



基于广泛激活深度残差网络的图像超分辨率重建

王凡超, 丁世飞

引用本文:

王凡超, 丁世飞. 基于广泛激活深度残差网络的图像超分辨率重建[J]. *智能系统学报*, 2022, 17(2): 440–446.

WANG Fanchao, DING Shifei. Image super-resolution reconstruction based on widely activated deep residual networks[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(2): 440–446.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202106023>

您可能感兴趣的其他文章

基于生成对抗网络的机载遥感图像超分辨率重建

Super-resolution reconstruction of airborne remote sensing images based on the generative adversarial networks
智能系统学报. 2020, 15(1): 74–83 <https://dx.doi.org/10.11992/tis.202002002>

基于稀疏表示与线性回归的图像快速超分辨率重建

Rapid super-resolution image reconstruction based on sparse representation and linear regression
智能系统学报. 2017, 12(1): 8–14 <https://dx.doi.org/10.11992/tis.201603039>

区域损失函数的孪生网络目标跟踪

Regional loss function based siamese network for object tracking
智能系统学报. 2020, 15(4): 722–731 <https://dx.doi.org/10.11992/tis.201910005>

基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects
智能系统学报. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

一种恢复图像自然色彩的重构方法

Reconstruction method for restoring the natural color of an image
智能系统学报. 2019, 14(5): 877–881 <https://dx.doi.org/10.11992/tis.201805011>

基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism
智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

 微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202106023

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20211012.1937.008.html>

基于广泛激活深度残差网络的图像超分辨率重建

王凡超¹, 丁世飞^{1,2}

(1. 中国矿业大学 计算机科学与技术学院, 江苏 徐州 221116; 2. 矿山数字化教育部工程研究中心, 江苏 徐州 221116)

摘要: 为了得到更好的图像评价指标, 均方误差损失是大多数现有的与深度学习方法结合的图像超分辨率技术都在使用的目标优化函数, 但大多数算法构建出来的图像因严重丢失高频信息和模糊的纹理边缘而不能达到视觉感受的需求。针对上述问题, 本文提出融合感知损失的广泛激活的非常深的残差网络的超分辨率模型, 通过引入感知损失、对抗损失, 并结合平均绝对误差组成新的损失函数, 通过调整不同损失项的权重对损失函数进行优化, 提高了对低分辨率图像的特征重构能力, 高度还原图像缺失的高频信息。本文选取峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似度 (structural similarity, SSIM) 两个国际公认的评判指标作为客观评判标准, 更换数据集进行实验分析、结果对比, 在主观视觉上直观观察效果, 结果从不同角度证明本文方法性能较对比模型有所提升, 证明了引入感知损失后, 模型更好地构建了低分辨率图的纹理细节, 可以获得更好的视觉体验。

关键词: 深度学习; 超分辨率; 广泛激活; 感知损失; 特征重构; 峰值信噪比; 结构相似度; 视觉体验

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1673-4785(2022)02-0440-07

中文引用格式: 王凡超, 丁世飞. 基于广泛激活深度残差网络的图像超分辨率重建 [J]. 智能系统学报, 2022, 17(2): 440-446.

英文引用格式: WANG Fanchao, DING Shifei. Image super-resolution reconstruction based on widely activated deep residual networks[J]. CAAI transactions on intelligent systems, 2022, 17(2): 440-446.

Image super-resolution reconstruction based on widely activated deep residual networks

WANG Fanchao¹, DING Shifei^{1,2}

(1. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou 221116, China; 2. Mine Digitization Engineering Research Center of Ministry of Education of the People's Republic of China, Xuzhou 221116, China)

Abstract: To obtain good image evaluation indexes, the mean squared error loss is used as an objective optimization function in image super-resolution technologies combined with the deep learning method. However, most constructed images cannot meet the visual experience requirement due to the serious loss of high-frequency signals and fuzzy texture edges. In response to the above problems, in this paper, we propose a super-resolution model for a widely activated deep residual network combining perceptual loss. A new loss function is formed by introducing perceptual and adversarial losses and is optimized by adjusting the weight of different losses. The loss function is optimized to improve the feature reconstruction ability of low-resolution images and highly restore the high-frequency information missing from the images. Two internationally recognized evaluation indicators, namely, peak signal-to-noise ratio and structural similarity, are selected as objective evaluation criteria. A comparative analysis is performed on different datasets, and the images produced are subjected to direct and subjective observations. The results show that the performance of the proposed method is improved in different aspects in comparison with the compared models. Hence, after the introduction of perceptual loss, the model can effectively reconstruct the texture details of low-resolution images and offer an outstanding visual experience.

Keywords: deep learning; super-resolution; extensive activation; perceptual loss; feature reconstruction; peak signal-to-noise ratio; structural similarity; visual experience

收稿日期: 2021-06-15. 网络出版日期: 2021-10-13.
基金项目: 国家自然科学基金项目 (61976216, 61672522).
通信作者: 丁世飞. E-mail: dingsf@cumt.edu.cn.

如今, 图像处理技术和信息交互快速发展, 图像作为极其重要的信息载体, 在公共安防、医学

诊疗、卫星遥感等应用领域中愈发重要,但环境和噪声等因素的干扰,导致图像出现质量低或者细节的纹理信息缺乏等问题^[1]。图像超分辨率(super resolution, SR),特别单图像超分辨率(single image super-resolution, SISR),几十年来受到广泛关注。单图像超分辨率为了通过技术手段将低分辨率(low-resolution, LR)图像重新构建为高分辨率(high-resolution, HR)图像,使其具有良好的高频、纹理和边缘信息。经过几十年发展,图像超分辨率逐渐被分为基于插值^[2-3]、基于重建^[4]和基于学习^[5]三大类。

基于插值和重建超分辨率算法属于传统算法,都运用经典的数学模型算法,但随着放大因子的增大,这两种算法提供的用于重新构建高分辨率图像的细节信息资源不足的弊端逐渐显露,使得这些传统算法很难达到重建高频信息的目的。

深度学习高速发展以来,图像重建领域也开始使用深度学习模型来完成特定任务。2014年,Dong等^[6]最先将卷积神经网络应用到图像超分辨率任务上,提出了超分辨率卷积神经网络(super-resolution convolutional neural network, SRCNN)模型,该模型进行端到端的图像对训练,在当时大大提高了重建效果,同时也开辟了超分辨率技术的新时代。2016年,Kim等^[7]将超分辨率任务与残差思想进行结合,提出了使用非常深的卷积网络超分辨率(super-resolution using very deep convolution network, VDSR)模型,同时将网络的深度增加到了20层,证明了深层网络能提取出更多的特征,取得更好的重建效果。Lim等^[8]通过对残差结构的改进,提出增强型深度超分辨率网络(enhanced deep super-resolution network, EDSR)模型,通过移除批规范化处理^[9](batch normalization, BN)操作,增加层数,提取更多特征,从而获得更满意的结果。

针对上述模型中存在的对不同层的图像信息使用率不足等问题,本文提出了融合感知损失的广泛激活的深度残差网络的超分辨率模型(widely-activated deep residual network for super-resolution combining perceptual loss, PWDSR),通过已训练的VGG模型^[10]提取激活前的特征,使用激活前的特征会克服两个缺点:第一,深层网络中,激活后得到的稀疏特征能够提供的监督效果非常弱,影响性能;第二,激活后的特征会使得重建后的图像与真实图像在亮度上有所差异。融合感知损失更着眼于纹理而不是目标物体。

本文的主要贡献包括以下方面:

- 1) 使用权重归一化代替批量归一化,提高了学习率和训练、测试准确率;
- 2) 使用预训练的深度模型提取激活前的特征得到感知损失,通过构建的图像与真实图像计算对抗损失,并结合图像的像素损失,构建了新损失函数;
- 3) 使用全局跳跃连接避免梯度消失的问题,促进梯度的反向传播,从而加快训练过程;
- 4) 实验证明了本文提出的损失函数改进在不同数据集上可以取得更好的评价指标,在主观视觉效果也有所提高;
- 5) 本文提供了高倍重建任务中优化损失函数、调整残差块等新的研究方向。

1 相关工作

1.1 超分辨率网络

从超分辨率卷积神经网络^[6](SRCNN)首次提出端到端卷积神经网络作为新的映射方式开始,超分辨率与卷积神经网络模型越来越密切。

1.1.1 上采样

超分辨率卷积神经网络^[6](SRCNN)对图像进行上采样,在高分辨率特征空间上计算卷积导致其效率是低分辨率空间的 S^2 倍(S 为放大因子),因此效率低下。另一个有效的替代方案是子像素卷积^[11](sub-pixel convolution),该卷积比反卷积层引入更少的伪像误差。

1.1.2 标准化

从3层的超分辨率卷积神经网络^[6](SRCNN)到160层的多尺度超分辨率网络^[8](multi-scale deep super-resolution, MDSR),图像超分辨率网络越来越深入,训练变得更加困难。批量归一化^[9](batch normalization, BN)在许多任务中能解决训练困难的问题,例如SRResNet^[12](super-resolution residual network)中使用了批量归一化。

BN通过再次校准中间特征的均值和方差来解决训练深度神经网络时内部协变量偏移的问题。简单来说,如果忽略BN中可学习参数的重新缩放,那么训练期间可用训练小批量的均值和方差对每一层的特征进行归一化:

$$\hat{x} = \frac{x_B - E_B[x_B]}{\sqrt{\text{Var}_B[x_B] + \varepsilon}} \quad (1)$$

式中: x_B 是当前训练批次的特征,它是一个较小的值用来避免零除。然后将一阶和二阶统计信息更新为全局统计信息:

$$E[x] \leftarrow E_B[x_B] \quad (2)$$

$$\text{Var}[x] \leftarrow \text{Var}_B[x_B] \tag{3}$$

其中 \leftarrow 表示分配移动线。这些全局统计信息将用于标准化:

$$\hat{x}_{\text{test}} = \frac{x_{\text{test}} - E[x]}{\sqrt{\text{Var}[x] + \varepsilon}} \tag{4}$$

但 BN 用于超分辨率任务时存在以下问题:

1) 图像 SR 通常仅使用较小的图像块 (48×48) 来加快训练速度, 小图像块的均值和方差相差很大, 影响统计数据; 2) 图像 SR 网络中没有使用正则化器会过度拟合训练数据集; 3) 图像 SR 网络的训练和测试时使用不同公式会降低密集像素值预测的准确性。

1.1.3 跳跃连接

跳跃连接在深层神经网络中具有优良表现, 其可以兼顾低级特征和高级特征。非常深的卷积网络超分辨率^[7](super-resolution using very deep convolution network, VDSR) 模型使用全局跳跃连接。残差密集网络^[13](RDN) 使用所有卷积层的分层特征。

1.1.4 分组卷积和深度可分离卷积

分组卷积将特征按通道划分为多个组, 并分别在组内执行卷积, 然后进行串联以形成最终输出。在组卷积中, 参数的数量可以减少 g 倍, 其中 g 是组数。

深度可分离卷积是深度非线性卷积 (即在输入每个通道上独立执行的空间卷积), 然后是点卷积 (即 1×1 卷积)。也可以将其视为特定类型的组卷积, 其中组数 g 是通道数。

1.2 广泛激活的深度残差网络

广泛激活的深度残差网络^[14](widely-activated deep residual network for super-resolution, WDSR) 在 ReLU 激活层之前扩展特征, 同时追求不会额外增加运算量。该模型压缩残差等价映射路径的特征, 同时扩展激活前的特征, 如图 1 所示, 并与增强型深度超分辨率网络^[8](EDSR) 中的基础残差块 (图 1(a)) 进行对比。WDSR-A 中的 (图 1(b)) 具有细长的映射路径, 在每个残差块中激活之前都具有较宽的 (2~4 倍) 通道。WDSR-B (图 1(c)) 具有线性低秩卷积堆栈, 同时在不额外增加运算量的情况下加宽激活范围 (6~9 倍)。首先使用 1×1 卷积核增加通道数, 在 ReLU 激活层之后使用有效的线性低秩卷积, 用两个低秩的卷积核替换一个大的卷积核 (1×1 卷积核降低通道数, 3×3 卷积核执行空间特征提取)。在 WDSR-A 和 WDSR-B 中, 所有 ReLU 激活层仅应用于两个较宽的功能部件 (具有较大通道的功能部件) 之间。

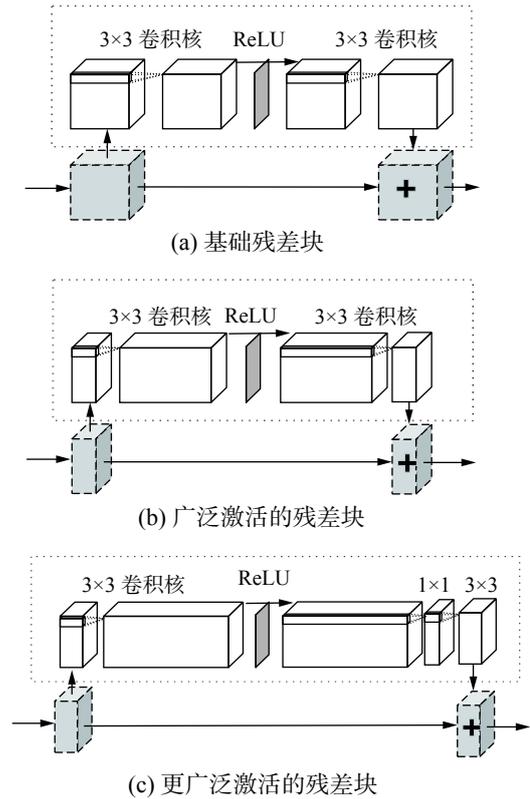


图 1 广泛激活的残差块与基础残差块的对比
Fig. 1 Comparison of the residual block with wide activation and the original residual block

2 本文模型

受增强型超分辨率生成对抗网络^[15](enhanced super-resolution generative adversarial networks, ESRGAN) 中使用感知域损失的启发, 本文针对广泛激活的深度残差网络存在对不同层级图像信息使用不足的问题, 引入感知域损失, 通过调整损失权重进行优化。本文网络模型如图 2 所示。

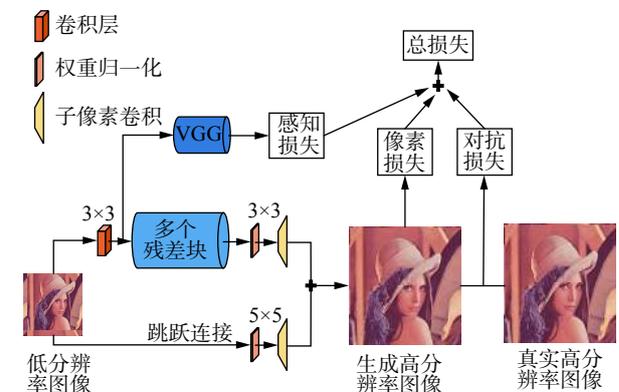


图 2 融合感知损失的广泛激活的深度残差网络的超分辨率模型
Fig. 2 Widely-activated deep residual network for super-resolution combining perceptual loss

本文调用已训练的 VGG19 模型作为特征提取器。VGG 网络模型^[10]如图 3 所示。

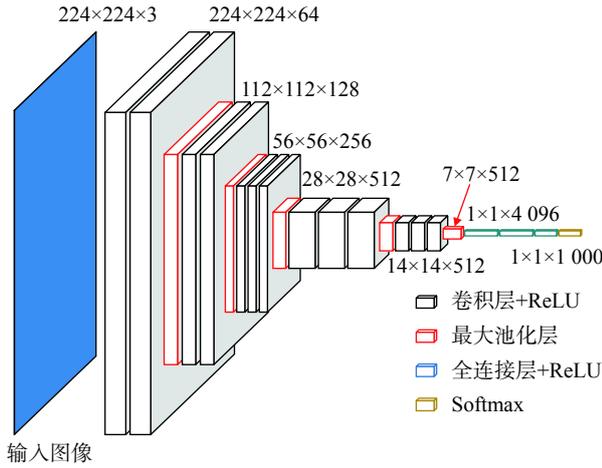


图 3 VGG 结构模型

Fig. 3 VGG structure model

2.1 权重归一化

由于批量归一化干扰图像超分辨率的准确性, 故本文使用权重归一化 (weigh normalization, WN) 来代替批量归一化。权重归一化是神经网络中权重向量的重新参数化, 将这些权重向量的长度与其方向解耦, 使其不会在小批量中引入示例之间的依赖关系, 并且在训练和测试中具有相同的表示形式。输出 y 的形式为

$$y = w \cdot x + b \quad (5)$$

式中: w 是 k 维权重向量; b 是标量偏差项; x 是输入的 k 维向量。WN 使用以下参数重新参数化权重向量:

$$w = \frac{g}{\|v\|} v \quad (6)$$

其中, v 是一个 k 维矢量, g 是一个标量, $\|v\|$ 表示 v 的欧几里得范数。通过这种形式化, 得到 $\|w\| = g$, 而与参数 v 无关。对于图像超分辨率, WN 只是一种重新参数化技术, 并且具有完全相同的表示能力, 因此不会影响准确性。同时, WN 可以提高学习率, 并提高训练和测试准确性。

2.2 损失函数改进

模型重新构建图像的结果很大程度上取决于损失函数的选择, 一般超分辨率模型选择平均绝对误差 (mean absolute error, MAE) 或者均方误差 (mean-square error, MSE) 作为优化的目标, 原因是在测试时可以获得较高的评价指标, 但是在进行 8 倍等大尺度的超分辨率重建任务中, 重新构建的图像缺失大量高频信息, 导致构建的图像不能达到视觉要求。本文模型使用感知损失函数 l_p , 像素损失函数 l_1 和对抗损失函数 l_g , 通过配置 3 种损失函数不同权重使其更好地对高频信息进行重新构建, 总损失函数可表示为

$$l = \lambda_1 l_p + \lambda_2 l_1 + \lambda_3 l_g \quad (7)$$

其中, λ_i 代表调节各个损失项权重的正则因子, $i =$

1, 2, 3。

为了在重建时保证重建图像与对应低分辨率图像在低频部分保持图像结构的一致性, 像素损失使用平均绝对误差, 公式为

$$l_1 = \frac{1}{n} \sum |S - H| \quad (8)$$

式中: S 代表生成图像; H 代表真实高分辨率图像; n 代表网络超参数 batch 的大小。

根据生成对抗网络的思想, 在网络重建出一幅高分辨率图像之后, 与其对应真实高分辨率图像进行比较计算, 假设 N 个 batch, 生成 n 个标签, 公式为

$$l_g = \{r_1, r_2, \dots, r_N\} \quad (9)$$

$$r_n = -[y_n \cdot \log(\sigma(x_n)) + (1 - y_n) \cdot \log(1 - \sigma(x_n))] \quad (10)$$

其中, $\sigma(x_n)$ 为 sigmoid 函数, 可以把 x 映射到 (0, 1) 之间:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (11)$$

本文使用的特征是预训练的深度网络激活层前的特征。公式为

$$l_p = \sum_{i=1}^N \frac{1}{C_j H_j W_j} [\|\phi_j(p) - \phi_j(q)\|_1] \quad (12)$$

式中: p 、 q 分别代表真实高分辨率图像和生成图像; ϕ 表示预训练的神经网络; j 表示该网络的第 j 层; $C_j \times H_j \times W_j$ 为第 j 层特征图的形状, 使用的是每个卷积模块的激活值。

为确定损失函数中不同分量的权重值, 本文进行了参数实验, 根据收敛情况确定权重值, 图 4~6 分别为感知损失、像素损失、对抗损失随权重的变化曲线, 损失函数收敛到最小值时对应值即为正则因子 λ_i 权重值, $i = 1, 2, 3$ 。

PWDSR 算法描述如下:

输入 低分辨率图像, 真实高分辨率图像;

输出 生成的高分辨率图像。

1) 低分辨率图像 \rightarrow GPU \Rightarrow 低分辨率图像对应张量, 真实高分辨率图像 \rightarrow GPU \Rightarrow 高分辨率图像对应张量;

2) 低分辨率图像对应张量 \rightarrow PWDSR \Rightarrow 超分辨率生成图像对应张量;

3) 不同层 feature map \rightarrow VGG $\Rightarrow l_p$;

4) [高分辨率图像对应张量 - 低分辨率图像对应张量] $\Rightarrow l_1, l_g$;

5) 通过公式 $l = \lambda_1 l_p + \lambda_2 l_1 + \lambda_3 l_g$ 计算损失;

7) 是否达到 epoch 最大值, 是则输出超分辨率图像对应张量; 否则继续训练;

6) PWDSR $\leftarrow l$;

8) 超分辨率图像对应张量 \rightarrow CPU \Rightarrow 生成的高分辨率图像。

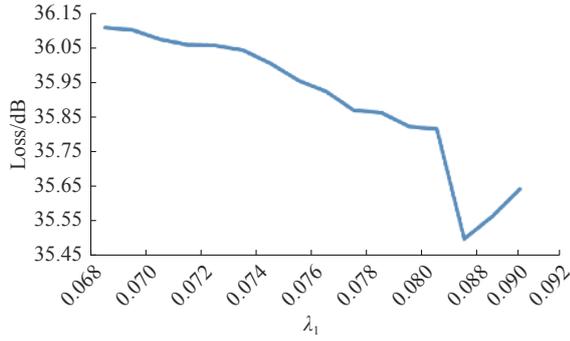


图 4 感知损失收敛曲线
Fig. 4 Perceptual loss convergence curve

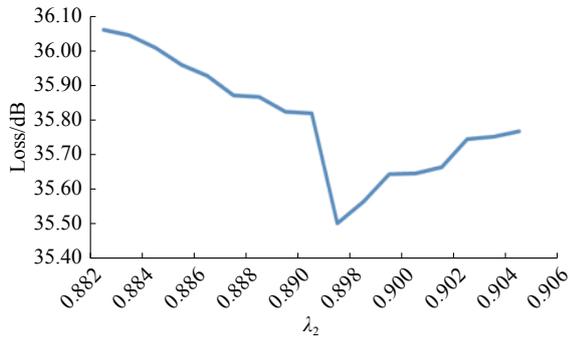


图 5 像素损失收敛曲线
Fig. 5 Pixel loss convergence curve

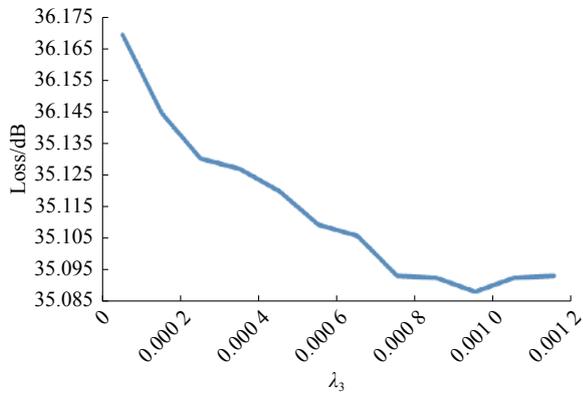


图 6 对抗损失收敛曲线
Fig. 6 Adversarial loss convergence curve

3 实验设置

本实验在 Windows 10 操作系统中使用 PyTorch 深度学习框架, Python 版本为 3.7.6, 硬件配置为 Intel(R) Core(TM) i5-7300HQ CPU @ 2.50 GHz, 24 GB 内存, 显卡为 NIDIA GeForce GTX 1050 Ti, 使用 CUDA 10.2 + cudnn 7.4.1 进行 GPU 加速, 在训练过程中, 采用 Adam 梯度下降算法, 设置初始学习率为 0.001, 模型的 epoch 设置为 200, 每 10 次保存一次网络模型。

3.1 数据集

DIV2K 数据集是用于 NTIRE 大赛的标准数

据集, 该数据集包含 1000 张 2K 分辨率的高清图像, 其中 800 张作为训练数据集, 100 张用于验证, 100 张用于测试。该数据集还包含高清分辨率图像对应的低分辨率图像 (使用插值法获得) 便于训练。本文中, 将 DIV2K 数据集中编号 1~800 的图像作为训练集, 编号 801~900 的图像作为验证集, 编号 901~1000 的图像作为测试集, 另选 Set5、Set14 为测试集进行对比。Set5 为 5 张动植物的图像, Set14 数据集包含 14 张自然景象的图像。

3.2 评价指标

图像超分辨率效果的客观评价指标为峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似性 (structural similarity, SSIM)。

3.2.1 峰值信噪比

峰值信噪比是图像超分辨率评价指标中使用最多的一种标准, 其使用均方误差来对图片质量作判断。对于单色 $m \times n$ 的高清原图 I 与超分辨率得到的图片 K , 两者之间的均方误差公式为

$$MSE = \frac{1}{mn} \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} [I(x,y) - K(x,y)]^2 \quad (13)$$

峰值信噪比的公式为

$$PSNR = 10 \cdot \lg \left(\frac{MAX_I^2}{MSE} \right) = 20 \cdot \lg \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (14)$$

式中: MAX_I 表示像素最大值, 如果采样点用 8 位表示, 则为 255。可见, MSE 与 PSNR 成反比, PSNR 越大代表重新构建的图像效果越好。

3.2.2 结构相似度

结构相似度是图像超分辨率重建的另一个应用较为广泛的测量指标, 其输入是两张图像, 其中一张是未经压缩的无失真图像 y , 另一张是重新构建出的图像 x , 那么 SSIM 公式为

$$SSIM(x,y) = [l(x,y)]^\alpha \cdot [c(x,y)]^\beta \cdot [s(x,y)]^\gamma \quad (15)$$

式中: $\alpha > 0, \beta > 0, \gamma > 0$, l 是亮度 (luminance) 比较, c 是对比度 (contrast) 比较, s 是结构 (structure) 比较:

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1}$$

$$c(x,y) = \frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2}$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3}$$

一般 $c_3 = c_2/2$, 其中 μ_x, μ_y 表示均值。 σ_x^2, σ_y^2 表示方差, σ_{xy} 表示 x 与 y 的协方差。在实际应用中通常设 $\alpha = \beta = \gamma = 1$, 故可将式 (15) 简化为

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

可以看出, SSIM 具有对称性, 即 $SSIM(x,y) =$

SSIM(y, x), SSIM ∈ [0, 1], SSIM 与输出图像和无失真图像的差距成反比, SSIM 越大图像质量越好。当两幅图像一模一样时, SSIM=1。

3.3 实验结果及分析

本实验从客观评价结果和主观评价结果两方面来表现改进模型的超分辨率能力, 通过不同方

式证明了本文提出的模型改进能优于其他模型。

3.3.1 客观评价结果

本文在 3 个公开数据集上测试了 WDSR-A、WDSR-B 以及本文模型, 分别计算在不同数据集上采用不同算法进行上采样 2、3、4 倍时的 PSNR 和 SSIM, 对比结果如表 1 所示。

表 1 在不同数据集上对比放大倍数为 2、3、4 的重构图像的 PSNR 和 SSIM

Table 1 Comparison of PSNR and SSIM of 2, 3, and 4 times reconstructed images on different datasets

数据集	放大倍数	WDSR-A			WDSR-B			PWDSR		
		PSNR/dB	SSIM	t/min	PSNR/dB	SSIM	t/min	PSNR/dB	SSIM	t/min
DIV2K	2	33.8613	0.8040	1054.0	34.8217	0.8357	1055.0	35.0280	0.8521	1056.5
	3	31.0211	0.7084	1332.8	31.0644	0.7145	1335.0	31.1129	0.7202	1337.5
	4	28.7837	0.6321	1695.5	28.9369	0.6367	1697.5	29.0079	0.6524	1702.5
Set5	2	35.0688	0.9002	89.4	35.1157	0.9020	89.5	35.3112	0.9084	89.6
	3	31.0657	0.7165	107.9	31.3806	0.7345	108.0	31.7262	0.7465	109.1
	4	29.3180	0.6730	144.2	29.3658	0.6851	144.7	29.5172	0.6964	145.6
Set14	2	31.0615	0.7131	262.5	31.1140	0.7252	264.6	31.1448	0.7308	266.0
	3	27.8946	0.6180	328.1	28.0032	0.6225	329.7	28.9950	0.6434	331.9
	4	26.2087	0.6043	420.8	26.3112	0.6087	422.1	27.2699	0.6176	424.4

注: 粗体字代表最好结果

从实验结果不难发现, 更换不同数据集, 本文提出的方法在 2、3、4 倍重建任务中, 都能够取得较好的 PSNR 和 SSIM 值, 相较于其他模型, 在客观指标上有所提升。

3.3.2 主观评价结果

本文分别选取了 DIV2K、Set5、Set14 数据集的 3 张高分辨率图像进行放大倍数为 2、3、4 的重构对比, 为了更好地体现对比结果, 本文将选取不同图片的不同细节进行放大对比: 图 7 选取 DIV2K 数据集中图像绿叶的右端枝叶部分进行 4 倍重建对比, 图 8 选取 Set5 数据集中图像婴儿的左眼及上方部分进行 3 倍重建对比, 图 9 选取 Set14 数据集中图像女孩的左眼及下方部分进行 4 倍重建对比。



图 7 编号 0803(DIV2K) 4 倍重建视觉比较

Fig. 7 Visual comparison of SR results of “0803” (DIV2K) with scale factor 4



图 8 baby(Set5) 3 倍重建视觉比较

Fig. 8 Visual comparison of SR results of “baby” (Set5) with scale factor 3



图 9 comic(Set14) 4 倍重建视觉比较

Fig. 9 Visual comparison of SR results of “comic” (Set14) with scale factor 4

从视觉对比结果可以看出, 3 种模型在 2、3、4 倍的重建任务中都可以完成重建高分辨率图像, 但本文提出模型在不同倍数重建任务中, 能够重建出更好的纹理细节, 达到更好的视觉效果, 说明重新构建出的图像更接近原始高分辨率图像。

4 结束语

本文在广泛激活的深度残差网络的基础上,融合感知损失、对抗损失、像素损失,对整体损失函数进行优化,使用已训练好的 VGG19 模型提取激活前的特征得到感知损失,避免了使用激活后的稀疏特征导致的性能不良等问题。本文使用权重归一化代替批量归一化,提高了学习率和训练、测试准确率。使用全局跳跃连接,避免梯度消失的问题,同时有助于梯度的反向传播,加快训练过程。从实验结果可以看出,本文提出的损失函数改进在不同数据集上可以取得更好的评价指标,在主观视觉效果也有所提高。在高倍重建任务上还有提升空间,后续工作以调整残差块和优化损失函数等方向进行展开。

参考文献:

- [1] 雷鹏程, 刘丛, 唐坚刚, 等. 分层特征融合注意力网络图像超分辨率重建 [J]. 中国图象图形学报, 2020, 25(9): 1773–1786.
LEI Pengcheng, LIU Cong, TANG Jiangang, et al. Hierarchical feature fusion attention network for image super-resolution reconstruction[J]. Journal of image and graphics, 2020, 25(9): 1773–1786.
- [2] HUANG Zhengzhong, CAO Liangcai. Bicubic interpolation and extrapolation iteration method for high resolution digital holographic reconstruction[J]. Optics and lasers in engineering, 2020, 130: 106090.
- [3] HAO Sai, DONG Xianghui. Interpolation-based plane stress anisotropic yield models[J]. International journal of mechanical sciences, 2020, 178: 105612.
- [4] 贺璟, 郝晓丽, 吕进来. 梯度插值与可变阈值改进的 POCS 算法 [J]. 中国科技论文, 2017, 12(14): 1655–1658, 1684.
HE Jing, HAO Xiaoli, LÜ Jinlai. POCS algorithm based on gradient interpolation and variable threshold[J]. China sciencepaper, 2017, 12(14): 1655–1658, 1684.
- [5] 白蔚, 杨撒博雅, 刘家瑛, 等. 基于显著性稀疏表示的图像超分辨率算法 [J]. 中国科技论文, 2014, 9(1): 103–107.
BAI Wei, YANG Saboya, LIU Jiaying, et al. Image super resolution based on salient sparse coding[J]. China sciencepaper, 2014, 9(1): 103–107.
- [6] DONG Chao, LOY C C, HE Kaiming, et al. Learning a deep convolutional network for image super-resolution [C]//Proceedings of the 13th European Conference on Computer Vision. Zurich, Switzerland, 2014: 184–199.
- [7] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 1646–1654.
- [8] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, USA, 2017: 136–144.
- [9] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on Machine Learning. Lille, France, 2015: 448–456.
- [10] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10)[2021-01-01]. <https://arxiv.org/abs/1409.1556>.
- [11] SHI Wenzhe, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 1874–1883.
- [12] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 4681–4690.
- [13] ZHANG Yulun, TIAN Yapeng, KONG Yu, et al. Residual dense network for image super-resolution[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 2472–2481.
- [14] YU Jiahui, FAN Yuchen, YANG Jianchao, et al. Wide activation for efficient and accurate image super-resolution[EB/OL]. (2018-12-21)[2021-01-01]. <https://arxiv.org/abs/1808.08718v1>.
- [15] WANG Xintao, YU Ke, WU Shixiang, et al. ESRGAN: enhanced super-resolution generative adversarial networks[C]//Proceedings of the European Conference on Computer Vision. Munich, Germany, 2018: 63–79.

作者简介:



王凡超, 硕士研究生, 主要研究方向为深度学习、图像超分辨率。



丁世飞, 教授, 博士生导师, CCF 杰出会员, 第八届吴文俊人工智能科学技术奖获得者, 主要研究方向为人工智能与模式识别、机器学习与数据挖掘。主持国家重点基础研究计划课题 1 项、国家自然科学基金面上项目 3 项。出版专著 5 部, 发表学术论文 200 余篇。