



计算机工程与应用
Computer Engineering and Applications
ISSN 1002-8331, CN 11-2127/TP

《计算机工程与应用》网络首发论文

题目：跨尺度多分支网络的单帧图像超分辨率重建
作者：刘南艳，马圣祥，魏鸿飞
网络首发日期：2022-04-28
引用格式：刘南艳，马圣祥，魏鸿飞. 跨尺度多分支网络的单帧图像超分辨率重建[J/OL]. 计算机工程与应用.
<https://kns.cnki.net/kcms/detail/11.2127.tp.20220427.1735.004.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。

跨尺度多分支网络的单帧图像超分辨率重建

刘南艳, 马圣祥, 魏鸿飞

西安科技大学 计算机科学与技术学院, 西安 710699

摘要: 超分辨率重建算法是计算机视觉领域的重点研究问题之一, 目前各个领域对图像要求逐渐提高, 现有的超分辨率重建算法通过加深单一网络深度来提升图像质量, 忽略了重建时的计算复杂度问题, 同时缺少利用图像自身信息来提升图像的重建效果。针对以上问题提出了一种跨尺度多分支的单帧图像超分辨率重建网络, 跨尺度模块探索图像内部相似区域的相关性, 用于提升重建图像细节信息的能力, 并获得更好的视觉效果。多分支结构将图像中重建难度不同的区域, 通过不同复杂度的网络分别进行重建, 解决了深度网络计算复杂度高的问题。实验结果表明该模型在 Urban100 和 Manga109 测试集上相比其他方法具有更高的峰值信噪比和结构相似度。

关键词: 超分辨率重建; 卷积神经网络; 跨尺度; 多分支

文献标志码:A 中图分类号:TP391.4 doi: 10.3778/j.issn.1002-8331.2203-0307

Super-resolution reconstruction of single-frame images in cross-scale multi-branching networks

LIU Nanyan, MA Shengxiang, WEI Hongfei

College of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an 710699, China

Abstract: Super-resolution reconstruction algorithm is one of the key research problems in the field of computer vision. At present, the image requirements in various fields are gradually increasing, and the existing super-resolution reconstruction algorithms improve the image quality by deepening a single network depth, ignoring the computational complexity problem when reconstructing, and lacking the use of image own information to improve the image reconstruction effect. To address the above problems, a cross-scale multi-branching single-frame image super-resolution reconstruction network is proposed. The cross-scale module explores the correlation of similar regions within the image and is used to enhance the ability to reconstruct image detail information and obtain better visual effects. The multi-branching structure reconstructs regions in the image with different reconstruction difficulty by networks of different complexity separately, solving the problem of high computational complexity of deep networks. Experimental results show that the model has higher peak signal-to-noise ratio and structural similarity on Urban100 and Manga109 test sets compared with other methods.

Key words: super-resolution reconstruction; convolutional neural networks; cross-scale; multi-branching

基金项目: 国家自然科学基金(62002285)。

作者简介: 刘南艳(1966 -),女,硕士,副教授,研究方向为图像处理、机器学习,E-mail: 513318452@qq.com;马圣祥(1998 -),男,硕士研究生,研究方向为图像处理、深度学习;魏鸿飞(1996 -),男,硕士研究生,研究方向为图像处理、模式识别。

超分辨率重建是当前计算机视觉领域的重要研究方向之一,受到国内外研究者的广泛关注。图像超分辨率(Super Resolution, SR)重建技术是将一幅模糊的低分辨率(Low Resolution, LR)图像通过一定的算法转换成一幅高分辨率(High Resolution, HR)图像的过程。随着社会发展和人们对于获得高质量图像需求的不断提高,超分辨率重建技术在生产生活的各个领域得到广泛应用,如:医学影像^[1-4]、人脸识别^[5-8]和卫星成像^[9-12]等领域,其中医学影像中有大量相似的小细胞结构,人脸中有许多相同的器官,卫星图像中存在相同的建筑设施,图像中重复的结构都能使用跨尺度来进行重建。目前超分辨率算法主要分为基于单帧图像的超分辨率重建算法^[13-18]和基于视频的超分辨率重建算法^[19-22],本文主要研究基于单帧图像的超分辨率算法。

1 相关工作

图像超分辨率重建技术最早可以追溯到 1964 年由 Harris 提出的频谱外推法,随后由 Tsai 和 Huang 等人^[23]最先提出运用基于重构方法中的频域法来解决超分辨率重建问题,将低分辨率图像和高分辨图像进行傅里叶变换,使其在频域之间建立线性关系来完成图像重建,之后此类问题受到研究人员的广泛关注。Freeman 等人在 2002 年首次提出基于样例方法的图像自相似性超分辨率重建,揭开了样例方法在超分辨率重建中的应用。然而传统基于重构和样例学习的方法计算复杂而且不能有效恢复出原始图像的信息,目前应用最广的深度学习方法取得了不错的结果,与传统超分辨率重建技术相比,深度学习对于重建出来的图像具有更好的视觉效果。

2014 年 Chao Dong 等^[24]提出了一种超分辨率卷积神经网络(Super-Resolution Convolutional Neural Network, SRCNN)模型,该方法仅仅使用了三层卷积层,通过 9×9 的卷积核来提取一组特征映射、 1×1 的卷积核进行特征非线性映射和 5×5 的卷积核结合空间邻域内预测来生成最终的高分辨率图像。作为深度学习方法在 SR 技术中的开山之作,为图像超分辨率重建技术在深度学习方面应用奠定了基础。为了解决之前 SRCNN 中重建时计算开销大的问题,Chao Dong 等^[25]又在 SRCNN 的基础上进行改进提出了一种加速超分辨率卷积神经网络(Fast Super-Resolution Convolutional Neural Networks, FSRCNN),虽然 SRCNN 在超分辨率重建中取得了不错的结果,但由于其大量的计算开销阻碍实际应用,为此原来的研究团队以卷

积核大小和映射层数量互换,用卷积核减小换取映射层增多。这样相较于之前提出的 SRCNN 虽然在效果上没有得到很大提升,但是大幅度节省了计算开销,实现超过 40 倍的最终加速度,在不损失恢复图像质量的前提下实现模型更高训练速度。为了解决之前方法重建之后图像的效果与原图像存在较大差距, Kim 等^[26]提出了一种高精度单图像超分辨率(Accurate Image Super-Resolution Using Very Deep Convolutional Networks, VDSR)方法,借助 VGG 网络的灵感,将网络层数加深到 20 层,确实也取得了不错的效果,但是随着网络层数的加深,收敛速度随之受到制约,为此其引入残差结构^[27]很好的避免了此问题,在计算开销和视觉效果面均取得了不错的成绩,但是在基准数据集上的性能仍有一定的提升空间。随后 Lim 等人^[28]提出了增强的深度超分辨率网络(Enhanced Deep Residual Networks for Single Image Super-Resolution, EDSR),去除了 SRResNet 中的批量归一化(Batch Normalization, BN)层,从而可以通过加大模型结构来提升性能。从 SRCNN 至今各种基于卷积网络的重建算法层出不穷,随着网络结构向着更深、更宽、更复杂方向发展,网络模型对于图像特征信息提取更加全面、具有更大的感受野,利用更多的特征信息也会得到更好的图像质量和视觉效果,但是正因为当前网络过于复杂的结构导致其计算开销较高。

生成对抗网络^[29-30](Generative Adversarial Networks, GAN)提出以后广泛应用于图像处理、模式识别等各个领域。该网络将零和博弈的思想运用到训练神经网络中,生成器根据随机数来不断创造出新的数据,进而生成一幅图像去交给鉴别器,鉴别器尝试去判断生成器所发数据的真实性,二者相互博弈,最后不断优化提升达到纳什均衡。Ledig 等^[31]首次提出将 GAN 网络运用于 SR 技术中,其使用 16 块深度 ResNet 针对 MSE 进行优化,通过 VGG 网络特征图上的计算损失替换 MSE 内容损失,同时提出了一种平均意见得分(Mean Opinion Score, MOS)评价指标,在 MOS 上取得了不错的成绩,为之后的研究者在 SR 技术上开辟了另一条道路,但是对于标准的峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似度(Structural Similarity, SSIM)评价指标,其重建结果并不如之前模型。为了解决 SRGAN 网络在图像重建过程中产生的伪影问题,Wang 等^[32]在 SRGAN 网络的基础上提出了一种增强的超分辨率生成对抗网络(Enhanced Super-Resolution Generative Adversarial Networks, ESRGAN),由于残差密集块具有易于训练

和更高容量的特点,因此将其引入原始 SRGAN 网络中,同时删除了原有结构中的 BN 层,使得改进的网络对于重建后的图像具有更锐利的边缘和更少的伪影,但是其对于图像细节部分的重建质量仍然不够理想。

针对网络模型计算复杂度高的问题,Ahn 等^[33]提出了一种级联残差网络(Fast, Accurate, and Lightweight Super-Resolution with Cascading Residual Network, CARN)更加轻量化,然而该网络的重建结果不够理想。EDSR 网络因为其深层次的网络结构重建效果相对理想,然而巨大的计算开销成为了该网络的主要问题。为此本文提出了一种跨尺度多分支的超分辨率重建算法,在 Urban100 和 Manga109 公共数据集上优于目前的重建模型,主要工作如下:

(1)构建了一种跨尺度模块,该模块可以利用图像中相似区域的相关性,通过图像内部的高分辨率区域来辅助重建图像中相似的低分辨率区域,从而加强恢复图像细节信息,提高图像整体质量,获得更好的视觉效果。

(2)构建了一个多分支网络结构,依据图像自身具有的结构特性,将不同复杂度的子图像分别用不同网络重建,来降低重建时的计算开销。

(3)在 Urban100 和 Manga109 测试集上验证该网络的有效性。

2 跨尺度多分支网络

2.1 跨尺度模块

图像内部包含大量的资源信息,观察图像可以发现每张图像中都有许多相似的边角结构。如图 1 所示,建筑物图像上有很多相同的窗户,虽然这些窗户在图像中的远近高低各不相同,但它们都有着相似的结构。在超分辨率重建时可以考虑图像中长距离特征的相似性,以加强恢复细小结构的纹理信息。



图 1 图像相似信息示意图

Fig.1 Image similarity information schematic

卷积层可以对图像中的特征进行提取,但是它提

取了图像中的全部特征,这样大尺度区域与小尺度区域以同样的方式被对待,导致图像小尺度区域重建效果不理想,引入该模块后可以利用提取到的大尺度区域特征辅助重建小尺度区域的纹理,为纹理信息分配更多的内容,进而使得超分辨率重建之后的高清图像在细节方面有更好的视觉效果,如图 2 所示能看出卷积只是简单提取出图像中的特征信息,跨尺度模块可以利用这些特征信息中的重复结构来提升重建效果。

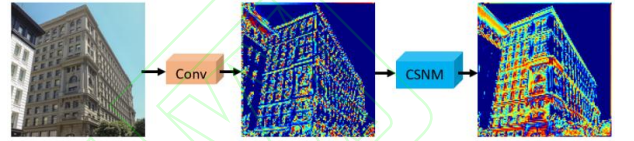


图 2 跨尺度模块和卷积特征图

Fig.2 Cross-scale modules and convolutional feature maps

非局部相关性可以通过总结整个图像的相关特征来探索自我范式。在形式上,给定图像特征映射 X ,非局部相关性定义为:

$$Z_{i,j} = \sum_{g,h} \frac{\exp(\phi(X_{i,j}, X_{g,h}))}{\sum_{u,v} \exp(\phi(X_{i,j}, X_{u,v}))} \varphi(X_{g,h}) \quad (1)$$

其中 $(i,j), (g,h)$ 和 (u,v) 为 X 坐标对, $\varphi(X_{g,h})$ 为特征变换函数, $\phi(\cdot)$ 为度量相似度的相关函数,该函数定义为:

$$\phi(X_{i,j}, X_{g,h}) = \theta(X_{i,j})^T \delta(X_{g,h}) \quad (2)$$

其中 $\theta(\cdot)$ 和 $\delta(\cdot)$ 是特征变换函数,但是该非局部相关性是在相同尺度下进行衡量的,无法利用大尺度结构来辅助重建小尺度结构。

因此本文研究跨尺度相关性方法,把上面非局部相关性公式扩展到跨尺度相关性,首先寻找 LR 图像中小尺度区域和大尺度区域之间的相关性,然后学习大尺度区域与小尺度区域的映射关系来辅助重建图像中小尺度区域。图像重建的上采样因子为 s , 给定一个大小为 (W, H) 的特征图 X , 将特征图 X 通过双线性插值方式进行 s 倍下采样得到新的特征图 Y , 然后将 X 中 $p \times p$ 的小块与下采样后的 Y 中同样大小的 $p \times p$ 小块通过匹配函数进行匹配。再将所有 X 中提取到大小为 (sp, sp) 的特征块进行加权相加,得到大小为 (sW, sH) 的特征块 Z 辅助 X 进行重建,跨尺度重建公式为:

$$Z_{si,si}^{sp \times sp} = \sum_{g,h} \frac{\exp(\phi(X_{i,j}^{p \times p}, Y_{g,h}^{p \times p}))}{\sum_{u,v} \exp(\phi(X_{i,j}^{p \times p}, Y_{u,v}^{p \times p}))} \varphi(X_{sg,sh}^{sp \times sp}) \quad (3)$$

其中, $Z_{si,sj}^{sp \times sp}$ 是位于 (si, sj) 处大小为 $sp \times sp$ 的加权结果, $X_{i,j}^{p \times p}$ 与 $Y_{g,h}^{p \times p}$ 通过相似度函数实现块之间的匹配, 利用特征图中提取的特征块 $X_{sg,sh}^{sp \times sp}$ 获得加权平均特征

$Z_{si,sj}^{sp \times sp}$, 主要是利用 X 在 Y 中的映射来对 X 中的小块进行重建, 从原图像内在资源中挖掘更多丰富的高频细节信息。

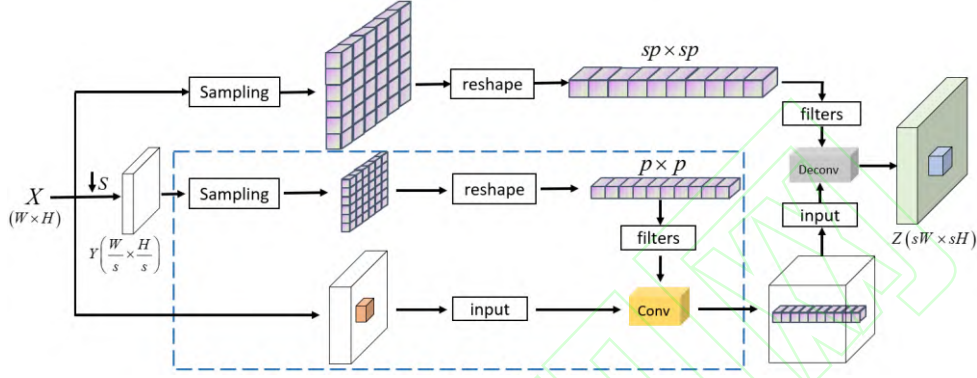


图3 跨尺度模块结构

Fig.3 Cross-scale module structure

如图3所示为跨尺度模块的网络结构图, 其中蓝色的框内主要用于跨尺度相似性匹配, 蓝色框上面的分支是从 LR 图像中提取原始的大尺度特征块, 由于空间维度的差异性, 使用特征相似性的度量方式是不可行的, 需要将 X 进行下采样得到对应的低分辨率图像 Y , 在 Y 中搜寻多个与 X 特征相似的特征块, 利用该特征的信息就能找到需要辅助重建的大尺度特征块。

图像通过跨尺度模块提取特征之后, 将输出结果扩大到对应上采样倍数的图像尺寸, 实现由小尺度到大尺度的映射操作。

如图4所示, 在不同上采样因子下, 将卷积生成的 r^2 个小尺度特征图重新排列组合, 重构放大后得到的 SR 图像。该算子可以描述为:

$$I_{SR} = PS(w * X_L + b) \quad (4)$$

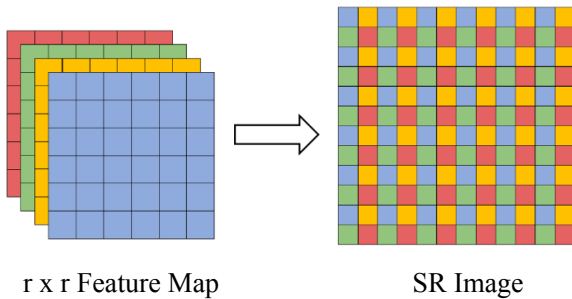


图4 小尺度与大尺度映射关系

Fig.4 Small-scale and large-scale mapping relationships

其中 I_{SR} 是小尺度 SR 后的空间张量。 X_L 为小尺度特征图, w 和 b 是可学习的网络权值和偏差。 $PS(\cdot)$ 是

一个周期变换算子, 它将一个形状为 $h \times w \times c \times r^2$ 的张量重新排列成形状为 $hr \times wr \times c$ 的张量, 以实现上采样。

本文使用 L_1 损失函数来计算重建后的小尺度区域与大尺度区域之间的误差, 公式可以表示为:

$$L_1 = \frac{1}{N} \sum_{i=1}^N \|I_{SR} - I_L\| \quad (5)$$

其中 I_L 为上文中大尺度特征 Z 的空间张量, N 为样本个数。

2.2 多分支网络

通过观察可以发现每幅图像都包含复杂的纹理区域和稀疏的平滑区域, 对于超分辨率重建而言, 平滑区域比复杂区域更容易处理, 利用这一特性, 我们可以在重建一整幅大图像时采用适当的 SR 网络来处理不同区域的子图像。如图5所示, 简单的网络处理平滑区域(绿色), 复杂的网络处理纹理区域(红色), 黄色区域复杂度介于前两者之间。



图5 图像分区域示意图

Fig.5 Image division area schematic

依据上图的特性, 本文构建了一种多分支网络结构, 整体网络框架如图6所示, 该网络以低分辨率图像为输入, 输出高分辨图像。多分支网络分为分类模块和超分辨率网络模块两个部分。首先将一整幅 LR

图像分割成多个子图像 $\{x_i\}_{i=1}^N$ ，然后用分类网络接收每一个子图像 x_i ，分类网络由五个卷积层，一个平均池化层和一个全连接层组成，通过底层特征将输入图像分为 M 类 ($M=3$) 并生成概率向量 $[P_1(x_i), \dots, P_M(x_i)]$ ，超分辨率模块包含 M 个 SR 网络分支 $\{f_{SR}^j\}_{j=1}^M$ 来处理不同的输入，通过选择最大概率值 $K = \arg \max_j P_j(x_i)$ 来确定使用的 SR 分支。轻量级网络重建速度快，但是缺少对细节信息的恢复能力，因

此平滑图像使用轻量级网络 CARN 进行超分辨率重建。深度网络重建速度慢，但是效果相对理想，因此较复杂图像使用 EDSR 网络进行重建。由于最复杂的子图像中往往包含着许多相似的边角结构，因此使用 2.1 节中提出的跨尺度模块构建了一种跨尺度超分辨率 (Cross-Scale Super-Resolution, CSSR) 网络进行重建，不同分支对 x_i 进行重建后得到 $y_i = f_{SR}^j(x_i)$ ，最后将输出的子图像合并得到最终的 SR 图像。

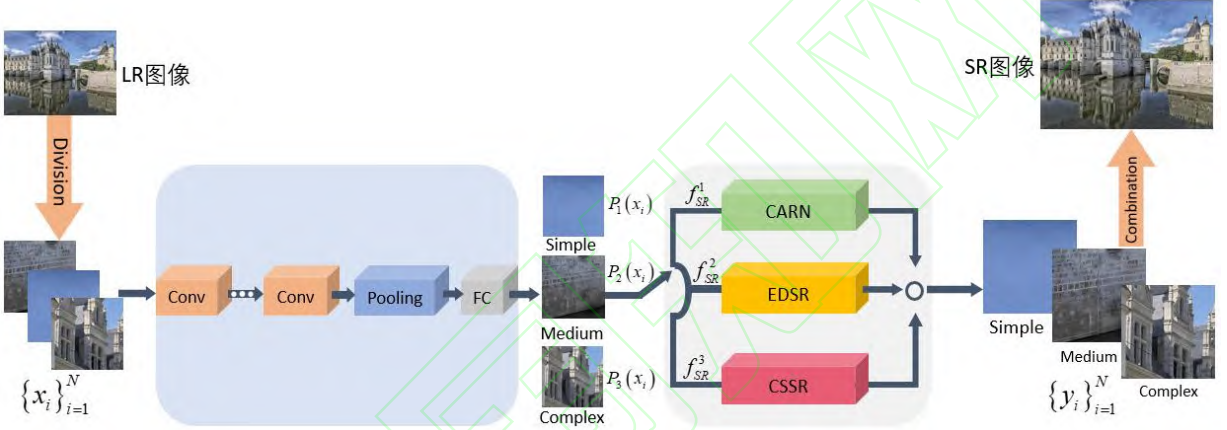


图 6 整体网络结构

Fig.6 Overall Network Structure

如图 7 所示，分别展示了三种分支的网络结构，轻量级网络 CARN 通过在局部和全局中使用级联机制集成多个网络层的特性，以 Cascading Block 作为网络的主要结构，来反映不同级别的输入表示，以便于接收更多的信息。

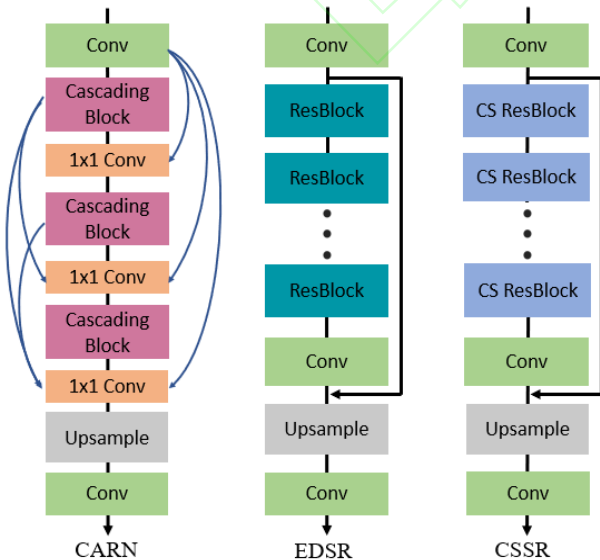


图 7 CARN、EDSR 和 CSSR 三种重建分支网络结构

Fig.7 Three reconstructed branch network structures of CARN, EDSR and CSSR

EDSR 网络将原始残差结构进行改进，得到该网络的 ResBlock 结构，通过堆叠 16 层 ResBlock 结构来加强图像的对比度信息，提升网络输出质量。

针对于纹理细节信息的恢复本文提出在 EDSR 网络的残差结构中引入跨尺度模块 (Cross-Scale Block, CSBlock)，专门处理复杂子图像中的相似边、角等结构，得到 CSSR 网络的主要结构跨尺度残差块 (Cross-Scale Residual Block, CS ResBlock)。

为了加强网络对于细节信息的捕获能力，提升网络超分辨率后图像的质量，本文在 EDSR 中每个 ResBlock 内部的卷积层提取完特征之后引入 CS Block，该模块针对细节部分提供更多的信息辅助重建，防止图像中细节信息丢失，提升特征对于高频信息的表达能力，加强图像重建后的视觉效果。相比 CARN、EDSR 在残差内部结构的改进有所不同，如图 8 是三种网络的残差结构内部进行对比，从图中可以看出 CARN 方法将原始残差结构内部的深度卷积替换成了分组卷积，选择适合的分组大小来降低参数得到该网络的残差结构，随后 EDSR 将传统方法中的 BN 层去除，因为 BN 层对于网络特征的规范限制了网络的灵活性，本文将

CS Block 与 ResBlock 相结合,得到用图像自身相似结构来重建图像的 CS ResBlock,提升图像中细节信息的重建能力。

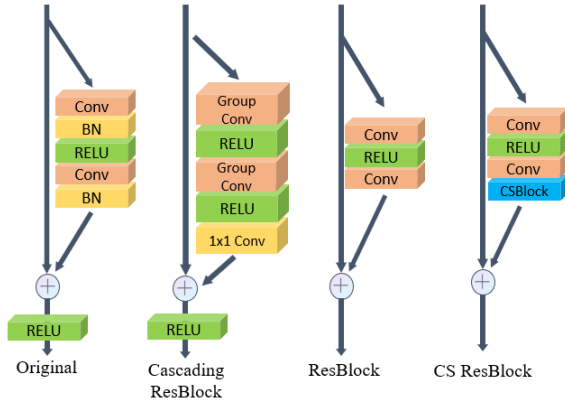


图8 文中方法与传统方法残差结构对比

Fig.8 Comparison of the residual structure of the method in the paper and the traditional method

3 实验结果及分析

3.1 实验数据及环境

为了获得更高质量的重建图像,本文采用DIV2K^[34]数据集对该网络模型进行训练,该数据集中总共包含1000张2K分辨率的RGB三通道图像,选用其中的800张图像用于训练,同时为了便于与经典网络模型在峰值信噪比和结构相似度评价指标上进行对比,本实验网络模型选用Urban100测试集和Manga109测试集中进行测试,其中Urban100测试集主要以建筑为主,包含各类相似结构的建筑设施图像数据集,Manga109数据集是以动漫为主要图像的数据集。

本次实验首先对训练图像进行预处理,将DIV2K数据集中的高分辨率图像通过Bicubic算法进行0.6、0.7、0.8和0.9倍的下采样生成更多的HR图像,以增大训练时的数据量,对这些图像进行进一步的下采样生成LR图像,然后从LR图像中裁剪尺寸为 32×32 的子图像,接下来输入网络中进行训练。训练过程中的批量大小设置为16,使用 L_1 损失函数和Adam优化器进行训练,其中 $\beta_1=0.9$, $\beta_2=0.999$,通过余弦退火策略来调整训练速率,初始学习速率设置为 10^{-3} ,最小的学习速率设置在 10^{-7} 。在测试过程中,将LR图像裁剪成 32×32 的子图像,步长为28,通过将4个像素点的重叠区域进行平均,以免重建后的图像出现明显分割痕迹。

PSNR作为重建图像的评价标准,一方面可以看出图像的重建难度,另一方面可以反映图像纹理结构

的复杂程度。如图9所示,通常PSNR值高的图像比较平滑,PSNR值低的图像具有复杂的纹理,本文通过大量实验后选用PSNR值为27.16dB和35.15dB作为分割子图像的依据,从图中可以看出取得了不错的效果。

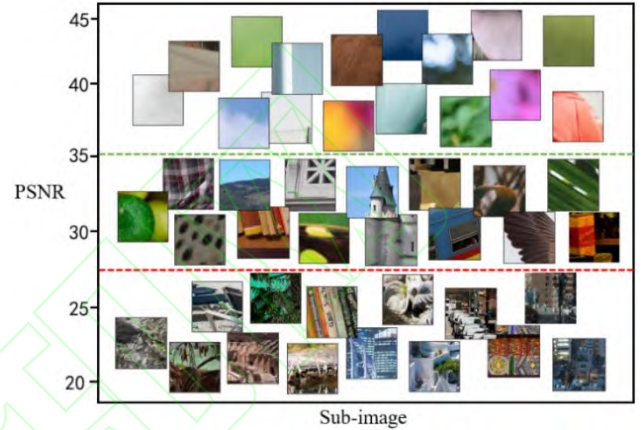


图9 子图像分类示意图

Fig.9 Sub-image classification schematic

为了保证实验质量,本次实验均采用统一的硬件配置,其中操作系统为Ubuntu系统、处理器为AMD EPYC 7642 48-Core Processor、显卡为NVIDIA GeForce RTX3090,并在Python3.6.13、CUDA11.0环境配置下运行,选用Pytorch1.7.0作为深度学习框架。

3.2 实验评价指标

通常图像在放大或压缩前后会出现不同程度的误差,为了客观的展现出超分辨后的图像与原始图像之间存在的差距,比较不同超分辨模型之间重建图像质量的差别,本文通过主观和客观两种不同方式对重建图像的质量进行评价,首先采用峰值信噪比和结构相似度对生成的图像进行客观评估,这两种方式是当前超分辨率重建方式中最普遍和广泛使用的客观评价指标,同时为了验证该模型良好的计算效率,本文使用浮点计算数来衡量模型计算的复杂度,对比突出多分支方法与以往单一分支方法计算复杂度的不同。

峰值信噪比是通过对比处理前后图像中对应像素点之间的差距进行计算来评估图像重建后的质量,其计算表达式为:

$$PSNR = 10 \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (6)$$

其中, n 为每像素之间的比特数,一般图像灰阶数为256,因此 n 通常取8,MSE为超分辨率重建前后图像的均方误差。由公式可以看出,重建前后图像的峰值

信噪比越大,说明图像失真越少,PSNR 单位为 dB。

与峰值信噪比不同,结构相似度是从图像亮度、对比度和结构三个方面来判断输出图像与原始图像之间的相似性,其计算表达式为:

$$SSIM(x, y) = l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma \quad (7)$$

式中, $l(x, y)$ 为亮度对比函数, $c(x, y)$ 为对比度对比函数, $s(x, y)$ 为结构对比函数,通过调节 α 、 β 、 γ 之间不同的权重分配来对以上这三个方面赋予不同的重要性。

亮度对比函数为:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (8)$$

对比度对比函数为:

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (9)$$

结构对比函数为:

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (10)$$

其中, μ_x, μ_y 表示图像 x, y 的均值, σ_x, σ_y 表示图像 x, y 的方差, σ_{xy} 表示图像 x, y 的协方差, C_1, C_2, C_3 为常数。

标准的客观评价指标虽然可以计算出图像重建前后的失真情况,但是其并没有考虑到人眼观察图像的视觉感受,往往人眼对于图像中像素点之间的差异敏感度较高,会出现客观评价指标与人眼主观感受不一致的情况。为此本文中新增主观评价指标,任意寻找 50 名参与者,从其他模型和本文模型的实验结果中各选取三幅重建图像,参与者在不知晓模型的情况下根据自己主观感受对比不同模型重建后的图像与原始图像进行打分,最终计算出参与者对于每张图像的平均分进行对比。

图像超分辨率重建技术除了衡量图像质量以外,其计算效率也是重要的衡量指标之一,高效、准确的重建出目标图像是当前该领域的重要目标,所以不仅要追求更好的视觉效果,对于批量处理图像时计算效率也是应当考虑的因素。因此本文采用浮点计算数(Floating Point Operations, FLOPs)作为衡量模型重建图像计算复杂度的标准,FLOPs 的计算公式如下:

$$FLOPs = 2HW(C_{in}K^2 + 1)C_{out} \quad (11)$$

其中, HW 代表输出图像特征图的大小, C_{in} 代表输入图像的通道数, C_{out} 为输出图像的通道数, K 代表卷积核的大小。

3.3 实验结果与分析

3.3.1 客观评价

为了验证平滑区域图像使用轻量级网络 CARN、较复杂区域图像使用 EDSR 网络与使用单一跨尺度网络重建之间可以获得大致相同的效果,选用三种不同分类的子图像分别放入 CARN、EDSR 和 CSSR 网络进行重建,再将三类子图像全部放入 CSSR 网络进行重建,对比重建后的峰值信噪比。

如表 1 所示,可以发现 CARN 和 CSSR 网络重建平滑子图像几乎没有差异,在较复杂子图像上,EDSR 与 CSSR 网络同样获得大致相同的性能。因此说明了通过多分支网络分区域重建图像来节约计算成本的有效性,且产生图像的质量大致相同。

表 1 不同子图像在三种分支下的 PSNR 对比

Table 1 Comparison of PSNR of different sub-images under three branches

Model	Simple	Medium	Hard
CARN	38.62dB	—	—
EDSR	—	26.43dB	—
CSSR	38.64dB	26.46dB	21.56dB

为了进一步验证实验的有效性,本文将文中方法与当前的经典网络重建图像效果进行对比,其中包括: Bicubic、SRCNN、VDSR、CARN、SRGAN、EDSR。同时将本文中提出的跨尺度单分支网络(CSSR)与跨尺度多分支网络(M-CSSR)也进行了比较,在 Urban100 数据集和 Manga109 数据集上分别对 2 倍、3 倍和 4 倍放大因子超分辨率重建对比峰值信噪比和结构相似度,结果如表 2 和表 3 所示,其中粗体字显示为当前比较中的最优结果。

从表中可以看出文中方法与其他方法相比具有更高的 PSNR 和 SSMI,其中客观结果表明 SRCNN 算法作为最初应用于深度学习领域的超分辨率算法,其理论是开创性的,但是重建效果不够理想,随后出现的深度学习超分辨率重建的在网络深度和宽度上都有一定程度的提升,VDSR、CARN 和 EDSR 通过对网

络结构的优化和对网络层数的加深,使得它们在客观评价指标上均有一定的提高,SRGAN 通过使用生成

对抗网络完成重建使性能得到一定提升,同时也为超分辨重建开辟了一种新的思路。

表2 不同算法在 Urban100 和 Manga109 数据集上 PSNR 测试结果

Table 2 PSNR test results of different algorithms on Urban100 and Manga109 datasets

数据集	倍数	Bicubic	SRCNN	VDSR	CARN	SRGAN	EDSR	CSSR (Ours)	M-CSSR (Ours)
Urban100	x2	26.88	29.50	30.76	31.92	32.23	32.93	33.42	33.39
	x3	24.46	26.24	27.14	28.06	28.20	28.80	29.36	29.36
	x4	23.14	24.52	25.18	26.07	26.12	26.64	27.42	27.42
Manga109	x2	30.80	35.60	37.22	38.36	38.05	39.10	39.25	39.25
	x3	26.95	30.48	32.01	33.50	33.54	34.17	34.49	34.46
	x4	24.89	27.66	28.83	30.47	30.48	31.02	31.38	31.38

表3 不同算法在 Urban100 和 Manga109 数据集上 SSMI 测试结果

Table 3 SSMI test results of different algorithms on Urban100 and Manga109 datasets

数据集	倍数	Bicubic	SRCNN	VDSR	CARN	SRGAN	EDSR	CSSR (Ours)	M-CSSR (Ours)
Urban100	x2	0.8403	0.8946	0.9140	0.9256	0.9295	0.9351	0.9384	0.9384
	x3	0.7349	0.7989	0.8279	0.8493	0.8535	0.8653	0.8723	0.8723
	x4	0.6577	0.7221	0.7524	0.7837	0.7871	0.8033	0.8157	0.8134
Manga109	x2	0.9339	0.9663	0.9750	0.9765	0.9607	0.9773	0.9786	0.9786
	x3	0.8556	0.9117	0.9340	0.9440	0.9448	0.9476	0.9483	0.9479
	x4	0.7866	0.8505	0.8809	0.9084	0.9087	0.9148	0.9259	0.9259

相较于前几种方法,文中算法在 Urban100 和 Manga109 数据集上,PSNR 和 SSMI 两种客观评价指标中都取得了不错的结果,本文也将单一的跨尺度网络与跨尺度多分支网络结构进行了比较,从结果中可以看出多分支结构与其单分支结构在峰值信噪比和结构相似度上略有差距,但是其差距并不明显,这两种数据集包含了丰富的重复块,例如一些边、小角,证明跨尺度非局部模块可以利用图像的内部信息提升重建效果,从而表明我们关注内部资源辅助重建是有效的。

传统的重建方法旨在通过更深的网络结构、更宽的网络宽度来得到更好的视觉效果,这种考虑方式不够全面,良好的视觉效果往往承受着巨大的计算代价。

因此在考虑图像峰值信噪比和结构相似度的提升时,还应当将计算开销作为重要的关注点,这样可以更全面的衡量模型的重建性能。本文用 Urban100 和 Manga109 测试集中图像重建的浮点计算数、实时性和模型参数量三个方面衡量模型的性能,如表4所示。从表中可以看出本文方法 CSSR 在 EDSR 的 ResBlock 中构建了 CS Block,因此 CSSR 网络相较于 EDSR 网络在参数上增加了 0.6M,实时性略微降低,重建时间增加了 0.12 秒。然而多分支结构的 M-CSSR 网络重建图像的实时性大幅度提升,单幅图像重建时间比 CSSR 减少了 0.28 秒,比 EDSR 减少了 0.16 秒。总体来说,本文方法从参数、浮点计算数和实时性方面整体优于 EDSR 方法。同时可以看出,子图像划分的数量对计算开销也有一定的影响,Manga109 测试集中存

在的平滑区域比 Urban100 测试集相对多,因此在划分子图像时 Manga109 测试集被划分到简单类别的子图像更多,重建时仅需要原来浮点计算数的 53%,而 Urban100 测试集浮点计算数只减少为原来的 56%。

表 4 不同模型性能对比结果

Table 4 Performance comparison results of different models				
Datasets	Indicators	EDSR	CSSR	M-CSSR
Urban100	Params[M]	12.6	13.2	25.1
	FLOPs[G]	854.6	1043.2	584.2
	PSNR[dB]	26.64	27.42	27.42
	Time[s]	1.31	1.43	1.15
Manga109	FLOPs[G]	847.9	1026.4	543.9
	PSNR[dB]	31.02	31.38	31.38
	Time[s]	1.27	1.41	1.09

3.3.2 主观评价

为了更好的体现不同峰值信噪比和结构相似度在图像中呈现出来的差距,本文从实验结果中选取三张高清图像,通过对于图像中的局部区域进行放大,对比不同方法在这三张图像的重建效果。高频信息是图像中颜色变化快的区域,例如图 10 中建筑物的窗户与墙面之间的边缘部分。可以从图中看出 Bicubic 重建出的图像相对模糊,重建效果与 HR 图像差距较大,对于低频信息与高频信息的捕获能力均不够理想。虽然 CARN 方法相较于 Bicubic 方法有所改进,但是其整体重建效果仍然模糊。随后提出的 EDSR 方法挖掘低频信息能力方面有所改进,从图中可以看出对于整体墙面的重建效果相对理想,但是图像中没有很好的利用内部信息对细小的结构进行辅助重建,仍然无法挖掘到颜色变化快的高频区域,如图中窗户与墙面之间的窗台区域,导致图像看上不够自然。因此,本文方法中的 CS Block 利用图像中内部资源的跨尺度相关性,主要针对前几种经典方法中挖掘高频信息不足,纹理细节区域重建效果不理想等问题进行改善,同时

使图像变得更加自然和真实,可以看出本文方法在图像中墙面与窗户等相关边界区域呈现出更多的信息。总体来看,本文方法从视觉效果上比以往方法更具有优势。

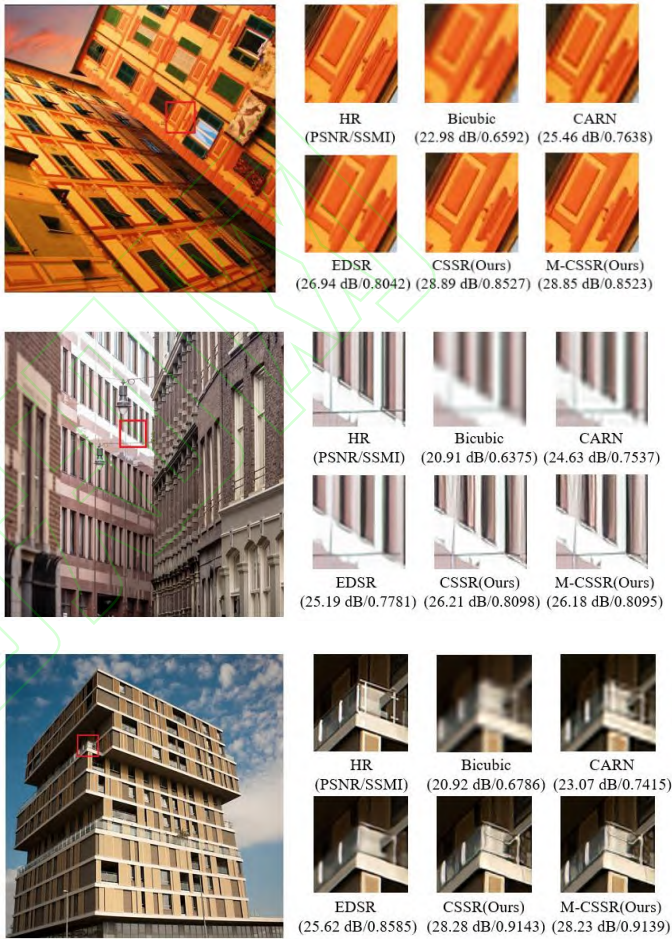


图 10 不同模型重建结果








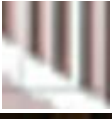


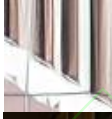







Fig.10 Reconstruction results of different models

同时,本文针对以上三幅图像采用主观打分的方式进行评价,将以上三幅图像中对应的放大区域进行随机排列且不被告知图像模型名称,让参与者仅凭借其主观感受,对不同模型从 1 至 10 分中进行打分,最终计算出 50 位参与者在不同模型上的平均得分来进行比较,如表 5 所示。

表 5 不同模型 FLOPs 对比结果

Table 5 Subjective evaluation results of different algorithms

Picture	Bicubic	CARN	EDSR	CSSR	M-CSSR
---------	---------	------	------	------	--------

HR	SR	Score	SR	Score	SR	Score	SR	Score	SR	Score
		5.48		6.48		7.52		8.74		8.68
		5.36		6.82		7.38		8.84		8.92
		5.74		6.72		6.88		8.46		8.36

4 结语

本文主要针对当前超分辨率模型纹理结构不清晰、细节信息不明显和网络过深导致重建计算成本过高的问题进行研究,提出了一种跨尺度多分支网络的单帧图像超分辨率重建方法。该模型依据不同区域图像重建难度不同,通过多分支网络的方式解决深度模型重建图像时计算成本过高的问题。此外在残差结构内部通过跨尺度模块,利用图像内部信息辅助重建图像中大量的细节区域,能使图像看起来更加真实自然、纹理细节更加清晰。

为了体现本文的超分辨率模型相比于以往超分辨率模型的优势,采用客观评价与主观评价相结合的方式对不同模型进行比较,在两个公开数据集上对它们之间峰值信噪比和结构相似度进行比较,同时采用主观评分的方式对模型重建图像的视觉效果进行评价。为了体现多分支结构可以降低模型的计算成本,本文采用浮点计算数对模型的复杂程度进行了比较,最终对比表明本文模型不仅能够有效恢复超分辨率图像的纹理细节部分,使图像重建后更加真实自然,具有更好的复原效果,而且模型计算复杂度相比于经典方法有所减少。

参考文献:

- [1] 席志红,侯彩燕,袁昆鹏.基于残差网络的医学图像超分辨率重建[J].计算机工程与应用,2019,55(19):191-197.
XI Zhihong, HOU Caiyan, YUAN Kunpeng. Medical Image Super Resolution Reconstruction Based on Residual Network[J]. Computer Engineering and Applications, 2019, 55(19): 191-197.
- [2] Li Y, Sixou B, Peyrin F. A review of the deep learning methods for medical images super resolution problems[J]. IRBM, 2021, 42(2): 120-133.
- [3] Qiu D, Zheng L, Zhu J, et al. Multiple improved residual networks for medical image super-resolution[J]. Future Generation Computer Systems, 2021, 116: 200-208.
- [4] 邢晓羊,魏敏,符颖.基于特征损失的医学图像超分辨率重建[J].计算机工程与应用,2018,54(20):202-207.
XING Xiaoyang, WEI Min, FU Ying. Super-resolution reconstruction of medical images using feature-based loss[J]. Computer Engineering and Applications, 2018, 54(20): 202-207.
- [5] 蒋文杰,罗晓曙,戴沁璇.基于对抗网络人脸超分辨率重建算法研究[J].计算机工程与应用,2021,57(11):219-223.
JIANG Wenjie, LUO Xiaoshu, DAI Qinxuan. Research on Face Super-Resolution Reconstruction Algorithm Based on Generative Adversarial Networks[J]. Computer Engineering and Applications, 2021, 57(11): 219-223.
- [6] Chen Y, Phonevilay V, Tao J, et al. The face image super-resolution algorithm based on combined representation learning[J]. Multimedia Tools and Applications, 2021, 80(20): 30839-30861.
- [7] Li M, Zhang Z, Yu J, et al. Learning face image super-resolution through facial semantic attribute transformation and self-attentive structure enhancement[J]. IEEE Transactions on Multimedia, 2020, 23: 468-483.
- [8] 刘颖,董占龙,卢津,王富平.多注意力域的稠密连接人脸超分算法[J].计算机工程与应用,2020,56(20):179-185.
LIU Ying, DONG Zhanlong, LU Jin, WANG Fuping. Densely Connected Face Super-Division Algorithm with Multiple Attention Domains[J]. Computer Engineering and Applications, 2020, 56(20): 179-185.
- [9] 夏皓,吕宏峰,罗军,蔡念.图像超分辨率深度学习研究及应用进展[J].计算机工程与应用,2021,57(24):51-60.
XIA Hao, LYU Hongfeng, LUO Jun, CAI Nian. Survey on

- Deep Learning Based Image Super-Resolution[J]. Computer Engineering and Applications, 2021, 57(24): 51-60.
- [10] Lu T, Wang J, Zhang Y, et al. Satellite image super-resolution via multi-scale residual deep neural network[J]. Remote Sensing, 2019, 11(13): 1588.
- [11] Shin C, Kim S, Kim Y. Satellite Image Target Super-Resolution With Adversarial Shape Discriminator[J]. IEEE Geoscience and Remote Sensing Letters, 2020.
- [12] 贺智, 贺丹. 基于深度学习的高分四号卫星图像超分辨率重建[J]. 遥感学报, 2020, 24(12): 1500-1510.
- He Z and He D. 2020. Deep learning-based super-resolution for GF-4 satellite imagery. Journal of Remote Sensing (Chinese), 24(12): 1500-1510.
- [13] Lan R, Sun L, Liu Z, et al. Cascading and enhanced residual networks for accurate single-image super-resolution[J]. IEEE Transactions on Cybernetics, 2020, 51(1): 115-125.
- [14] Kong X, Zhao H, Qiao Y, et al. Classsr: A general framework to accelerate super-resolution networks by data characteristic[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021: 12016-12025.
- [15] Hu Y, Gao X, Li J, et al. Single image super-resolution with multi-scale information cross-fusion network[J]. Signal Processing, 2021, 179: 107831.
- [16] 李现国, 冯欣欣, 李建雄. 多尺度残差网络的单幅图像超分辨率重建[J]. 计算机工程与应用, 2021, 57(7): 215-221.
- LI Xianguo, FENG Xinxin, LI Jianxiong. Single Image Super-Resolution Reconstruction Based on Multi-scale Residual Network[J]. Computer Engineering and Applications, 2021, 57(7): 215-221.
- [17] 刘颖, 朱丽, 林庆帆, 李莹华, 王富平, 卢津. 图像超分辨率技术的回顾与展望[J]. 计算机科学与探索, 2020, 14(2): 181-199.
- LIU Ying, ZHU Li, LIM Kengpan, LI Yinghua, WANG Fuping, LU Jin. Review and Prospect of Image Super-Resolution Technology[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(2): 181-199.
- [18] 唐家军, 刘辉, 胡雪影. 功能型复合深度网络的图像超分辨率重建[J]. 计算机科学与探索, 2020, 14(8): 1368-1379.
- TANG Jiajun, LIU Hui, HU Xueying. Image Super-resolution Reconstruction of Functional Composite Deep Network[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(8): 1368-1379.
- [19] Isobe T., Jia X., Gu S., Li S., Wang S., Tian Q. (2020) Video Super-Resolution with Recurrent Structure-Detail Network. In: Vedaldi A., Bischof H., Brox T., Frahm JM. (eds) Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science, vol 12357. Springer, Cham.
- [20] Liu C, Sun D. A bayesian approach to adaptive video super resolution[C]//CVPR 2011. IEEE, 2011: 209-216.
- [21] Tian Y, Zhang Y, Fu Y, et al. Tdan: Temporally-deformable alignment network for video super-resolution[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 3360-3369.
- [22] Habib Z. DSTnet: Deformable Spatio-Temporal Convolutional Residual Network for Video Super-Resolution[J]. Mathematics, 2021, 9.
- [23] Tsai R Y and Huang T S. Multipleframe image restoration and registration[A]. Advances in computer vision and image processing. Greenwich, CT: JAI Press Inc., 1984: 317-339.
- [24] Chao D, Chen C L, He K, et al. Learning a Deep Convolutional Network for Image Super-Resolution[J]. ECCV, 2014.
- [25] Chao D, Chen C L, Tang X. Accelerating the Super-Resolution Convolutional Neural Network[C]// European Conference on Computer Vision. Springer, Cham, 2016.
- [26] Kim J, Lee J K, Lee K M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks[C]//IEEE Conference on Computer Vision & Pattern Recognition. IEEE, 2016.
- [27] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. IEEE, 2016.
- [28] Lim B, Son S, Kim H, et al. Enhanced Deep Residual Networks for Single Image Super-Resolution[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2017.
- [29] 汪鑫耘, 李丹. 生成式对抗网络在超分辨率图像重建中的应用[J]. 计算机科学与探索, 2020, 14(4): 680-687.
- WANG Xinyun, LI Dan. Application of Generative Adversarial Network in Super-Resolution Image Reconstruction[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(4): 680-687.
- [30] 彭晏飞, 高艺, 杜婷婷, 桑雨, 訾玲玲. 生成对抗网络的单图像超分辨率重建方法[J]. 计算机科学与探索,

- 2020, 14(9): 1612-1620.
- PENG Yanfei, GAO Yi, DU Tingting, SANG Yu, ZI Lingling. Single Image Super-Resolution Reconstruction Method for Generative Adversarial Network[J]. Journal of Frontiers of Computer Science and Technology, 2020, 14(9): 1612-1620.
- [31] Ledig C , Theis L , F Huszar, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[J]. IEEE Computer Society, 2016.
- [32] Wang X, Yu K, Wu S, et al. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks[J]. Springer, Cham, 2018.
- [33] Ahn N, Kang B, Sohn K A. Fast, accurate, and lightweight super-resolution with cascading residual network[C]// Proceedings of the European conference on computer vision (ECCV). 2018: 252-268.
- [34] Agustsson E, Timofte R. Ntire 2017 challenge on single image super-resolution: Dataset and study[C]//Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017: 126-135.