

基于多尺度特征融合反投影网络的图像超分辨率重建

孙超文¹ 陈晓^{1,2}

摘要 针对现有图像超分辨率重建方法恢复图像高频细节能力较弱、特征利用率不足的问题,提出了一种多尺度特征融合反投影网络用于图像超分辨率重建.该网络首先在浅层特征提取层使用多尺度的卷积核提取不同维度的特征信息,增强跨通道信息融合能力;然后,构建多尺度反投影模块通过递归学习执行特征映射,提升网络的早期重建能力;最后,将局部残差反馈结合全局残差学习促进特征的传播和利用,从而融合不同深度的特征信息进行图像重建.对图像进行 $\times 2 \sim \times 8$ 超分辨率的实验结果表明,本方法的重建图像质量在主观感受和客观评价指标上均优于现有图像超分辨率重建方法,超分辨率倍数大时重建性能相比更优秀.

关键词 图像超分辨率重建,多尺度卷积,特征融合,反投影

引用格式 孙超文,陈晓.基于多尺度特征融合反投影网络的图像超分辨率重建.自动化学报,2021,47(7):1689–1700

DOI 10.16383/j.aas.c200714

Multiscale Feature Fusion Back-projection Network for Image Super-resolution

SUN Chao-Wen¹ CHEN Xiao^{1,2}

Abstract Aiming at the problems that existing image super-resolution reconstruction methods have weak ability to restore image high-frequency details and insufficient feature utilization, a multi-scale feature fusion back projection network is proposed for image super-resolution reconstruction. The network first uses multi-scale convolution kernels in the shallow feature extraction layer to extract feature information of different dimensions to enhance cross-channel information fusion; then builds a multi-scale back projection module to perform feature mapping through recursive learning to improve the early reconstruction capabilities of the network; Finally, local residual feedback is combined with global residual learning to promote the spread and utilization of features, thereby fusing feature information of different depths for image reconstruction. The experimental results of $\times 2 \sim \times 8$ SR on the images show that the quality of SR image of this method is better than the existing image super-resolution method in subjective perception and objective evaluation index, and the reconstruction performance is relatively better when the scale factors is large.

Key words Image super-resolution, multi-scale convolution, feature fusion, back-projection

Citation Sun Chao-Wen, Chen Xiao. Multiscale feature fusion back-projection network for image super-resolution. *Acta Automatica Sinica*, 2021, 47(7): 1689–1700

图像的超分辨率重建 (Super-resolution, SR) 主要分为单图像超分辨率重建和多图像超分辨率重建^[1], 本文研究单图像超分辨率重建. 作为一种重要的图像处理技术, SR 从低分辨率 (Low-resolution,

LR) 图像生成对应的高分辨率 (High-resolution, HR) 图像^[2]. 它在图像压缩^[3]、医学成像^[4]、安全监控^[5]、卫星遥感图像^[6] 等领域有着广泛的应用. 近年来基于深度学习的图像超分辨率重建方法由于其优秀的重建性能受到广泛关注^[7].

SRCNN^[8–9] 首次将卷积神经网络 (Convolutional neural network, CNN) 用于 SR. 此后许多研究提出了一系列基于 CNN 的 SR 算法. 增加 CNN 网络的深度虽然可以提供更大的感受野和更多的上下文信息, 但是容易引起两个问题: 梯度消失和大量参数^[10]. 为了有效地缓解梯度消失的问题, VDSR^[11] 和 EDSR^[12] 算法借助残差学习^[13] 训练了更深层的网络; SR-DenseNet^[14]、RDN^[15] 和 DBPN^[16] 算法则引入密集连接缓解梯度消失的问题^[17]. 为了减少网络参数, DRCN^[18] 和 DRRN^[19]

收稿日期 2020-09-02 录用日期 2021-02-09

Manuscript received September 2, 2020; accepted February 9, 2021

江苏省 333 高层次人才培养工程项目 (2625); 江苏高校优势学科建设工程资助项目资助

Supported by 333 High Level Personnel Training Project Jiangsu Province of China (2625); the Priority Academic Program Development of Jiangsu Higher Education Institutions

本文责任编辑 杨健

Recommended by Associate Editor YANG Jian

1. 南京信息工程大学电子与信息工程学院 南京 210044 2. 南京信息工程大学江苏省大气环境与装备技术协同创新中心 南京 210044

1. School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044 2. Jiangsu Collaborative Innovation Center for Atmospheric Environment and Equipment Technology, Nanjing University of Information Science and Technology, Nanjing 210044

等网络采用递归学习实现权重共享. 但只是以递归的方式堆叠多个相同的层或单元, 容易导致出现深度网络模型的退化问题. 基于递归网络的 SR 算法以前馈的方式从输入的 LR 图像中学习分层表示, 这种单向映射依赖于 LR 图像中的有限特征. 基于插值的前馈网络如 FSRCNN^[20] 和 EPSCN^[21], 迁移到其他放大倍数需要繁琐的预处理, 容易引入新的噪声, 难以适应更大倍数的放大. 虽然 MSRN^[22] 和 LapSRN^[23] 可以执行 8 倍放大但重建效果仍有很大的改善空间. 尽管 DBPN 算法^[16] 引入反投影机制实现了深度 SR, 却需要庞大的训练集和时间, 包括 DIV2K (800 张 2K 图像)^[24], Flickr2K (2650 张 2K 图像)^[12] 和 ImageNet 数据集 (超过 1 400 万张图像)^[25]. 并且, 以单一尺寸卷积提取图像特征减弱网络恢复高频细节的能力, 以单尺度的投影单元执行特征映射降低网络的学习能力, 在放大倍数大时这个问题更突出.

针对以上问题, 本文提出了一种多尺度特征融合反投影网络用于图像超分辨率重建, 其结构如图 1 所示. 作为一种模块化的端到端网络, 只需对参数进行少量调整就可以扩展到任意放大倍数, 实现了

灵活地定义和训练不同深度的网络. 本文主要工作如下:

1) 在浅层特征提取阶段, 基于 GoogleNet 的思想^[26], 设计多尺度特征提取模块. 采用多个不同的卷积核对输入 LR 图像进行卷积操作, 提取图像的多尺度特征, 再进行特征级联实现多尺度特征融合, 最后利用 1×1 卷积进行降维. 多尺度特征提取模块使得网络在浅层特征提取阶段获取不同维度的特征信息, 可以使网络学习到图像更多的高频信息, 提高重建性能.

2) 在特征映射阶段, 构建多尺度投影单元实现自适应获取不同尺度的图像特征. 在两条支路分别引入不同大小的卷积核进行上 (下) 采样, 不同支路间可以信息共享, 实现充分利用图像的局部特征. 采用 1×1 卷积完成数据降维和特征的跨通道融合, 在学习前一层的表示中添加非线性激励来提高网络的泛化能力. 这种局部多路径学习方式增强了支路信息交流, 丰富了投影特征, 进一步增大网络的感受野和提高模型学习复杂映射的能力.

3) 实现高层信息和低层信息的融合. 多尺度上下投影单元通过递归学习从输入的 LR 图像中生

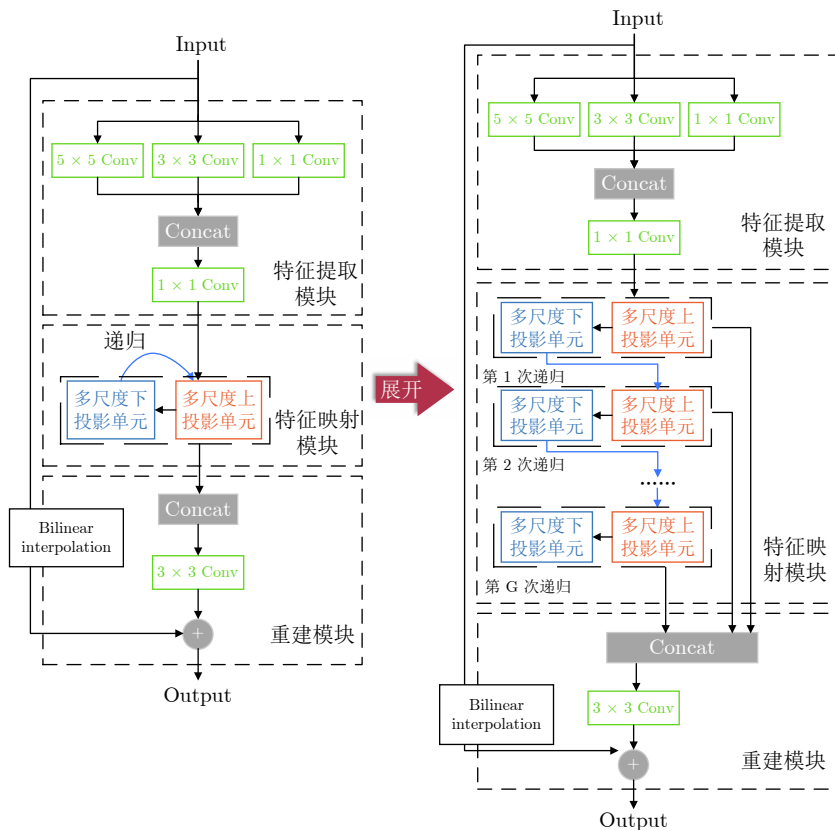


图 1 本文算法网络结构图
Fig.1 The structure of the proposed network

成 HR 特征再投影回 LR 空间, 实现残差反馈过程. 这种反馈机制将高层信息传输回前层并调整输入, 在丰富高层特征的同时完善低层特征表示. 携带反馈的递归结构不断调整 LR 和 HR 图像之间的相互关系, 在增加网络深度的同时有效地减少网络冗余. 网络通过组合不同深度的 HR 特征生成重建图像, 高层与低层特征信息的融合使得网络具有强大的早期重建能力.

4) 实现全局信息和局部信息的融合. 局部残差信息结合全局残差信息可以极大程度地利用图像特征并克服特征在传输过程中消失的缺陷. 本网络不仅应用迭代上下采样方法为特征映射模块的投影残差提供局部残差反馈, 为早期层的学习提供丰富的局部特征信息, 还在重建阶段通过全局残差跳跃连接将残差图像从 LR 输入添加到全局恒等映射中, 使得网络仅需要学习图像残差, 极大地降低学习难度并加速了网络的收敛. 局部残差反馈和全局残差学习的结合促进了特征的传播和利用, 为最终的重建提供更多的特征信息.

1 多尺度特征融合反投影网络

多尺度特征融合反投影网络主要由 3 个部分组成, 分别为浅层特征提取模块, 特征映射模块和重建模块, 如图 1 所示.

由于应用了全局残差学习, 因此整个网络将 LR 图像作为输入, 仅需要学习 HR 图像和插值 LR 图像之间的残差. 令 $conv(f, n)$ 表示卷积层, 其中 f 为卷积核大小, n 为通道数, 三个模块的详细阐述如下.

1.1 浅层特征提取模块

首先, 使用不同尺度的卷积核在三条支路中分别对输入的 LR 图像 I^{LR} 进行卷积操作提取图像初始特征, 然后将特征图级联实现多特征的跨通道融合并利用 1×1 卷积层进行数据降维, 得到浅层 LR 特征图 L^0 :

$$\begin{aligned} L_{1 \times 1}^0 &= f_{1 \times 1}(I^{LR}) \\ L_{3 \times 3}^0 &= f_{3 \times 3}(I^{LR}) \\ L_{5 \times 5}^0 &= f_{5 \times 5}(I^{LR}) \end{aligned} \quad (1)$$

$$L^0 = f_0([L_{1 \times 1}^0, L_{3 \times 3}^0, L_{5 \times 5}^0]) \quad (2)$$

式中, $f_{1 \times 1}$ 、 $f_{3 \times 3}$ 和 $f_{5 \times 5}$ 分别表示 $conv(1, n_0)$ 、 $conv(3, n_0)$ 和 $conv(5, n_0)$, n_0 为初始特征提取阶段的通道数, f_0 表示 $conv(1, n)$, n 是多尺度投影单元的输入通道数, f_0 将总通道数 $3 \times n_0$ 降为 n .

1.2 特征映射模块

LR 特征图 L^0 流入特征映射模块得到多个不同

深度的 HR 特征图 H^g :

$$H^g = f_{FM}^g(L^{g-1}), \quad 1 \leq g \leq G \quad (3)$$

式中, G 表示递归总次数; 当 $g = 1$ 时, 浅层 LR 特征图 L^0 作为当前输入; 当 $g > 1$ 时, 将第 $g-1$ 次递归产生的 LR 特征图 L^{g-1} 作为本次输入. f_{FM}^g 表示第 g 次递归的特征投影过程, 详细解释如下: 将一个多尺度上投影单元级联一个多尺度下投影单元实现将 LR 特征投影到 HR 空间, 再投影回 LR 空间. 本网络仅使用一对多尺度上下投影单元, 通过递归学习实现权重共享并构成迭代反馈机制. 在多尺度投影单元中构建两条支路并且对不同的支路应用不同尺度的卷积核. 这种局部多路径学习方式可以实现不同旁路之间的信息共享并充分利用图像局部特征. 基于迭代上下采样框架, 设计多尺度上投影单元和多尺度下投影单元分别用于上采样和下采样操作.

1.2.1 多尺度上投影单元

多尺度上投影单元是将 LR 特征图 L^{g-1} 映射到 HR 空间中, 计算流程如下:

步骤 1. 输入 LR 特征图 L^{g-1} , 利用不同尺度的反卷积在两条支路上分别进行上采样操作, 生成 HR 特征图 H_{u1}^g 和 H_{u2}^g .

$$\begin{aligned} H_{u1}^g &= D_{u1}^\uparrow(L^{g-1}) \\ H_{u2}^g &= D_{u2}^\uparrow(L^{g-1}) \end{aligned} \quad (4)$$

其中, D_{u1}^\uparrow 和 D_{u2}^\uparrow 分别表示反卷积层 $Deconv1(k_1, n)$ 和 $Deconv2(k_2, n)$, k_1 和 k_2 表示卷积核大小, n 表示卷积核数量.

步骤 2. 级联 HR 特征图 H_{u1}^g 和 H_{u2}^g , 在两条支路上利用不同尺度的卷积分别进行下采样操作, 将产生的 LR 特征图级联实现跨通道信息融合, 然后利用 1×1 卷积进行特征池化和降维并生成 LR 特征图 L_u^g .

$$\begin{aligned} L_u^g &= C_u([C_{u1}^\downarrow([H_{u1}^g, H_{u2}^g]), \\ &C_{u2}^\downarrow([H_{u1}^g, H_{u2}^g])]) \end{aligned} \quad (5)$$

其中, C_{u1}^\downarrow 和 C_{u2}^\downarrow 分别表示卷积层 $Conv1(k_1, 2n)$ 和 $Conv2(k_2, 2n)$. 每条支路的通道数由 n 变成 $2n$. C_u 表示 $Conv(1, n)$, 将两条支路的总通道数由 $4n$ 降为 n . 所有 1×1 卷积在前一层的学习表示上添加了非线性激励, 提升网络的表达能力.

步骤 3. 计算输入的 LR 特征图 L^{g-1} 和重建的 LR 特征图 L_u^g 之间的残差, 并构建两条支路利用不同尺度的反卷积分别进行上采样操作, LR 空间中的残差被映射到 HR 空间中, 生成新的残差特征 H_{e1}^g 和 H_{e2}^g .

$$\begin{aligned} H_{e1}^g &= D_{e1}^\uparrow(L_u^g - L^{g-1}) \\ H_{e2}^g &= D_{e2}^\uparrow(L_u^g - L^{g-1}) \end{aligned} \quad (6)$$

其中, D_{e1}^\uparrow 和 D_{e2}^\uparrow 分别表示反卷积层 $Deconv1(k_1, n)$ 和 $Deconv2(k_2, n)$, 每条支路的通道数为 n .

步骤 4. 级联 H_{e1}^g 和 H_{e2}^g 并与级联的 H_{u1}^g 和 H_{u2}^g 进行叠加, 通过 1×1 卷积进行特征融合和降维, 多尺度上投影单元输出 HR 特征图 H^g .

$$H^g = C_h([H_{u1}^g, H_{u2}^g] + [H_{e1}^g, H_{e2}^g]) \quad (7)$$

其中, C_h 表示 $Conv(1, n)$, 叠加后总通道数为 $2n$, 通过 C_h 输出通道数变为 n , 与输入通道数保持一致.

1.2.2 多尺度下投影单元

多尺度下投影单元与多尺度上投影单元十分相似, 是多尺度上投影单元的逆过程, 目的是将 HR 特征投影到 LR 空间中, 计算过程如下:

步骤 1. 将 HR 特征图 H^g 作为输入, 利用不同尺度的卷积在两条支路上分别进行下采样操作, 生成 LR 特征图 L_{d1}^g 和 L_{d2}^g .

$$\begin{aligned} L_{d1}^g &= C_{d1}^\downarrow(H^g) \\ L_{d2}^g &= C_{d2}^\downarrow(H^g) \end{aligned} \quad (8)$$

其中, C_{d1}^\downarrow 和 C_{d2}^\downarrow 分别表示卷积层 $Conv1(k_1, n)$ 和 $Conv2(k_2, n)$.

步骤 2. 级联 LR 特征图 L_{d1}^g 和 L_{d2}^g , 在两条支路上利用不同尺度的卷积分别进行上采样操作, 将产生的 HR 特征图级联实现跨通道信息融合, 然后利用 1×1 卷积进行特征池化和降维并生成 HR 特征图 H_d^g .

$$\begin{aligned} H_d^g &= C_d([D_{d1}^\uparrow([L_{d1}^g, L_{d2}^g]), \\ &D_{d2}^\uparrow([L_{d1}^g, L_{d2}^g])]) \end{aligned} \quad (9)$$

其中, D_{d1}^\uparrow 和 D_{d2}^\uparrow 分别表示反卷积层 $Deconv1(k_1, 2n)$ 和 $Deconv2(k_2, 2n)$. 每条支路的通道数由 n 变成 $2n$, C_u 表示 $Conv(1, n)$, 将两条支路的总通道数由 $4n$ 降为 n .

步骤 3. 计算输入的 HR 特征图 H^g 和重建的 HR 特征图 H_d^g 之间的残差, 并构建两条支路利用不同尺度的卷积分别进行下采样操作, HR 空间中的残差被反投影回 LR 空间中, 生成新的残差特征 L_{e1}^g 和 L_{e2}^g .

$$\begin{aligned} L_{e1}^g &= C_{e1}^\downarrow(H^g - H_d^g) \\ L_{e2}^g &= C_{e2}^\downarrow(H^g - H_d^g) \end{aligned} \quad (10)$$

其中, C_{e1}^\downarrow 和 C_{e2}^\downarrow 分别表示具有不同尺度卷积核的卷积层 $Conv1(k_1, n)$ 和 $Conv2(k_2, n)$, 每条支路的通道数为 n .

步骤 4. 级联 LR 残差特征 L_{e1}^g 和 L_{e2}^g , 并与步骤 2 中级联的 LR 特征图 L_{d1}^g 和 L_{d2}^g 进行叠加, 通过 1×1 卷积进行特征融合和降维, 得到多尺度下投影单元最终的输出 LR 图 L^g .

$$L^g = C_l([L_{d1}^g, L_{d2}^g] + [L_{e1}^g, L_{e2}^g]) \quad (11)$$

其中, C_l 表示 $Conv(1, n)$, 相加后总通道数为 $2n$, 通过 $Conv(1, n)$ 输出通道数降低为 n , 输入和输出保持相同的通道数.

令多尺度上、下投影单元的输入和输出都具有相同的通道数, 这种结构允许多个多尺度投影单元相互交替连接.

为了控制参数数量并降低计算复杂度, 许多网络往往避免使用大尺寸卷积进行特征映射, 这虽然可以从一定程度上提升收敛速度减少运算成本, 但却以牺牲重建性能为代价. 而以递归学习的方式实现多尺度上、下投影单元的迭代能够大大促进权重共享减少网络冗余. 这种设计不仅可以抑制由大尺寸卷积核带来的收敛速度慢的缺陷, 还可以通过扩大感受野提取更多图像特征以提高重建效果. 因此多尺度投影单元的每条支路均使用了大尺寸卷积核.

1.3 重建模块

重建模块融合特征映射模块生成的不同深度的 HR 特征图并输出残差图像 I^{Res} ,

$$I^{\text{Res}} = f_{RB}([H^1, H^2, \dots, H^g]) \quad (12)$$

其中, $[H^1, H^2, \dots, H^g]$ 表示多个 HR 特征的深度级联, f_{RB} 表示重建模块的操作, 即对级联的 HR 特征应用 $conv(3, 3)$ 进行卷积处理实现分层特征融合.

最后, 通过全局残差跳跃连接输出超分辨率图像 I^{SR} , 其过程表示为:

$$I^{SR} = I^{\text{Res}} + f_{US}(I^{LR}) \quad (13)$$

其中, f_{US} 表示插值上采样操作, 采用双线性插值将原始输入图像 I^{LR} 放大至目标尺寸 (也可以选用其他插值算法, 如双三次插值法等), 然后绕过网络主体部分将插值 LR 图像传递到网络末端与 I^{Res} 相加, 生成最终的重建图像 I^{SR} .

2 实验与结果分析

2.1 实验设置

实验平台的 CPU 为 Intel Core i5-7500, GPU 为 NVIDIA RTX-2080. 使用深度学习框架 Pytorch 1.2.0.

把 DIV2K 中的 800 张训练图像作为训练集^[24], 为了充分利用数据, 采用了旋转和翻转进行了数据

扩充^[12]. 选用 5 个标准数据集评估模型性能, 分别为 Set5^[27], Set14^[28], BSD100^[29], Urban100^[30] 和 Manga109^[31]. 用峰值信噪比 (PSNR) 和结构相似性 (SSIM)^[32] 作为客观性能评价指标. 批处理大小设置为 16. 为了提高计算效率, 根据不同的放大倍数选择输入不同大小的 RGB 图像, 如表 1 所示. 将 Bicubic 作为退化模型, 采用 He 等^[33] 提出的方法初始化权重, 使用 ADAM^[34] 优化器优化参数. 将学习率初始化为 0.0001, 每 200 轮学习率衰减为原来的一半, 总共训练 1 000 轮. 选择 $L1$ 损失来训练本文提出的网络.

在多尺度投影单元的两条支路中设计不同大小的卷积核, 并对应 4 种放大倍数设置了相应的卷积核和步幅. 在最终模型中, 将递归参数设置为 7, 多尺度投影单元的通道数设置为 64, 即输出 64 个 HR 特征图, 因此 64 个特征进行 7 次迭代产生 448 个 HR 特征图作为重建模块的输入. 表 1 详细列出了不同放大倍数下的超参数设置. 网络的输入和输出均使用 RGB 通道. 除网络末端的重建层外, 所有卷积和反卷积层后面的激活函数都使用参数校正线性单元 (PReLU)^[32].

2.2 模型分析

首先对本文算法进行模型分析, 然后与其他 SR 算法的重建结果 ($\times 2$ 、 $\times 3$ 、 $\times 4$ 和 $\times 8$) 进行定性和定量比较.

2.2.1 特征提取模块卷积核尺寸分析

卷积核大小能够直接改变网络的感受野, 从而对网络的超分辨率性能产生重要影响^[35-36]. 为了达到最佳实验结果, 本文对特征提取层三条支路上的卷积核大小进行研究. 将卷积核大小分别设置 1×1 、 3×3 和 5×5 表示为 Ours_135, 将卷积核大小分别

设置 3×3 、 5×5 和 7×7 表示为 Ours_357, 结合图 2 和表 2, 对于 8 倍放大, 在 Set5 测试集下当网络收敛后 Ours_135 比 Ours_357 的 PSNR 结果高 0.04 dB; Ours_135 在 5 个标准测试集中获得了最高的 SSIM 值, 在 4 个测试集中均具有最佳的 PSNR 值, 仅在 Set14 中比 Ours_357 低 0.01 dB, 总体平均仍高出 0.051 dB. 此外, 更大的卷积核意味着更多的参数, 从而降低网络的推理速度. 实验表明, 当采用 1×1 、 3×3 和 5×5 的卷积组合时, 网络能够充分提取特征, 并获得更高的评价指标, 具有更高的泛化能力.

2.2.2 多尺度投影单元的卷积核分析

为了分析多尺度投影单元两条支路上的卷积核大小对重建性能的影响, 我们选取了 $(8 \times 8, 10 \times 10)$ 、 $(10 \times 10, 12 \times 12)$ 和 $(12 \times 12, 14 \times 14)$ 三种不同的多尺度卷积组合在 $\times 8$ SR 上进行研究, 实验结果如表 3 所示, 当采用大尺度卷积核 ($12 \times 12, 14 \times 14$) 的组合方案时, 网络具有最佳的重建结果.

2.2.3 多尺度投影单元数量分析

为了分析多尺度投影单元数量对超分辨率性能的影响, 本文采用 set5 测试集在 8 倍放大下对多尺度投影单元数量进行实验, 递归次数对应了网络中一个多尺度上、下投影单元的数量. 实验结果如表 4 所示, 每增加 1 次递归, 网络增加 10 层, 参数数量增加 1 728, 但随着网络的逐渐加深, 网络之间的性能差距越来越小, 7 次递归和 8 次递归的网络重建结果几乎相同, 考虑到模型复杂度和计算成本, 最终选了 7 次递归的模型.

2.2.4 参数规模分析

表 1 显示了本算法在不同放大倍数下的参数数量, 为了进一步对本文算法的复杂度进行分析, 本

表 1 输入块大小、参数数量和网络超参数设置

Table 1 The settings of input patch size, number of parameters and network hyperparameters

放大倍数	$\times 2$	$\times 3$	$\times 4$	$\times 8$
参数数量	5 016 211	6 490 771	8 161 939	16 812 691
输入块尺寸	60×60	50×50	40×40	20×20
特征提取模块	特征提取层	$f_{1 \times 1}: Conv(128, 1, 1, 0); f_{3 \times 3}: Conv(128, 3, 1, 1); f_{5 \times 5}: Conv(128, 5, 1, 2)$		
	特征融合层	$Conv(128 \times 3, 1, 1, 1)$		
特征映射模块	支路1	$Conv1(64, 6, 2, 2)$	$Conv1(64, 7, 3, 2)$	$Conv1(64, 8, 4, 2)$
	支路2	$Conv2(64, 8, 2, 3)$	$Conv2(64, 9, 3, 3)$	$Conv2(64, 10, 4, 3)$
重建模块	$Conv(64 \times 7, 3, 1, 1)$			
递归次数	7			
深度	73			

注: $Conv(C, K, S, P)$: C 表示通道数, K 表示卷积核大小, S 表示步长, P 表示填充.

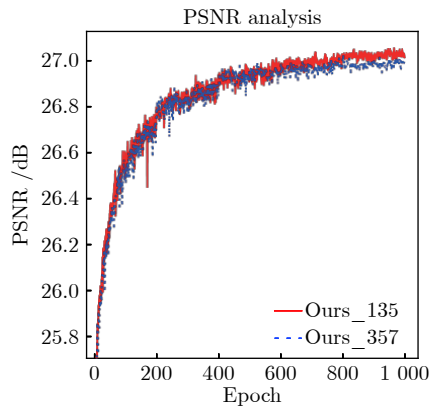


图 2 8 倍放大下对特征提取模块卷积核大小的分析
Fig.2 Analysis of kernel size in the feature extraction module on $\times 8$ enlargement

文选取了 6 种主流的基于深度学习的 SR 算法和本算法的参数数量进行对比, 包括 SRCNN、FSR-CNN、VDSR、LapSRN、MSRN 和 EDSR. 如图 3 所示, 本文算法重建图像的质量比 EDSR 的重建结果提高了 0.17 dB, 但参数数量却减少了约 61%; DBPN 算法虽然通过训练庞大的图像集获得了比本文算法略高的重建性能 (0.08 dB), 但同时参数规模也提升了 38%. 综上所述, 本文算法属于中等规模网络, 与 EDSR 和 DBPN 等具有大规模参数的网络相比, 仍然可以产生具有竞争力的重建结果, 在网络性能和参数量之间具有最佳的平衡.

2.3 与目前先进方法的比较

2.3.1 定量分析

DBPN 算法使用了大量的训练图像实现了复杂深度模型的高性能. 为公平起见, 本文对 DBPN 算法进行了参数调整和重新训练, 保持和本文算法

相同的训练集和迭代次数, 使用一对单尺度投影单元进行 7 次递归, 每个投影单元输出的 64 个 HR 特征图, 共 448 个特征图作为重建层的输入, 并将其命名为 DBPN-R64-7. 本文算法与 DBPN 算法的对比结果如图 4 所示.

从图 4(b) ~ (d) 训练时损失函数的迭代结果可以看出, 本算法相比 DBPN-R64-7 具有更快的收敛速度, 收敛后的 loss 值更低. 随着放大倍数的增加, 这种优势更加明显. 根据图 4(a) ~ (d) 对客观评价指标的分析可以得出, 在 Set 5 数据集上, 对于 2 到 8 倍多个放大倍数, 网络收敛后本文算法比 DBPN-R64-7 获得了更好的 PSNR 和 SSIM 值, 尤其在放大倍数大时. 放大倍数越大重建难度也越大. 而本算法具有更好的重建性能, 比 DBPN-R64-7 算法能够更好地适应重建难度大的放大倍数.

对于 $\times 2$ 、 $\times 3$ 、 $\times 4$ 的 SR, 将本文算法与 13 种 SR 方法进行了对比, 如表 5 所示. 由于很多算法不适用于 8 倍放大, 因此在 $\times 8$ SR 中与 10 种方法进行了对比, 如表 6 所示. 在对比算法中, Bicubic 为传统的双三次插值算法, SRCNN、ESPCN 和 FSR-CNN 为基于 CNN 的浅层线性网络, VDSR 为深层线性网络, DRCN、DRRN 和 DRFN^[37] 为递归网络, 其中 DRFN 是设计处理大放大倍数 SR 的深度网络, LapSRN 为渐进式重建网络, IDN^[38] 和 MSRN 为多路径网络, SRMDNF^[39] 为多重退化网络, MRFN^[40] 和 EDSR 为残差网络. 结合表 5 和表 6 可以看出, 本文算法仅在 2 倍放大时与 MRFN 存在微弱的差距, 但 $\times 3$ 、 $\times 4$ 和 $\times 8$ SR 时在 5 个具有不同图像特点的数据集中均具有最高的客观评价指标, 显示了在处理大放大倍数 SR 时的有效性, 能够比其他算法更好地保存 HR 分量. 在 5 个数据集中, Set5、Set14 和 B100 主要包含自然场景; Urban100

表 2 对特征提取模块卷积核大小的分析
Table 2 Analysis of the kernel size of the feature extraction module

Scale	Method	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSD100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
$\times 8$	Ours_135	27.13/0.7819	25.02/0.6445	24.86/0.5992	22.59/0.6231	24.85/0.7885
	Ours_357	27.09/0.7806	25.03/0.6437	24.86/0.5986	22.57/0.6219	24.78/0.7859

表 3 对多尺度投影单元的卷积核大小分析
Table 3 Analysis of the kernel size of the multi-scale projection unit

Scale	(卷积核尺寸、步长、填充)		PSNR (dB)				
	支路1	支路2	Set5	Set14	BSD100	Urban100	Manga109
$\times 8$	(8, 8, 0)	(10, 8, 1)	27.00	24.95	24.82	22.45	24.68
	(10, 8, 1)	(12, 8, 2)	27.08	24.99	24.84	22.53	24.77
	(12, 8, 2)	(14, 8, 3)	27.13	25.02	24.86	22.59	24.85

表 4 $\times 8$ 模型在 Set5 和 Set14 测试集上的深度分析
Table 4 The depth analysis of the $\times 8$ model on Set5 and Set14 datasets

递归次数	网络层数	参数数量	PSNR (dB)	
			Set5	Set14
1	13	16802323	26.50	24.53
3	33	16805779	26.98	24.89
4	43	16807507	27.03	24.94
5	53	16809235	27.05	24.96
6	63	16810963	27.07	24.98
7	73	16812691	27.13	25.02
8	83	16814419	27.13	25.02

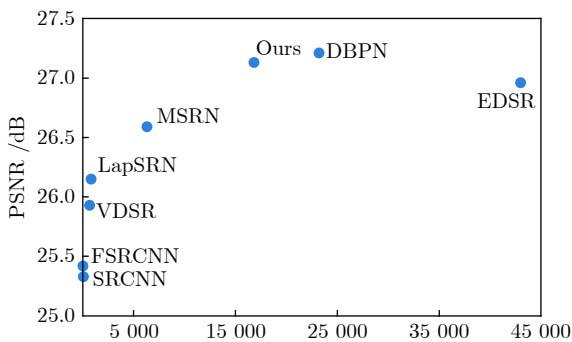


图 3 主流重建算法在 Set5 数据集上对于 $\times 8$ SR 的平均 PSNR 和参数数量对比

Fig. 3 Comparison of the average PSNR and the number of parameters of the mainstream reconstruction algorithm for $\times 8$ SR on Set5

由城市场景中的规则图案和人造结构组成, 包含不同频段的细节特征; Manga109 是漫画数据集. 本文算法不仅擅长重建规则的人工图案, 还擅长不规则的自然图案. 尤其是本训练集中没有包含任何漫画图像, 但在 Manga109 中展现出出色的实验性能. 这表明本文算法在处理漫画人物等具有精细结构的图像上也具有优秀的重建效果.

纵向来看, 在 Set5 数据集上对于 $\times 2 \sim \times 8$ SR, 本文算法的 PSNR 值相较于 DBPN-R64-7 分别提高了 0.25 dB、0.41 dB、0.39 dB、0.31 dB. 横向来看, 当放大倍数为 8 时, 本文算法 PSNR 值在 5 个数据集上比 EDSR 算法分别提高了 0.17 dB、0.11 dB、0.05 dB、0.08 dB 和 0.16 dB. 实验结果表明, 在 2 到 8 倍超分辨率重建时, 本文算法对于不同类型的图像和不同的放大倍数均具有良好的重建性能, 更好地适应重建难度更大的放大倍数, 而且放大倍数越大本文算法的优势相比越明显.

2.3.2 定性分析

选取 Bicubic、SRCNN、FSRCNN、DRCN、

DRRN、VDSR 和 LapSRN 这 7 种算法与本文算法进行了定性对比, 图 5 到图 9 显示了在 8 倍放大下对于 5 个标准测试集的超分辨率视觉效果对比图.

图 5 显示了对于自然图像 “Woman” 的可视化结果, 本文算法重建出更加清晰准确的眉毛图案, 而其他算法损失了更多的细节; 图 6 显示了对于自然图像 “Zebra” 的可视化结果, 其他算法由于对高频分量的恢复能力较弱所以重建出了错误的条纹方向, 并且遭受不同程度的模糊和棋盘伪影, 而本文算法则成功消除了这些负面干扰并重建出更接近原始 HR 图像的纹理信息. 从图 7 可以看出, 本文算法对于不规则图案超越了其他算法, 可以预测更多的边缘细节. 从图 8 可以看出, 本文算法对包含更多中频和低频信息的规则结构图案的重建性能具有出色的效果, 最大程度地还原了高频细节并极大减少振铃效应的干扰. 图 9 显示了对于漫画图像 “Touyou-Kidan” 的可视化结果. 对于包含更多复杂结构和精细纹理的漫画图像, 其他算法由于特征利用不足难以估计高频信息, 降低了对文本细节的重建能力, 重建图像易遭受平滑边缘和严重的模糊伪影, 但本文算法抑制了平滑分量, 重建出清晰锐利的文字, 明显预测了更多高频细节, 具有更加清晰的边缘和轮廓特征.

综上所述, 由于图像退化过程中信息的丢失, 尤其是高频信息的丢失, 这些基于 CNN 的 SR 算法只能还原平滑的图像边缘, 并且放大倍数越大, 边缘模糊越为严重. 而本文算法能够抑制平滑分量, 预测更多的高频分量和细化特征, 使 SR 图像具有更加清晰锐利的边缘和轮廓, 在很大程度上缓解棋盘伪影和振铃效应的干扰, 并且在 8 倍放大中仍然保留这种优势, 产生最接近原始真实图像的结果.

3 结论

本文中提出了一种基于多尺度特征融合反投影网络用于图像的超分辨率重建. 与单尺度网络相比, 网络中的多尺度特征提取模块在多个上下文尺度上提取图像特征, 捕获更多的细节信息, 还可以从输入的 LR 图像中创建上下文信息, 促进 LR 图像的恢复; 与前馈网络相比, 本网络可以利用高层特征细化低层表示, 并进一步丰富高级信息, 更深层次的挖掘 LR 和 HR 图像间的相互依赖关系. 本文算法不仅注重局部信息和全局信息的融合, 还侧重于将低级细节与高级抽象语义相结合, 面对 8 倍放大仍能很好地生成 HR 图像. 实验分析表明, 本文算法减轻了边缘结构的锯齿效应和模糊伪影的影响, 具有优秀的重建性能, 优于其他 SR 方法, 尤其在 8 倍放大中. 未来研究方向是将其应用于更高倍数 (如

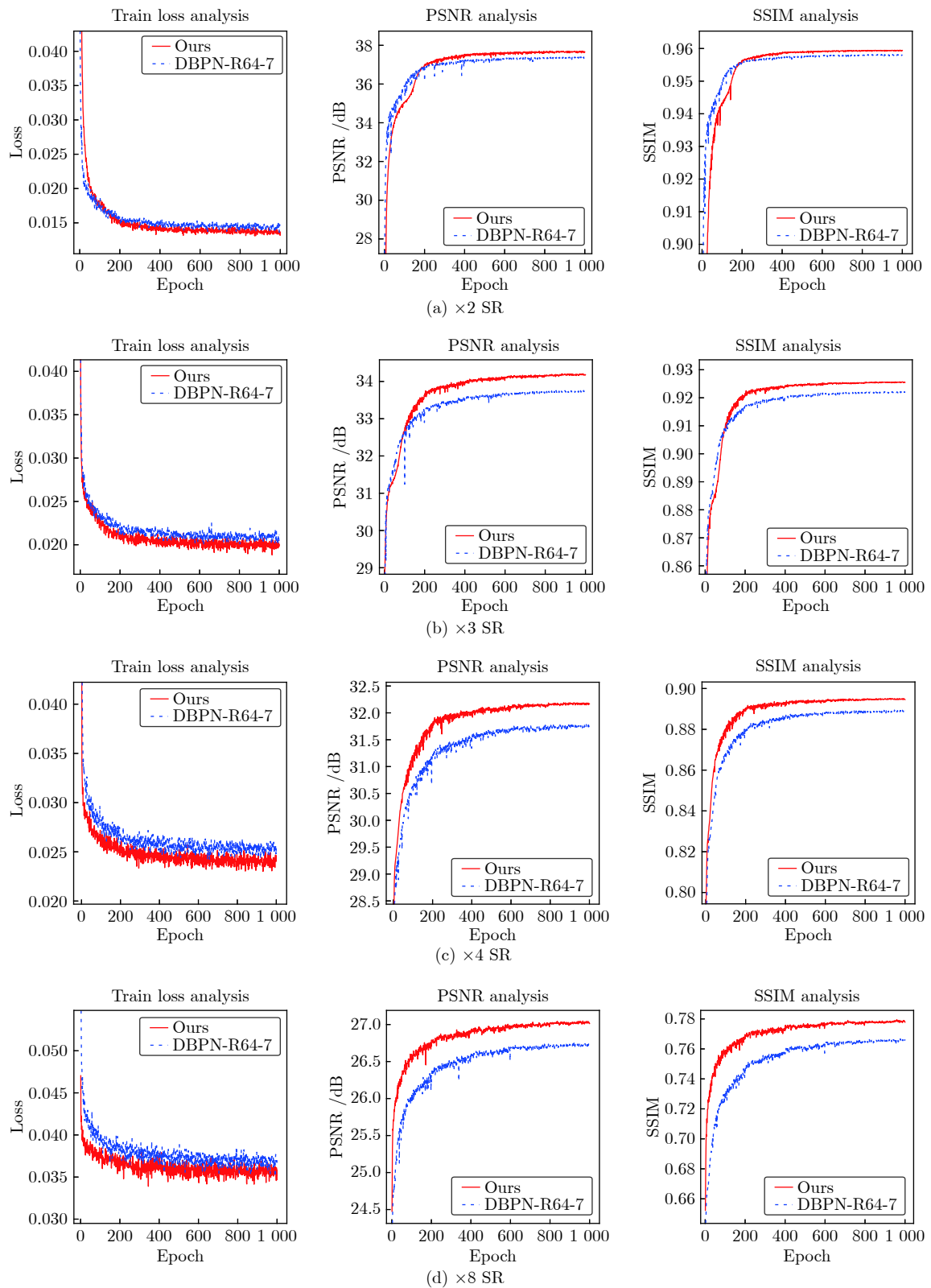


图 4 Set5 测试集下对不同网络的性能分析

Fig.4 Analysis of different networks under Set5

16 倍) 的 SR. 此外, 实际情况中往往需要使用任意放大倍数进行超分辨率重建, 因此设计和实现图像

任意放大倍数超分辨率重建的单个网络也是潜在的研究方向.

表 5 不同 SR 算法在 $\times 2$ 、 $\times 3$ 和 $\times 4$ 上的定量评估
 Table 5 Quantitative comparison of different algorithms on $\times 2$, $\times 3$, and $\times 4$

Scale	Method	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSD100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
$\times 2$	1. Bicubic	33.68/0.9304	30.24/0.8691	29.56/0.8435	26.88/0.8405	31.05/0.9350
$\times 2$	2. SRCNN	36.66/0.9542	32.45/0.9067	31.36/0.8879	29.51/0.8946	35.72/0.9680
$\times 2$	3. ESPCN	37.00/0.9559	32.75/0.9098	31.51/0.8939	29.87/0.9065	36.21/0.9694
$\times 2$	4. FSRCNN	37.06/0.9554	32.76/0.9078	31.53/0.8912	29.88/0.9024	29.88/0.9024
$\times 2$	5. VDSR	37.53/0.9587	33.05/0.9127	31.90/0.8960	30.77/0.9141	37.16/0.9740
$\times 2$	6. DRCN	37.63/0.9588	33.06/0.9121	31.85/0.8942	30.76/0.9133	37.57/0.9730
$\times 2$	7. LapSRN	37.52/0.9591	32.99/0.9124	31.80/0.8949	30.41/0.9101	37.53/0.9740
$\times 2$	8. DRRN	37.74/0.9591	33.23/0.9136	32.05/0.8973	31.23/0.9188	37.92/0.9760
$\times 2$	9. DBPN-R64-7	37.57/0.9589	33.09/0.9132	31.83/0.8951	30.75/0.9133	37.65/0.9747
$\times 2$	10. IDN	37.83/0.9600	33.30/0.9148	32.08/0.8985	31.27/0.9196	38.02/0.9749
$\times 2$	11. SRMDNF	37.79/0.9601	33.32/0.9159	32.05/0.8985	31.33/0.9204	38.07/0.9761
$\times 2$	12. DRFN	37.71/0.9595	33.29/0.9142	32.02/0.8979	31.08/0.9179	33.42/0.9123
$\times 2$	13. MRFN	37.98/0.9611	33.41/0.9159	32.14/0.8997	31.45/0.9221	38.29/0.9759
$\times 2$	Ours	37.82/0.9599	33.35/0.9156	32.04/0.8980	31.49/0.9218	38.23/0.9762
$\times 3$	1. Bicubic	30.40/0.8686	27.54/0.7741	27.21/0.7389	24.46/0.7349	26.95/0.8560
$\times 3$	2. SRCNN	32.75/0.9090	29.29/0.8215	28.41/0.7863	26.24/0.7991	30.48/0.9120
$\times 3$	3. ESPCN	33.02/0.9135	29.49/0.8271	28.50/0.7937	26.41/0.8161	30.79/0.9181
$\times 3$	4. FSRCNN	33.20/0.9149	29.54/0.8277	28.55/0.7945	26.48/0.8175	30.98/0.9212
$\times 3$	5. VDSR	33.66/0.9213	29.78/0.8318	28.83/0.7976	27.14/0.8279	32.01/0.9340
$\times 3$	6. DRCN	33.82/0.9226	29.77/0.8314	28.80/0.7963	27.15/0.8277	32.31/0.9360
$\times 3$	7. LapSRN	33.82/0.9227	29.79/0.8320	28.82/0.7973	27.07/0.8271	32.21/0.9350
$\times 3$	8. DRRN	34.03/0.9244	29.96/0.8349	28.95/0.8004	27.53/0.8377	32.74/0.9390
$\times 3$	9. DBPN-R64-7	33.90/0.9236	29.99/0.8353	28.87/0.7991	27.35/0.8336	32.59/0.9373
$\times 3$	10. IDN	34.11/0.9253	29.99/0.8354	28.95/0.8013	27.42/0.8359	32.69/0.9378
$\times 3$	11. SRMDNF	34.12/0.9254	30.04/0.8382	28.97/0.8025	27.57/0.8398	33.00/0.9403
$\times 3$	12. DRFN	34.01/0.9234	30.06/0.8366	28.93/0.8010	27.43/0.8359	30.59/0.8539
$\times 3$	13. MRFN	34.21/0.9267	30.03/0.8363	28.99/0.8029	27.53/0.8389	32.82/0.9396
$\times 3$	Ours	34.31/0.9265	30.29/0.8408	29.05/0.8035	27.94/0.8472	33.37/0.9433
$\times 4$	1. Bicubic	28.43/0.8109	26.00/0.7023	25.96/0.6678	23.14/0.6574	25.15/0.7890
$\times 4$	2. SRCNN	30.48/0.8628	27.50/0.7513	26.9/0.7103	24.52/0.7226	27.66/0.8580
$\times 4$	3. ESPCN	30.66/0.8646	27.71/0.7562	26.98/0.7124	24.60/0.7360	27.70/0.8560
$\times 4$	4. FSRCNN	30.73/0.8601	27.71/0.7488	26.98/0.7029	24.62/0.7272	27.90/0.8517
$\times 4$	5. VDSR	31.35/0.8838	28.02/0.7678	27.29/0.7252	25.18/0.7525	28.82/0.8860
$\times 4$	6. DRCN	31.53/0.8854	28.03/0.7673	27.24/0.7233	25.14/0.7511	28.97/0.8860
$\times 4$	7. LapSRN	31.54/0.8866	28.09/0.7694	27.32/0.7264	25.21/0.7553	29.09/0.8900
$\times 4$	8. DRRN	31.68/0.8888	28.21/0.7720	27.38/0.7284	25.44/0.7638	29.46/0.8960
$\times 4$	9. DBPN-R64-7	31.92/0.8915	28.41/0.7770	27.42/0.7304	25.59/0.7681	29.92/0.9003
$\times 4$	10. IDN	31.82/0.8903	28.25/0.7730	27.41/0.7297	25.41/0.7632	—
$\times 4$	11. SRMDNF	31.96/0.8925	28.35/0.7787	27.49/0.7337	25.68/0.7731	30.09/0.9024
$\times 4$	12. DRFN	31.55/0.8861	28.30/0.7737	27.39/0.7293	25.45/0.7629	28.99/0.8106
$\times 4$	13. MRFN	31.90/0.8916	28.31/0.7746	27.43/0.7309	25.46/0.7654	29.57/0.8962
$\times 4$	Ours	32.31/0.8963	28.71/0.7843	27.66/0.7383	26.30/0.7922	30.84/0.9126

表 6 不同 SR 算法在×8 上的定量评估
Table 6 Quantitative comparison of different algorithms on ×8

Scale	Method	Set5 PSNR/SSIM	Set14 PSNR/SSIM	BSD100 PSNR/SSIM	Urban100 PSNR/SSIM	Manga109 PSNR/SSIM
× 8	1. Bicubic	24.40/0.6580	23.10/0.5660	23.67/0.5480	20.74/0.5160	21.47/0.6500
× 8	2. SRCNN	25.33/0.6900	23.76/0.5910	24.13/0.5660	21.29/0.5440	22.46/0.6950
× 8	3. ESPCN	25.75/0.6738	24.21/0.5109	24.37/0.5277	21.59/0.5420	22.83/0.6715
× 8	4. FSRCNN	25.42/0.6440	23.94/0.5482	24.21/0.5112	21.32/0.5090	22.39/0.6357
× 8	5. VDSR	25.93/0.7240	24.26/0.6140	24.49/0.5830	21.70/0.5710	23.16/0.7250
× 8	6. LapSRN	26.15/0.7380	24.35/0.6200	24.54/0.5860	21.81/0.5810	23.39/0.7350
× 8	7. DRFN	26.22/0.7400	24.57/0.6250	24.60/0.5870	—	—
× 8	8. MSRN	26.59/0.7254	24.88/0.5961	24.70/0.5410	22.37/0.5977	24.28/0.7517
× 8	9. DBPN-R64-7	26.82/0.7700	24.77/0.6346	24.72/0.5928	22.22/0.6033	24.19/0.7664
× 8	10. EDSR	26.96/0.7762	24.91/0.6420	24.81/0.5985	22.51/0.6221	24.69/0.7841
× 8	Ours	27.13/0.7819	25.02/0.6445	24.86/0.5992	22.59/0.6231	24.85/0.7885



图 5 在 Set5 上×8 SR 的可视化结果 (woman)
Fig.5 Visualized results of ×8 SR on Set5 (woman)

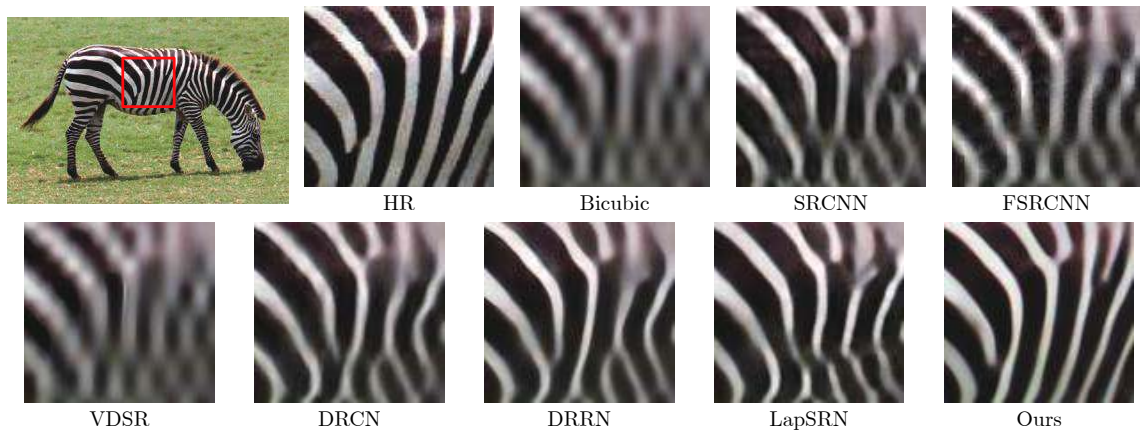


图 6 在 Set14 上×8 SR 的可视化结果 (zebra)
Fig.6 Visualized results of ×8 SR on Set14 (zebra)

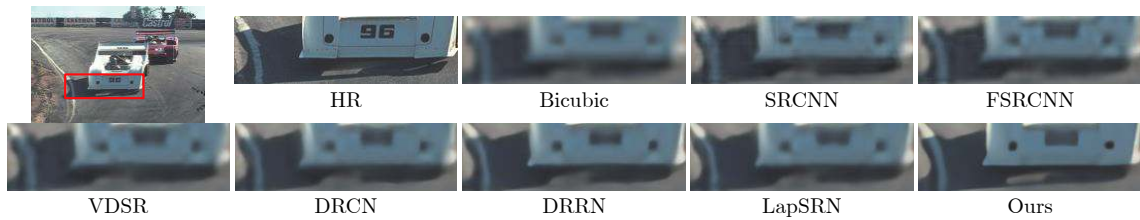
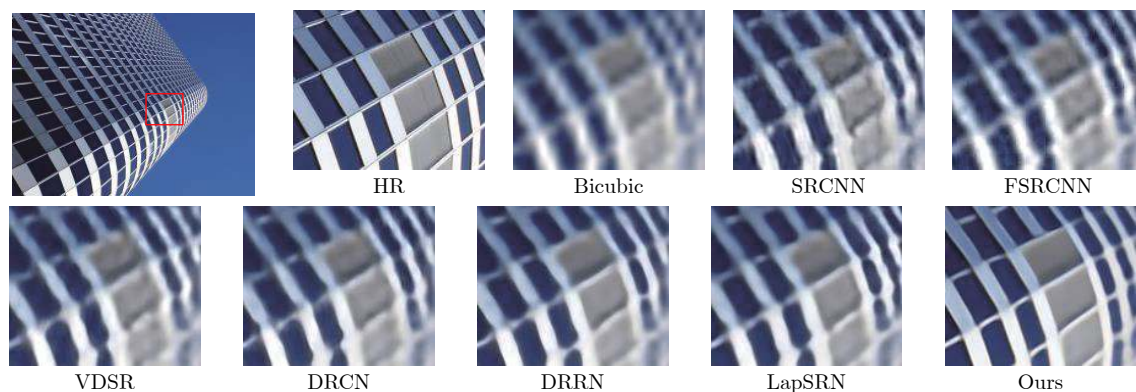
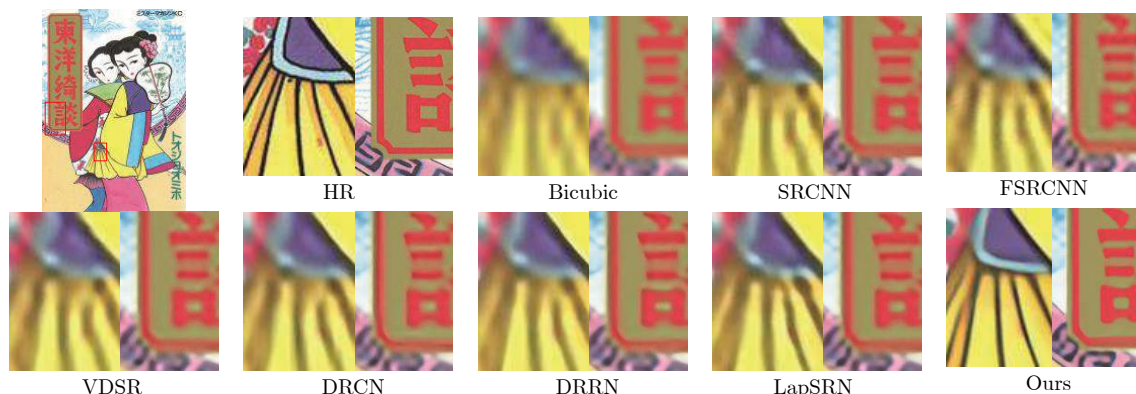


图 7 在 BSD100 上×8 SR 的可视化结果 (210779)
Fig.7 Visualized results of ×8 SR on BSD100 (210779)

图 8 在 Urban100 上 $\times 8$ SR 的可视化结果 (img005)Fig.8 Visualized results of $\times 8$ SR on Urban100 (img005)图 9 在 Manga109 上 $\times 8$ SR 的可视化结果 (TouyouKidan)Fig.9 Visualized results of $\times 8$ SR on Manga109 (TouyouKidan)

References

- Zhang Ning, Wang Yong-Cheng, Zhang Xin, Xu Dong-Dong. A review of single image super-resolution based on deep learning. *Acta Automatica Sinica*, 2020, **46**(12): 2479–2499 (张宁, 王永成, 张欣, 徐东东. 基于深度学习的单幅图片超分辨率重建研究进展. *自动化学报*, 2020, **46**(12): 2479–2499)
- Zhang Yi-Feng, Liu Yuan, Jiang Cheng, Cheng Xu. A curriculum learning approach for single image super resolution. *Acta Automatica Sinica*, 2020, **46**(2): 274–282 (张毅锋, 刘袁, 蒋程, 程旭. 用于超分辨率重建的深度网络递进学习方法. *自动化学报*, 2020, **46**(2): 274–282)
- Tan Y, Cai J, Zhang S, Zhong W, Ye L. Image compression algorithms based on super-resolution reconstruction technology. In: Proceedings of the 2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC), 2019. 162–166
- You C, Li G, Zhang Y, Zhang X, Shan H, Li M, Ju S, Zhao Z, Zhang Z, Cong W, Vannier M W, Saha P K, Hoffman E A, Wang G. CT super-resolution GAN constrained by the identical, residual, and cycle learning ensemble (GAN-CIRCLE). *IEEE Transactions on Medical Imaging*, 2020, **39**(1): 188–203
- Pang Y, Cao J, Wang J, Han J. JCS-Net: Joint classification and super-resolution network for small-scale pedestrian detection in surveillance images. *IEEE Transactions on Information Forensics and Security*, 2019, **14**(12): 3322–3331
- Zhou Deng-Wen, Zhao Li-Juan, Duan Ran, Chai Xiao-Liang. Image super-resolution based on recursive residual networks. *Acta Automatica Sinica*, 2019, **45**(6): 1157–1165 (周登文, 赵丽娟, 段然, 柴晓亮. 基于递归残差网络的图像超分辨率重建. *自动化学报*, 2019, **45**(6): 1157–1165)
- Sun Xu, Li Xiao-Guang, Li Jia-Feng, Zhuo Li. Review on deep learning based image super-resolution restoration algorithms. *Acta Automatica Sinica*, 2017, **43**(5): 697–709 (孙旭, 李晓光, 李嘉峰, 卓力. 基于深度学习的图像超分辨率复原研究进展. *自动化学报*, 2017, **43**(5): 697–709)
- Dong C, Loy C C, He K, Tang X. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, **38**(2): 295–307
- Dong C, Loy C C, He K, Tang X. Learning a deep convolutional network for image super-resolution. In: Proceedings of the 2014 European Conference on Computer Vision (ICCV), Springer, Cham, 2014. 184–199
- Liu Jian-Wei, Zhao Hui-Dan, Luo Xiong-Lin, Xu Jun. Research progress on batch normalization of deep learning and its related algorithms. *Acta Automatica Sinica*, 2020, **46**(6): 1090–1120 (刘建伟, 赵会丹, 罗雄麟, 许隼. 深度学习批归一化及其相关算法研究进展. *自动化学报*, 2020, **46**(6): 1090–1120)
- Kim J, Kwon Lee J, Mu Lee K. Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016. 1646–1654
- B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee. Enhanced deep residual networks for single image super-resolution. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, 2017. 136–144
- K. He, X. Zhang, S. Ren and J. Sun. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, 770–778
- Tong T, Li G, Liu X, Gao Q. Image super-resolution using dense skip connections. In: Proceedings of the 2017 IEEE International

- al Conference on Computer Vision (ICCV), Venice, 2017. 4809–4817
- 15 Zhang Y, Tian Y, Kong Y, Zhong B, Fu Y. Residual dense network for image super-resolution. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018. 2472–2481
- 16 Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for single image super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, DOI: [10.1109/TPAMI.2020.3002836](https://doi.org/10.1109/TPAMI.2020.3002836), 2020.
- 17 Huang G, Liu Z, Van Der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017. 2261–2269
- 18 Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016. 1637–1645
- 19 Tai Y, Yang J, Liu X. Image super-resolution via deep recursive residual network. In: Proceeding of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017. 2790–2798
- 20 Dong C, Loy C C, Tang X. Accelerating the super-resolution convolutional neural network. In: Proceedings of the European Conference on Computer Vision, Springer, Cham, 2016. 391–407
- 21 Shi W, Caballero J, Huszár F, Totz J, Aitken A P, Bishop R, Rueckert D, Wang Z. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016. 1874–1883
- 22 Li Jun-Cheng, Fang Fa-Ming, Mei Kang-Fu, Zhang Gui-Xu. Multiscale residual network for image super-resolution. In: Proceedings of the European Conference on Computer Vision, Springer, Cham, 2018. 527–542
- 23 Lai W, Huang J, Ahuja N, Yang M. Deep Laplacian pyramid networks for fast and accurate super-resolution. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017. 5835–5843
- 24 Agustsson E, Timofte R. NTIRE 2017 challenge on single image super-resolution: Dataset and study. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, 2017. 1122–1131
- 25 Deng J, Dong W, Socher R, Li L, Li Kai and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In: Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009. 248–255
- 26 Szegedy C, Liu Wei, Jia Yang-Qing, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015. 1–9
- 27 Bevilacqua M, Roumy A, Guillemot C, Alberi-Morel M L. Low-complexity single image super-resolution based on nonnegative neighbor embedding. In: Proceedings of the British Machine Vision Conference, 2012. 1–10
- 28 Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations. In: Proceedings of the International Conference on Curves and Surfaces, Springer, Berlin, Heidelberg, 2010. 711–730
- 29 Arbeláez P, Maire M, Fowlkes C, Malik J. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(5): 898–916
- 30 Huang J, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015. 5197–5206
- 31 Matsui Y, Ito K, Aramaki Y, Fujimoto A, Ogawa T, Yamasaki T, Aizawa K. Sketch-based manga retrieval using Manga109 dataset. *Multimedia Tools & Applications*, 2017, **76**(20): 21811–21838
- 32 Zhou Wang, Bovik A C, Sheikh H R, Simoncelli E P. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600–612
- 33 He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015. 1026–1034
- 34 Kingma D P, Ba J. Adam: A method for stochastic optimization, arXiv preprint, arXiv: 1412.6980, 2014.
- 35 Bi Min-Min. Research on image super-resolution technology based on deep learning [Master thesis]. Harbin Institute of Technology, China, 2020.
(毕敏敏. 基于深度学习的图像超分辨率技术研究 [硕士学位论文]. 哈尔滨工业大学, 中国, 2020.)
- 36 Li Bin, Yu Xia-Qiong, Wang Ping, Fu Rui-Gang, Zhang Hong. A survey of single image super-resolution reconstruction based on deep learning. *Computer Engineering and Science*, 2021, **43**(01): 112–124
(李彬, 喻夏琼, 王平, 傅瑞罡, 张虹. 基于深度学习的单幅图像超分辨率重建综述. *计算机工程与科学*, 2021, **43**(01): 112–124)
- 37 Yang X, Mei H, Zhang J, Xu K, Yin B, Zhang Q, Wei X. DRFN: Deep recurrent fusion network for single-image super-resolution with large factors. *IEEE Transactions on Multimedia*, 2019, **21**(2): 328–337
- 38 Hui Z, Wang X, Gao X. Fast and accurate single image super-resolution via information distillation network. In: Proceeding of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018.723–731
- 39 Zhang K, Zuo W, Zhang L. Learning a single convolutional super-resolution network for multiple degradations. In: Proceeding of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, 2018. 3262–3271
- 40 He Z, Cao Y, Du L, Xu B, Yang J, Cao Y, Tang S, Zhuang Y. MRFN: Multi-receptive-field network for fast and accurate single image super-resolution. *IEEE Transactions on Multimedia*, 2020, **22**(4): 1042–1054



孙超文 南京信息工程大学电子与信息工程学院硕士研究生。2018 年获得南京理工大学紫金学院电子工程与光电技术系学士学位。主要研究方向为深度学习和计算机视觉。

E-mail: 20181219071@nuist.edu.cn

(**SUN Chao-Wen** Master student

at the School of Electronic and Information Engineering, Nanjing University of Information Science and Technology. She received her bachelor degree in Electronic Engineering and Optoelectronic Technology from Zijin College, Nanjing University of Science and Technology in 2018. Her research interest covers deep learning and computer vision.)



陈晓 南京信息工程大学电子与信息工程学院教授。主要研究方向为现代电子系统设计, 信号与信息处理, 图像处理, 超声成像等。本文通信作者。

E-mail: chenxiao@nuist.edu.cn

(**CHEN Xiao** Professor at the

School of Electronic and Information Engineering, Nanjing University of Information Science and Technology. His research interest covers modern electronic system design, signal and information processing, image processing, and ultrasonic imaging. Corresponding author of this paper.)