866 5 | 0.903 3 | 0.905 5 | **0.916 3** |
| Set14 | ×2 | 0.868 7 | 0.903 9 | 0.905 2 | **0.905 4** |
| | ×3 | 0.773 6 | 0.814 5 | 0.816 7 | **0.824 2** |

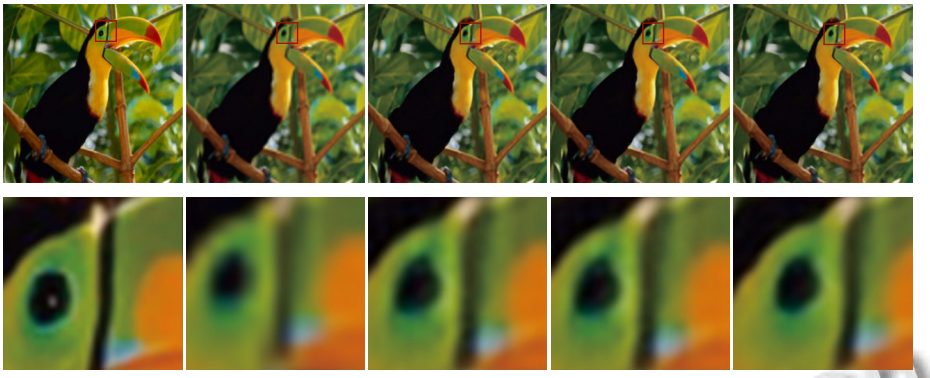
如图 9 和图 10 所示, 图像通过不同的超分辨率重建算法获得了不同的效果图, 可以看到, 相较于传统的 Bicubic 算法, SRCNN 对图像重建质量提升了很多, FSRCNN 改善了图像的纹理细节并且避免了噪声对图像细节的影响, 而本文所提出的算法较之原来的方法保留更多的图像细节特征, 这主要是得益于多尺度特征融合通道对特征提取能力的提高和子像素卷积对人工冗余信息的抑制作用.

## 5 结语

本文提出了一种基于多尺度卷积神经网络的图像超分辨率重建算法, 在算法的特征提取阶段, 设计了多通道多尺度特征提取方法, 分别使用 1×1、3×3 和 5×5 大小的卷积核, 对图像的不同特征运用不同尺寸的卷积操作, 并结合跳跃链接和特征融合操作, 提高对图像特征信息的感知能力. 在图像重建阶段, 使用子像素卷积避免反卷积层引用过多的人工冗余信息, 提高重建图像的质量. 下一步将研究在网络结构中结合图像的全局信息和局部信息, 在训练阶段充分利用高分辨率图像中的高频有用信息指导图像重建过程, 从而提高重建图像的质量.



© 中国科学院软件研究所 http://www.c-s-a.org.cn



(a) Original (b) Bicubic (c) SRCNN (d) FSRCNN (e) Ours

图 9 不同算法重建结果图对比

(a) Original (b) Bicubic (c) SRCNN (d) FSRCNN (e) Ours

图 10 不同算法重建效果图对比

## 参考文献


1 张芳, 赵东旭, 肖志涛, 等. 单幅图像超分辨率重建技术研究进展. 自动化学报, 2021, 47(X): 10001–10021.

2 Xi ZH, Yuan KP. Super-resolution image reconstruction based on residual channel attention and multilevel feature fusion. Laser & Optoelectronics Progress, 2020, 57(4): 041504.

3 Yang WM, Zhang XC, Tian YP, *et al*. Deep learning for single image super-resolution: A brief review. IEEE Transactions on Multimedia, 2019, 21(12): 3106–3121. [doi: 10.1109/TMM.2019.2919431]

4 Qiu DF, Cheng YH, Wang XS. Gradual back-projection residual attention network for magnetic resonance image super-resolution. Computer Methods and Programs in Biomedicine, 2021, 208: 106252. [doi: 10.1016/j.cmpb.2021.106252]

5 庞亚红, 毛幼菊. 一种基于凸集投影 (POCS) 的数字图像超分辨率重建算法. 计算机工程与应用, 2005, 41(4): 69–71. [doi: 10.3321/j.issn:1002-8331.2005.04.022]

6 Jakhetiya V, Lin WS, Jaiswal SP, *et al*. Maximum a posterior and perceptually motivated reconstruction algorithm: A generic framework. IEEE Transactions on Multimedia, 2017, 19(1): 93–106. [doi: 10.1109/TMM.2016.2609419]

7 龚文斌, 石章松, 韦华. 高分辨率遥感卫星图像不均匀云雾去除算法. 激光与光电子学进展, 2020, 57(6): 061504.

8 Ha VK, Ren JC, Xu XY, *et al*. Deep learning based single image super-resolution: A survey. International Journal of Automation and Computing, 2019, 16(4): 413–426. [doi: 10.1007/s11633-019-1183-x]

9 Yang JC, Wright J, Huang TS, *et al*. Image super-resolution via sparse representation. IEEE Transactions on Image Processing, 2010, 19(11): 2861–2873. [doi: 10.1109/TIP.2010.




© 中国科学院软件研究所 http://www.c-s-a.org.cn

2050625]

10 Dong C, Loy CC, He KM, *et al*. Learning a deep convolutional network for image super-resolution. Proceedings of the 13th European Conference on Computer Vision. Zurich: Springer, 2014. 184–199.

11 Dong C, Loy CC, Tang XO. Accelerating the super-resolution convolutional neural network. Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016. 391–407.

12 Shi WZ, Caballero J, Huszár F, *et al*. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 1874–1883.

13 Thakur RS, Yadav RN, Gupta L. PReLU and edge-aware filter-based image denoiser using convolutional neural network. IET Image Processing, 2020, 14(15): 3869–3879. [doi: 10.1049/iet-ipr.2020.0717]

14 Bevilacqua M, Roumy A, Guillemot C, *et al*. Single-image super-resolution via linear mapping of interpolated self-examples. IEEE Transactions on Image Processing, 2014, 23(12): 5334–5347. [doi: 10.1109/TIP.2014.2364116]

15 Zhang KB, Gao XB, Tao DC, *et al*. Single image super-resolution with non-local means and steering kernel regression. IEEE Transactions on Image Processing, 2012, 21(11): 4544–4556. [doi: 10.1109/TIP.2012.2208977]

16 Huynh-Thu Q, Ghanbari M. Scope of validity of PSNR in image/video quality assessment. Electronics Letters, 2008, 44(13): 800–801. [doi: 10.1049/el:20080522]

17 Wang Z, Bovik AC, Sheikh HR, *et al*. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 2004, 13(4): 600–612. [doi: 10.1109/TIP.2003.819861]

(校对责编: 孙君艳)

www.c-s-a.org.cn



© 中国科学院软件研究所 http://www.c-s-a.org.cn