



计算机应用
Journal of Computer Applications
ISSN 1001-9081, CN 51-1307/TP

《计算机应用》网络首发论文

题目：基于空间元学习的放大任意倍的超分辨率重建方法
作者：孙忠凡，周正华，赵建伟
收稿日期：2020-06-12
网络首发日期：2020-10-23
引用格式：孙忠凡，周正华，赵建伟. 基于空间元学习的放大任意倍的超分辨率重建方法[J/OL]. 计算机应用.
<https://kns.cnki.net/kcms/detail/51.1307.TP.20201022.1005.010.html>



网络首发：在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

出版确认：纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

基于空间元学习的放大任意倍的超分辨率重建方法

孙忠凡, 周正华, 赵建伟*

(中国计量大学 信息与数学系, 杭州 310018)

(*通信作者电子邮箱 zhaojw@amss.ac.cn)

摘要: 针对现有的基于深度学习的超分辨率重建方法主要研究放大整数倍的重建, 对放大任意倍即非整数倍重建情况讨论较少的问题, 提出一种基于空间元学习的放大任意倍的超分辨率重建方法。首先, 该方法利用坐标投影找出高分辨率图像与低分辨率图像坐标间的对应关系。其次, 在元学习网络的基础上, 考虑特征图的空间信息, 将提取的空间特征与坐标位置相结合作为权值预测网络的输入。最后, 将权值预测网络预测出的卷积核与特征图结合, 有效地放大特征图的尺寸, 得到放大任意倍的高分辨率图像。所提的空间元学习模块可以与其他深度网络相结合, 得到放大任意倍的超分辨率图像重建方法。本文提供了任意倍(非整数倍)超分辨率重建方法, 解决了实际生活中放大尺寸固定且非整数倍的重建问题。实验结果表明, 所提的重建方法在空间复杂度(网络参数)相当的情况下, 时间复杂度(计算量)是其他重建方法的 25%~50%, 且峰值信噪比(PSNR)比其他一些方法提高了 0.01~5、结构相似度(SSIM)提高了 0.03~0.11。

关键词: 超分辨率; 深度学习; 空间元学习; 残差密集模块; 权值预测

中图分类号: TP391

文献标志码: A

Super-resolution reconstruction method with arbitrary magnification based on spatial meta-learning

SUN Zhongfan, ZHOU Zhenghua, ZHAO Jianwei *

(Department of Information and Mathematics, China Jiliang University, Zhejiang Hangzhou 310008, China)

Abstract: For the problem that the existing deep-learning based super-resolution reconstruction methods mainly study the reconstruction problem of amplifying integer times, not on the case of amplifying arbitrary times (e.g. non-integer times), a super-resolution reconstruction method based on spatial meta-learning was proposed. Firstly, a coordinate projection was used to find the correspondence between the coordinates of the high resolution image and the low resolution image. Secondly, based on the meta-learning network, considering the spatial information of the feature map, the extracted spatial features and coordinate positions were combined as the input of the weighted prediction network. Finally, the convolution kernels predicted by the weighted prediction network were combined with the feature map to amplify the size of the feature maps effectively, and obtain the high resolution image with arbitrary times. The proposed spatial meta-learning module can be combined with other deep networks to obtain super-resolution reconstruction methods with arbitrary magnification. This paper provides a super-resolution reconstruction method with arbitrary magnification, which solves the reconstruction problem with a fixed but non-integer scale in the real life. The experimental results show that when the space complexity (network's parameters) is equivalent, the time complexity of proposed method is 25%-50% of the other reconstruction methods, the Peak-Signal-to-Noise Ratio (PSNR) of proposed method is 0.01-5 higher than the others, and the structural similarity (SSIM) of proposed method is 0.03-0.11 higher than the others.

Keywords: super-resolution; deep learning; spatial meta-learning; residual dense module; weight prediction

在医学图像处理^[1]、航空影像^[2]、安全监测等领域有着重要的应用。

0 引言

图像超分辨率重建的主要任务是从一张或多张低分辨率图像中重建出高分辨率图像, 是机器视觉中的研究热点之一,

目前, 人们基于插值、基于学习方法^[3]和基于重建方法, 提出了许多高效的超分辨率重建方法。特别地, 随着深度学习的快速发展和广泛应用, 超分辨率重建得到了迅速发展。

收稿日期:2020-06-12;修回日期:2020-08-31; 录用日期:2020-09-21。

基金项目: 国家自然科学基金资助项目(61571410);浙江省自然科学基金资助项目(LY18F020018, LSY19F020001)。

作者简介: 孙忠凡(1995—), 男, 安徽蚌埠人, 硕士研究生, 主要研究方向: 深度学习、图像处理; 周正华(1977—), 男, 陕西商洛人, 副教授, 博士, CCF 会员, 主要研究方向: 深度学习、图像处理; 赵建伟(1977—), 女, 浙江金华人, 教授, 博士, CCF 会员, 主要研究方向: 深度学习、图像处理。

基于深度学习的超分辨率重建方法通常包括深度特征提取和特征图放大两部分。近年来,关于特征提取部分已有各种模式,使得提取出来的特征更有代表性。然而,关于特征图放大的方法却比较少,主要有前端上采样、转置卷积^[4]和像素重组^[5]等。本文主要关注特征图的放大方法。

前端上采样是在低分辨率图像输入网络之后,通过双三次插值等方法将低分辨率图像放大到指定图像尺寸。超分辨率卷积神经网络(Super-Resolution Convolutional Neural Network, SRCNN)^[6]第一次将深度学习应用到超分辨率图像重建中,用三层卷积神经网络拟合放大后的低分辨率图像和高分辨率图像之间的映射关系,并取得了很好的重建效果。使用极深卷积网络的精确图像超分辨率(accurate image Super-Resolution using Very Deep convolutional networks, VDSR)^[7]将放大后的图像输入到网络中,构建20层的深度网络重建高分辨率图像。深度递归卷积网络(Deeply-Recursive Convolutional Network, DRCN)^[8]首次将现有的递归神经网络结构应用于放大后的低分辨率图像,节约网络权值的参数量。深度递归残差网络(Deep Recursive Residual Network, DRRN)^[9]结合了递归学习和残差学习的思想,设计了更深层次的网络结构作用在放大的低分辨率图像上来提高网络重建性能。基于双层可变形卷积网络的超分辨率(Super-Resolution using Bilayer Deformable Convolution Network, SRBDCN)^[10]为了提取不同尺度的图像特征,利用双层可变形卷积单元来构建网络。由于前端上采样在放大低分辨率图像的过程中会引入错误信息,因此由该方法重建的高分辨率图像质量较差。

转置卷积通过在低分辨率图像中插入零值并通过卷积操作来放大图像尺寸。Laplacian 金字塔超分辨率网络(Laplacian pyramid Super-Resolution Network, LapSRN)^[11]通过转置卷积在每个金字塔级上将图像放大一倍,逐步放大低分辨率图像。深度反投影网络(Deep Back-Projection Network, DBPN)^[12]利用转置卷积构造上投影单元和下投影单元,充分利用了高分辨率图像和低分辨率图像之间的关系。嵌入块残差网络(Embedded Block Residual Network, EBRN)^[13]利用转置卷积,考虑图像的纹理特征,利用更深的网络来恢复图像的高频信息。转置卷积的放大方法比前端上采样快,但重建图像会产生棋盘效应。

像素重组是在低分辨率图像经过深度网络特征提取后,利用像素信息进行图像尺寸的放大。有效的亚像素卷积神经网络(Efficient Sub-Pixel Convolutional Neural Network, ESPCN)^[6]首次提出像素重组的方法放大特征图的尺寸,产生更好的高分辨率图像。多尺度残差网络(Multi-Scale Residual Network, MSRN)^[14]利用不同尺度的卷积核构建多尺度残差模块,充分利用图像的特征,并用像素重组放大特征图的尺寸。残差通道关注网络(Residual Channel Attention Network, RCAN)^[15]设计了残差中的残差结构深化网络,使用多个跳越连接使网络专注于高频信息,并用像素重组放大特

征图的尺寸。残差密集网络(Residual Dense Network, RDN)^[16]使用密集连接来提取冗余的局部特征,全局特征融合来提取整体的分层特征,并用像素重组放大特征图的尺寸。虽然像素重组的放大方法可以获得较好的重建结果,但是它仅限于放大整数倍的重建情形。

以上的基于深度学习的超分辨率图像重建主要考虑特征图放大整数倍的情形,较少考虑放大任意倍的情况。然而,现实生活中会遇到高分辨率图像的尺寸并不一定是低分辨率图像的整数倍情况。也有研究者开始尝试解决这个问题,文献[17]基于元学习网络给出了放大任意倍的超分辨率重建方法。该方法为了产生更好的效果,将残差密集网络(Residual Dense Network, RDN)^[16]作为特征提取模块,然后通过构建元学习放大模块放大特征图为任意倍。但是,该方法中的元学习只使用了坐标信息对卷积核进行预测,没有考虑特征图像的空间特性。

鉴于此,本文在文献[17]的基础上,考虑特征图像的空间信息,通过压缩操作提取空间特征,将空间特征和坐标信息相结合,作为权值预测网络的输入,给出空间元学习模块,提出基于空间元学习的超分辨率重建(Spatial Meta-Learning based Super-Resolution, SMLSR)方法。所提的空间元学习模块有效地开发了隐藏在图像中的空间信息,用它可以进行更精准的卷积核预测,产生重建效果更好的高分辨率图像。

综上,本文的主要工作包括以下几个方面:

- 1) 利用深层特征的空间信息和坐标特性,提出空间元学习模块,可以更精准地预测卷积核权值。
- 2) 提出的基于空间元学习的超分辨率重建方法可以实现图像放大任意尺寸,丰富超分辨率重建的方法。
- 3) 本文所提的空间元学习模块可以与其他深度网络相结合,得到放大任意尺寸的超分辨率图像重建方法。

1 SMLSR 方法

在本章中,将详细介绍本文所提的基于空间元学习的超分辨率重建(SMLSR)方法。众所周知,基于深度学习的超分辨率重建包括特征提取和特征图像放大两部分。由于本文的重点在于讨论特征图的放大方法,因此,特征提取部分选择残差密集模块作为网络的主要构成部分。对于特征图放大部分,本文在元学习模块的基础上,考虑特征图的空间信息,通过压缩操作提取空间信息,将空间信息和坐标信息相结合,预测卷积核的权值,给出空间元学习模块。本文所提SMLSR的网络结构如图1所示。接下来,将详细介绍网络各部分的结构。

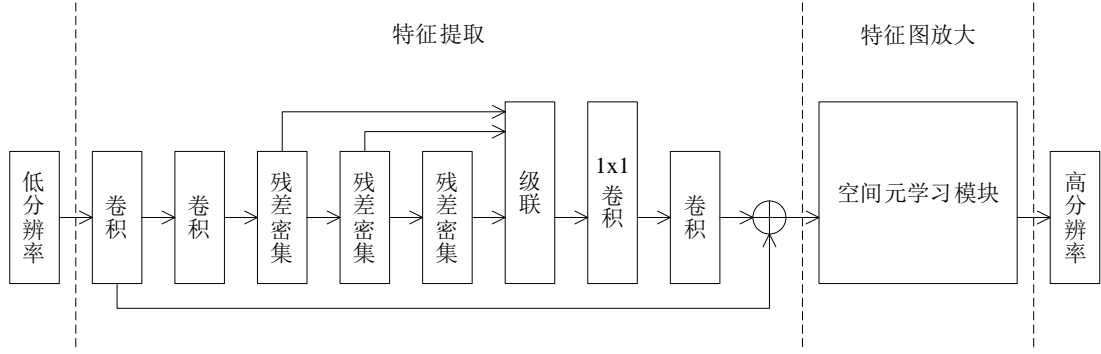


图1 本文所提的基于残差密集网络模块的空间元学习的超分辨率重建方法

Fig. 1 Architecture of proposed spatial meta-learning super-resolution based on residual dense network modules

1.1 特征提取部分

令 I_{LR} 和 I_{SR} 分别表示尺寸大小为 $H \times W$ 和 $rH \times rW$ 的低分辨率图像和重建出的高分辨率图像, 其中 r 为大于1的放大尺寸因子。首先, 将低分辨率图像 I_{LR} 输入到网络, 经过两个卷积层作用, 得到

$$X_1 = \text{Conv}(I_{LR}) \quad (1)$$

$$X_2 = \text{Conv}(X_1); \quad (2)$$

其中: X_1 、 X_2 分别表示通过两层卷积层操作后提取到的特征图; Conv 表示卷积操作。

其次, 将提取到的特征图 X_2 输入到三个串联的残差密集模块, 并将三个模块的输出级联在一起, 经过一个 1×1 的卷积操作将特征图的通道数压缩到与 X_1 通道数相同, 再经过一层卷积操作并通过残差连接, 最终得到深层特征图 F_{LR} 。整个过程可以由如下数学公式描述:

$$B_i = \text{RDB}(B_{i-1}), i=1, 2, 3 \quad (3)$$

$$B = \text{Conv}_{1 \times 1}([B_1, B_2, B_3]) \quad (4)$$

$$F_{LR} = \text{Conv}(B) + X_1; \quad (5)$$

其中: $B_0 = X_2$, B_i 代表经过第 i 个残差密集模块的输出; RDB 表示第 i 个模块的操作作用, $[B_1, B_2, B_3]$ 代表级联操作; F_{LR} 为特征提取部分最后提取的深层特征。

最后, 介绍残差密集模块的操作过程, 其具体结构如图2所示。每个残差密集模块中包含3组卷积层和线性整流函数 (Rectified Linear Unit, ReLU), 输入特征经过每组卷积操作和激活函数后, 将得到的结果输入到后面各个卷积层。将上述特征图的输出进行级联, 并通过 1×1 的卷积操作将级联的特征图的通道数压缩到与模块输入相同的通道数, 最后将其与模块输入相加。整个 RDB 的操作过程可由如下数学公式表示:

$$L_i = \text{ReLU}(\text{Conv}([L_0, L_1, \dots, L_{i-1}]))), i=1, 2, 3 \quad (6)$$

$$B_i = L_0 + \text{Conv}([L_0, L_1, L_2, L_3]); \quad (7)$$

其中: ReLU 为激活函数; $L_0 = B_{i-1}$, L_i 代表着被第 i 个 ReLU 激活函数作用后的输出。

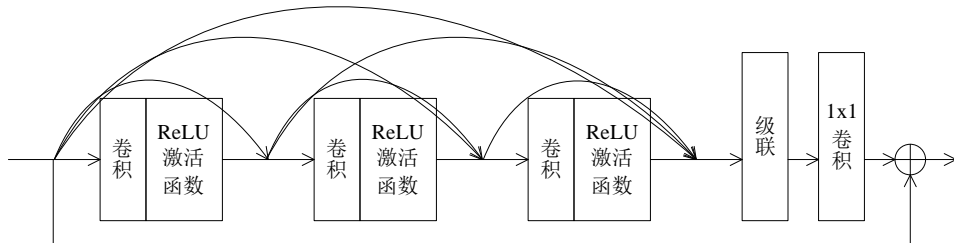


图2 残差密集模块(RDB)的结构

Fig. 2 Architecture of Residual Dense Block (RDB).

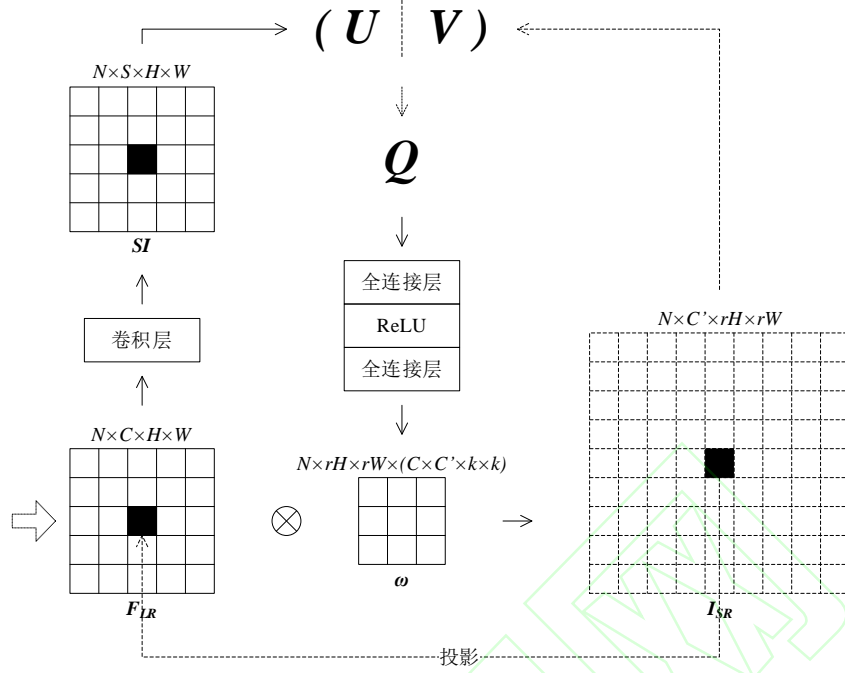


图3 我们所提出的空间元学习模块的结构

Fig. 3 The architecture of our proposed spatial meta learning Module.

1.2 特征图放大部分

特征提取部分提取到深度特征 F_{LR} 之后，将其输入到空间元学习模块。本文所提的空间元学习模块的结构如图3所示，其包含六个部分：坐标投影、空间信息压缩、坐标偏移、信息联合、权值预测和函数映射。该模块首先利用坐标投影的方法在低分辨率图像上找到与高分辨率坐标相关的坐标点；其次，运用空间信息压缩操作生成空间信息特征，并联合坐标偏移量，生成权值预测网络的输入量，通过权值预测网络生成卷积核的权值；最后，将该卷积核作用在深度特征 F_{LR} 上，生成最终的高分辨率图像。具体细节如下。

1) 坐标投影。

基于深度学习重建高分辨率图像，即利用网络确定高分辨率图像中每个像素点的值。对于放大 r 倍的高分辨率图像 I_{SR} ，由于 r 不一定是整数，因此需要通过坐标投影在深度特征 F_{LR} 上找到与 (i', j') 最相关的像素位置 (i, j) ，这样才方便确定卷积的作用区域 $A(i, j)$ ，使得卷积作用在 $A(i, j)$ 上得到 (i', j') 的像素值。给定坐标点 (i', j') 和放大因子 r ，通过以下投影可以在深度特征 F_{LR} 上找到位置 (i, j) ：

$$(i, j) = P(i', j') := (\lfloor i' / r \rfloor, \lfloor j' / r \rfloor); \quad (8)$$

其中： P 是坐标投影， $\lfloor \cdot \rfloor$ 是下取整操作。利用坐标投影，建立了低分辨率图像与放大任何尺寸的图像之间的对应关系

。根据坐标 (i, j) ，可以在深度特征图 F_{LR} 上确定卷积的作用区域 $A(i, j)$ 。

2) 空间信息压缩。

考虑在元学习模块中充分利用特征图 F_{LR} 的空间信息，使用卷积层来压缩提取空间本质信息 U ，学习滤波器的关键信息，移除冗余信息。这里将提取出的深层特征 F_{LR} 压缩成四个通道，该过程表示为：

$$SI = \text{Conv}(F_{LR}); \quad (9)$$

其中 SI 表示卷积后得到的四通道特征图，且 $SI = (S_1, S_2, S_3, S_4)$ ，令

$$U(i, j) = (S_1(i, j), S_2(i, j), S_3(i, j), S_4(i, j)) \quad (10)$$

作为权值预测网络输入的一部分。

3) 坐标偏移。

为了得到较好的滤波权值，需要利用高分辨率图像与低分辨率图像坐标点之间的对应关系。对于放大整数倍的超分辨率重建，通过 (ir, jr) 可以将低分辨率图像上的每个点精确映射到高分辨率图像的对应坐标。然而，对于任意大小的尺度因子，可能会得到一个分数而不是整数。通过添加坐标偏移量，可以利用对应整数位置的偏移程度来预测分数位置的滤波器权值。对于不同的尺度因子 r ，图像具有其不同的尺度特性。因此加入尺度信息，最终构建 $V(i', j')$ 作为权值输入网络的一部分：

$$V(i', j') = (i' / r - \lfloor i' / r \rfloor, j' / r - \lfloor j' / r \rfloor, 1 / r) \quad (10)$$

4) 信息联合。

联合式(2)中压缩出的空间信息和式(3)中的坐标偏移两部分信息,得到权值预测网络的输入于每一批输入图像,构造出权值预测网络的输入矩阵 Q :

$$Q = (U, V) = \begin{pmatrix} S_1(0,0) & S_2(0,0) & S_3(0,0) & S_4(0,0) & R(0) & R(0) & 1/r \\ S_1(0,1) & S_2(0,1) & S_3(0,1) & S_4(0,1) & R(0) & R(1) & 1/r \\ S_1(0,2) & S_2(0,2) & S_3(0,2) & S_4(0,2) & R(0) & R(2) & 1/r \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ S_1(i,j-1) & S_2(i,j-1) & S_3(i,j-1) & S_4(i,j-1) & R(i') & R(j'-1) & 1/r \\ S_1(i,j) & S_2(i,j) & S_3(i,j) & S_4(i,j) & R(i') & R(j') & 1/r \\ S_1(i,j+1) & S_2(i,j+1) & S_3(i,j+1) & S_4(i,j+1) & R(i') & R(j'+1) & 1/r \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (11)$$

其中 $R(k) = k / r - \lfloor k / r \rfloor$ 。上述权值网络的输入,不仅考虑了坐标特性,而且添加了不同像素位置的空间特性,这使得后续预测出的滤波器权值更加地精确。

5) 权值预测。

有了联合信息 Q ,利用子网络来预测相应的滤波器的权值 $\omega(i', j')$ 。这里预测网络 Φ 由两个全连接层和一个激活函数层构成,其数学公式为:

$$\Phi(Q(i', j'); \theta) = F_2 \left(ReLU \left(F_1(Q(i', j'); \theta) \right) \right) \quad (12)$$

其中: F_1, F_2 分别代表第一、第二个全连接层; θ 为网络中需要训练的权值的集合。

利用权值预测网络,可得滤波器权值:

$$\omega(i', j') = \Phi(Q(i', j'); \theta) \quad (13)$$

其中: $\omega(i', j')$ 是对应像素 (i', j') 的滤波器权值。

6) 函数映射。

通过坐标投影和权值预测网络,得到了特征图 F_{LR} 上的卷积作用区域 $A(i, j)$ 和卷积 $\omega(i', j')$ 。利用下面的卷积操作,得到重建的高分辨率图像。

$$I_{SR}(i', j') = A(i, j) * \omega(i', j') \quad (14)$$

其中 $*$ 是卷积操作。

2 实验与结果分析

2.1 实验环境设置

本文使用 291 张图像作为实验的训练集,其中 91 张图像来自文献[18],其他 200 张图片来自 Berkeley Segmentation Dataset^[19]。为了充分利用训练数据,用三种方式增强数据:

1) 将图像缩放为原来的 0.5 和 0.7; 2) 将图像旋转 90° 、 180° 和 270° ; 3) 将图像水平和垂直翻转。另外,选择 Set5^[20]、

Set14^[21]、B100^[22]和 Manga109^[23]作为测试集。将增强后的数据集从 RGB 彩色空间转成 YCbCr 彩色空间。

为了促进训练,将低分辨率图像裁剪成 30×30 的图像块,每 16 个图像块为一批输入进网络。对于任意尺度的超分辨率,在区间 $[1, 4]$ 里选取每 0.2 为一个间隔。使用 Adam^[24] 优化算法进行网络训练,其中的参数分别设置为 $\beta_1 = 0.9, \beta_2 = 0.999, \varepsilon = 10^{-8}$,学习率初始设置为 10^{-4} ,并且每 40 代衰减一半。本文所有实验利用峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似度(Structural Similarity, SSIM)作为图像超分辨率重建效果的评价标准。

所有的实验都在 Pytorch 1.1.0 框架下实施,在 Intel® Xeon® Silver 4108 CPU @ 1.80GHZ、内存为 48GB、GPU 为 GTX2080ti、显存为 11G、操作系统为 Windows10 的设备上运行。

2.2 消融实验

本文所提的超分辨率重建方法在构建元学习模块时,联合空间信息与坐标信息,学习更为合理的卷积核,从而重建出效果更好的高分辨率图像。为了说明空间信息和坐标信息对重建效果的影响,本小节给出了数据集 Manga109^[22]上关于本文方法的消融实验,其中网络第一部分仍取 1.1 节中的特征提取模块。具体实验结果见表 1。

表1 本文所提方法关于空间信息和坐标信息的消融实验。
Tab. 1 Ablation experiment of proposed method on spatial information and coordinate information.

放大因子	空间信息	坐标信息	PSNR(dB)	SSIM
3 倍	√	×	24.96	0.8280
	×	√	32.17	0.9357
	√	√	32.23	0.9361

如表 1 所示,元学习中仅结合坐标信息的算法性能比仅结合空间信息的算法性能要好,主要原因是空间信息的构造方式不适合单独用来预测卷积核。如图 4 所示,通过式(8)的坐标投影操作,可能会将高分辨率图像上的两个点 (k', l') 和 $(k', l'+1)$ 投影到低分辨率图像上的同一个点 (k, l) ,导致高分辨率上的两个点 (k', l') 和 $(k', l'+1)$ 有不同的坐标信息 $V(k', l')$ 和 $V(k', l'+1)$,却有相同的空间信息 $U(k, l)$ 。因此,仅用空间信息 $U(k, l)$ 进行预测会使得 (k', l') 和 $(k', l'+1)$ 所对应的卷积核相同,从而生成相同的像素值,使重建效果变差。但是,当空间信息与坐标信息联合使用时,对于映射在相同低分辨率坐标 (k, l) 上的高分辨率的点 (k', l') 和 $(k', l'+1)$ 会有不同的坐标信息,使得这两个点的对应的权值预测网络的输入不同,从而可以预测出不同的卷积核,提高网络的重建效果。

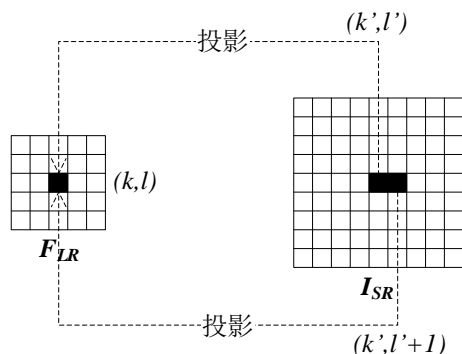


图4 高分辨率图像与低分辨率图像像素点映射关系

Fig. 4 Mapping relation of pixels from high resolution image to low resolution image.

2.3 放大任意尺度的超分辨率方法

为了验证本文所提出的空间元学习模块的有效性,将网络的特征提取部分统一设置为残差密集模块,而特征图放大部分选取不同的方法。假设将低分辨率图像放大 r 倍:第一种比较方法选取双三次插值;第二种基线是前端上采样基线,用双三次下采样将低分辨率图像放大 r 倍,将扩大后的图像输入到网络进行重建;第三种方法是文献[17]中基于元学习的超分辨率方法。本文所提方法与上述三种方法在各个放大尺寸下的重建结果见表2。

如表2所示,虽然双三次插值可以不需要进行网络的训练就可以得到超分辨率图像,但是仅仅得到质量比较差的超分辨率图像。前端上采样将双三次预处理后的图像输入网络,以至于通过网络后直接可以得到相应尺寸的图像。然而,这种方法非常耗时,产生的超分辨图像的质量也不是很好。用于超分辨率的放大任意倍网络(A Magnification-arbitrary Network for Super-Resolution, Meta-SR)^[17]在预测滤波器

时没有考虑到原始图像的空间信息,所以它不能实现最好的效果。本文所提出的 SMLSR 几乎在所有的尺度上产生了比其他任意倍的图像放大方法更高的 PSNR。实验结果表明本文所提出的空间元学习放大方法具有一定的优越性。

表2 数据集 B100 上放大任意倍的超分辨率方法的峰值信噪比(单位: dB)

Tab. 2 PSNR of super-resolution methods with arbitrary magnification on test dataset B100 (Unit: dB).

方法	放大因子				
	1.2 倍	1.4 倍	1.6 倍	1.8 倍	2.0 倍
双三次插值	35.1060	32.9707	31.5172	30.4315	29.5621
前端上采样	40.0056	36.7095	34.5986	33.1064	31.9902
Meta-SR ^[17]	40.0357	36.7434	34.6354	33.1467	32.0277
本文方法	40.0348	36.7446	34.6336	33.1521	32.0354
方法	放大因子				
	2.2 倍	2.4 倍	2.6 倍	2.8 倍	3.0 倍
双三次插值	28.9179	28.3823	27.9320	27.5192	27.2074
前端上采样	31.0714	30.3784	29.779	29.266	28.8162
Meta-SR ^[17]	31.1103	30.4204	29.8235	29.3141	28.8648
本文方法	31.1174	30.424	29.8346	29.3162	28.874
方法	放大因子				
	3.2 倍	3.4 倍	3.6 倍	3.8 倍	4.0 倍
双三次插值	26.8987	26.5943	26.3313	26.1291	25.9593
前端上采样	28.455	28.1049	27.7741	27.5136	27.2443
Meta-SR ^[17]	28.5008	28.1475	27.8287	27.5706	27.2922
本文方法	28.4998	28.1621	27.8323	27.5721	27.3047

表3 放大整数倍的超分辨率方法比较

Tab. 3 Comparison of super-resolution methods with integral magnification

方法	放大因子	Set5		Set14		B100		Manga109	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
双三次插值	2 倍	33.69	0.9284	30.24	0.8675	29.56	0.8433	30.82	0.9332
前端上采样	2 倍	37.73	0.9595	33.2	0.913	31.99	0.8989	37.57	0.9748
转置卷积 ^[4]	2 倍	37.74	0.9594	33.25	0.9131	32.03	0.8992	37.75	0.9749
像素重组 ^[5]	2 倍	37.78	0.9595	33.25	0.9139	32.03	0.8997	37.76	0.9751
Meta-SR ^[17]	2 倍	37.77	0.9594	33.26	0.9131	32.03	0.8993	37.73	0.975
本文方法	2 倍	37.79	0.9593	33.24	0.913	32.04	0.8994	37.75	0.975
双三次插值	3 倍	30.41	0.8655	27.55	0.7721	27.21	0.7342	26.96	0.8556
前端上采样	3 倍	33.71	0.9209	29.82	0.8322	28.82	0.7978	31.97	0.9338
转置卷积 ^[4]	3 倍	—	—	—	—	—	—	—	—
像素重组 ^[5]	3 倍	33.94	0.9226	29.88	0.8344	28.87	0.8004	32.25	0.9358
Meta-SR ^[17]	3 倍	33.93	0.9225	29.89	0.8343	28.86	0.8006	32.17	0.9357
本文方法	3 倍	33.9	0.9226	29.9	0.8346	28.87	0.8001	32.23	0.9361

双三次插值	4 倍	28.43	0.8022	26	0.6936	25.96	0.6514	24.91	0.7826
前端上采样	4 倍	31.25	0.878	27.97	0.762	27.24	0.7163	28.6	0.883
转置卷积 ^[4]	4 倍	31.45	0.8818	28.1	0.7678	27.3	0.7216	29	0.8886
像素重组 ^[5]	4 倍	31.54	0.8827	28.09	0.7676	27.29	0.7216	29.04	0.8887
Meta-SR ^[17]	4 倍	31.53	0.8822	28.1	0.7662	27.29	0.7204	29.02	0.8886
本文方法	4 倍	31.49	0.882	28.11	0.7672	27.3	0.7212	28.95	0.8884

2.4 放大整数尺度的超分辨率方法

因为当前基于深度学习的超分辨率重建方法主要是放大整数倍的尺度因子。本实验将本文的方法和这些放大整数倍的方法进行比较。设置了一系列基线,包括双三次基线、前端上采样基线、转置卷积^[4]基线、像素重组^[5]基线和 Meta-SR^[17]基线。在这里,转置卷积^[4]的参数设置遵循 DBPN^[12],因此没有放大 3 倍的参数配置。在 Set5^[20]、Set14^[21]、B100^[22]、Manga109^[23]四个数据集上进行实验,其实验结果见表 3,加粗的数值为最好的结果。

如表 3 所示,不同整数倍方法在不同数据集不同的尺度上有不同的优势。双三次插值和前端上采样在整数倍放大上和任意倍放大一样,产生的高分辨率图像效果较差。转置卷积在网络的最后利用转置卷积层放大图像,速度相比于前端上采样更快,产生了相比前端上采样更好的结果。Meta-SR 相比于前三种方法进一步提高了性能。总体上,在整数放大因子为 2 倍和 4 倍时,像素重组相比其他方法产生了最优的结果;而在整数放大因子为 3 倍时,本文的方法产生了最优的结果。与这些先进的放大整数倍的方法相比,本文提出的空间元学习具有一定的优势。实验结果表明本文提出的空间元学习不仅产生了最优的任意尺度因子超分辨率效果,在整数尺度因子上也能达到先进水平。

2.5 复杂度分析

神经网络的时间复杂度和空间复杂度通常用计算量:浮点运算数(Floating Point Operations, FLOPs)和网络参数量来衡量。本小节给出放大 2 倍下,不同重建方法构造出的网络模型在训练阶段所需的计算量和参数量,实验结果见表 4,其中 1GFLOPs 等于十亿次浮点运算数。

表4 不同重建方法的浮点运算数和参数量

Tab. 4 Number of floating point operations and parameters for different reconstruction methods

方法	计算量/ GFLOPs	参数量/ 10^4
前端上采样	2.9	81
转置卷积 ^[4]	0.73	81
像素重组 ^[5]	1.26	95
Meta-SR ^[17]	0.73	95
本文方法	0.73	96

表 4 结果表明,在计算量方面,因为前端上采样方法在整个模型上的运算都是高分辨率图像,所以计算量最大。转置卷积^[4]、像素重组^[5]、Meta-SR^[17]和本文方法仅仅在模型的

最后放大图像,但是像素重组^[5]利用卷积层将通道扩大 $r \times r$ 倍,极大地增加了计算量;而 Meta-SR^[17]和本文方法所用的权值预测网路所包含的两个全连接层的计算量较低,加上本文方法提取空间信息的卷积层,计算量也基本和转置卷积^[4]计算量相当,远低于像素重组^[5]的计算量。在参数量方面,前端上采样和转置卷积^[4]较低,因为前端上采样将双三次插值放大后的图像输入网络,仅仅在模型的最后用一层卷积将通道进行压缩。在卷积核大小相同的情况下,转置卷积层和普通卷积层参数量相同,故转置卷积^[4]参数量也不高。像素重组^[5]利用两个卷积层分别扩维和压缩维度,与 Meta-SR、本文方法所用的权值参数量基本相当。实验结果表明,在任意倍放大图像的方法当中,本文所提出的方法几乎在所有的尺度上产生了最好的效果,却与其他的方法拥有几乎一样的参数量和最少的计算量。在整数倍放大图像的方法当中,本文所提出的重建方法与像素重组具有相当的参数量,在不同整数倍上的效果各有优势,然而计算量几乎只有像素重组的一半。本文所提出的方法与其他方法相比,在参数量计算量都较少的情况下产生了更好的效果。

3 结语

本文在元学习网络的基础上,考虑特征图的空间信息,联合该空间信息与坐标信息,设计空间元学习模块,并结合残差密集网络结构,提出了基于空间元学习的超分辨率图像重建方法。所提的空间元学习模块可以与其他深度网络相结合,得到放大任意倍的超分辨率图像重建方法。实验结果表明,本文所提的重建方法优于现有的基于相同特征提取网络但不同放大模块的超分辨率重建方法。

本文所提出的基于空间元学习的放大任意倍超分辨率重建方法虽然比其他放大方法更有优势,但是在生成卷积核时需要将一维数据转换成多个维度,极大的考验了计算机硬件。因此,下一步工作将考虑如何构建更简单高效的放大任意倍重建方法,更方便地重建出高分辨率图像。

参考文献

- [1] SHI W Z, CABALLERO J, LEDIG C, et al. Cardiac image super-resolution with global correspondence using multi-atlas PatchMatch [C]// MICCAI 2013: Proceedings of the 16th International Conference on Medical Image Computing and Computer Assisted Intervention. Berlin: Springer, 2013: 9-16.
- [2] YILDIRIM D, GÜNGÖR O. A novel image fusion method using IKONOS satellite images [J]. Journal of Geodesy and Geoinformation, 2012, 1(1): 75-83.

- [3] 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.(HENG S, ZHOU J, ZHANG Z H. Survey of super-resolution image reconstruction methods[J]. Acta Automatica Sinica, 2013, 39(8):1202-1213.)
- [4] DONG C, LOY C C, TANG X O. Accelerating the super-resolution convolutional neural network [C]// ECCV 2016: Proceedings of the 14th European Conference on Computer Vision. Berlin: Springer, 2016: 391-407.
- [5] 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]// CVPR 2016: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2016: 1874-1883.
- [6] DONG C, LOY C C, HE K M, et al. Learning a deep convolutional network for image super-resolution [C]// ECCV 2014: Proceedings of 13th European Conference on Computer Vision. Berlin: Springer, 2014: 184-199.
- [7] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]// CVPR 2016: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 1646-1654.
- [8] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution [C]// CVPR 2016: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 1637-1645.
- [9] TAI Y, YANG J, LIU X M. Image super-resolution via deep recursive residual network [C]// CVPR 2017: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 3147-3155.
- [10] 黄陶冶,赵建伟,周正华. 双层可变形卷积网络的超分辨率图像重建[J]. 计算机应用, 2019, 39(z2):68-74. (HUANG T Y, ZHAO J W. ZHOU Z H. Super-resolution reconstruction using bilayer deformable convolutional networks [J]. Journal of Computer Applications, 2019, 39(z2):68-74.)
- [11] LAI W S, HUANG J B, AHUJN N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution [C]// CVPR 2017: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 624-632.
- [12] HARIS M, SHAKHNAROVICH G, UKITA N. Deep back-projection networks for super-resolution [C]//CVPR 2018: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 1664-1673.
- [13] QIU Y J, WANG R X, TAO D P, et al. Embedded Block Residual Network: A Recursive Restoration Model for Single-Image Super-Resolution [C]// ICCV 2019: Proceedings of the 2019 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2019: 4180-4189.
- [14] LI J C, FANG F M, MEI K F, et al. Multi-scale residual network for image super-resolution [C]// ECCV 2018: Proceedings of 15th European Conference on Computer Vision. Berlin: Springer, 2018: 517-532.
- [15] ZHANG Y L, LI K P, LI K, et al. Image super-resolution using very deep residual channel attention networks [C]// ECCV 2018: Proceedings of 2018 15th European Conference on Computer Vision. Berlin: Springer, 2018: 286-301.
- [16] ZHANG Y L, TIAN Y P, KONG Y, et al. Residual dense network for image super-resolution [C]// CVPR 2018: Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 2472-2481.
- [17] HU X C, MU H Y, ZHANG X Y, et al. Meta-SR: a magnification-arbitrary network for super-resolution [C]// CVPR 2019: Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019:1575-1584.
- [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] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics [C]// ICCV 2001: Proceedings of the 2001 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2001: 416-423.
- [20] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding [C]// BMVC 2012: Proceedings of the 23rd British Machine Vision Conference. Durham: BMVA Press, 2012: 135.1-135.10.
- [21] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations [C]// Proceedings of the 7th International Conference on Curves and Surfaces. Berlin: Springer, 2010: 711-730.
- [22] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.
- [23] MATSUI Y, ITO K, ARAMAKI Y, et al. Sketch-based manga retrieval using manga109 dataset [J]. Multimedia Tools and Applications, 2017, 76(20): 21811-21838.
- [24] KINGMA D, BA J. Adam: a method for stochastics optimization [EB/OL]. [2020-02-07]. <https://arxiv.org/abs/1412.6980>. pdf.

This work is partially supported by the Natural Science Foundation of Zhejiang Province (LY18F020018, LSY19F020001), the National Natural Science Foundation of China (61571410).

SUN Zhongfan, born in 1995, M. S. candidate. His research interests include deep learning, image processing.

ZHOU Zhenghua, born in 1977, Ph. D., associate professor. His research interests include deep learning, image processing.

ZHAO Jianwei, born in 1977, Ph. D., professor. Her research interests include deep learning, image processing.