

控制与决策

Control and Decision

基于自注意力生成对抗网络的图像超分辨率重建

王雪松, 晁杰, 程玉虎

引用本文:

王雪松, 晁杰, 程玉虎. 基于自注意力生成对抗网络的图像超分辨率重建[J]. *控制与决策*, 2021, 36(6): 1324–1332.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.1290>

您可能感兴趣的其他文章

Articles you may be interested in

基于条件对抗生成孪生网络的目标跟踪

Conditional generative adversarial siamese networks for object tracking

控制与决策. 2021, 36(5): 1110–1118 <https://doi.org/10.13195/j.kzyjc.2019.1215>

基于生成对抗网络学习被遮挡特征的目标检测方法

Object detection via learning occluded features based on generative adversarial networks

控制与决策. 2021, 36(5): 1199–1205 <https://doi.org/10.13195/j.kzyjc.2019.1319>

一种基于深度学习的时间序列预测方法

A time series prediction method based on deep learning

控制与决策. 2021, 36(3): 645–652 <https://doi.org/10.13195/j.kzyjc.2019.0809>

Anchor-free的尺度自适应行人检测算法

Anchor-free scale adaptive pedestrian detection algorithm

控制与决策. 2021, 36(2): 295–302 <https://doi.org/10.13195/j.kzyjc.2020.0124>

结合注意力机制的循环神经网络复述识别模型

Recurrent neural networks based paraphrase identification model combined with attention mechanism

控制与决策. 2021, 36(1): 152–158 <https://doi.org/10.13195/j.kzyjc.2019.0638>

基于自注意力生成对抗网络的图像超分辨率重建

王雪松, 晁杰, 程玉虎[†]

(1. 中国矿业大学 地下空间智能控制教育部工程研究中心, 江苏 徐州 221116;
2. 中国矿业大学 信息与控制工程学院, 江苏 徐州 221116)

摘要: 针对如何恢复重建后超分辨率图像的纹理细节问题, 提出基于自注意力生成对抗网络的图像超分辨率重建模型 (SRAGAN). 在 SRAGAN 中, 基于自注意力机制和残差模块相结合的生成器用于将低分辨率图像变换为超分辨率图像, 基于深度卷积网络构成的判别器试图区分重建后的超分辨率图像和真实超分辨率图像间的差异. 在损失函数构造方面, 一方面利用 Charbonnier 内容损失函数来提高图像的重建精度, 另一方面使用预训练 VGG 网络激活前的特征值来计算感知损失以实现超分辨率图像的精确纹理细节重构. 实验结果表明, SRAGAN 在峰值信噪比和结构相似度分数上均优于当前流行算法, 能够重构出更为真实和具有清晰纹理的图像.

关键词: 图像超分辨率重建; 自注意力机制; 生成对抗网络; 损失函数

中图分类号: TP273 文献标志码: A

DOI: 10.13195/j.kzyjc.2019.1290

引用格式: 王雪松, 晁杰, 程玉虎. 基于自注意力生成对抗网络的图像超分辨率重建 [J]. 控制与决策, 2021, 36(6): 1324-1332.

Image super-resolution reconstruction based on self-attention GAN

WANG Xue-song, CHAO Jie, CHENG Yu-hu[†]

(1. Engineering Research Center of Intelligent Control for Underground Space, Ministry of Education, China University of Mining and Technology, Xuzhou 221116, China; 2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China)

Abstract: Aiming at how to recover the texture details of the reconstructed super-resolution image, an image super-resolution reconstruction based on the self-attention generative adversarial network (SRAGAN) is proposed. In the SRAGAN, a generator based on a combination of the self-attention mechanism and the residual module is used to transform low-resolution into super-resolution images, while a discriminator based on the deep convolutional network tries to distinguish the difference between the reconstructed and real super-resolution images. In terms of loss function construction, on the one hand, the Charbonnier content loss function is used to improve the accuracy of image reconstruction; on the other hand, the eigenvalues before the activation layer in the pre-trained VGG network are used to calculate the perceptual loss to achieve accurate texture detail reconstruction of super-resolution images. Experiments show that the proposed SRAGAN is superior to the current popular algorithms in peak signal-to-noise ratio and structural similarity score, reconstructing more realistic images with clear textures.

Keywords: image super-resolution reconstruction; self-attention mechanism; generative adversarial network; loss function

0 引言

图像的超分辨率重建问题是指从一个低分辨率图像估计出超分辨率图像, 它是一个极具挑战性的任务. 在计算机视觉领域, 图像超分辨率重建一直受到极大的关注, 并且具有广泛的应用, 如医学成像^[1]、卫星成像^[2]、人脸识别^[3]等. 超分辨率重建算法可以在不升级成像系统硬件设备的前提下, 从软件算法层面来满足实际应用中图像分辨率增长的需求, 具有重要

的研究意义.

随着图像像素放大倍数的增加, 图像超分辨率重建问题的不确定性变得尤为明显. 当前图像超分辨率重建算法主要是基于学习的方法, 包括邻域嵌入^[4]、稀疏编码^[5]和随机森林^[6]等. 由于深度卷积神经网络 (CNN)^[7]的强大学习能力, Dong 等^[8]提出基于卷积神经网络的图像超分辨率重建算法 (SRCNN), 将卷积神经网络应用于图像超分辨率重建问题, 实

收稿日期: 2019-09-10; 修回日期: 2020-03-12.

基金项目: 国家自然科学基金项目 (61772532, 61976215).

[†]通讯作者. E-mail: chengyuhu@163.com.

现了低分辨率图像和超分辨率图像之间的端到端变换. 由于SRCNN的成功, 以至于后来多种类型的深度卷积网络模型被陆续应用于图像超分辨率重建问题中. 受VGG网络的启发, Kim等^[9]提出了基于超深卷积网络的图像超分辨率重建模型(VDSR). 为加快模型训练速度, Kim等^[10]和Tai等^[11]分别通过在VDSR的基础上增加递归连接来实现权重共享和模型压缩, 提出深度递归卷积网络模型(DRCN). Shi等^[12]通过采用有效的子像素卷积替换双三次上采样操作, 提出高效子像素卷积网络模型(ESPCN), 从而能够很好地实现超分辨率图像像素的重建. Chao等^[13]通过采用一种相较于ESPCN的层数更多且参数更少的沙漏形CNN, 提出基于高效卷积网络的图像超分辨率重建模型(FSRCNN). 拉普拉斯金字塔网络(LapSRN)^[14]是用于解决图像超分辨率重建问题的最新框架之一, 该模型使用拉普拉斯金字塔的特征提取和图像重建部分的级联框架, 可以很好地实现超分辨率图像纹理细节的重建. Lim等^[15]通过去除残差块中不必要的批量归一化层并扩展模型尺寸, 提出基于深度增强残差网络的图像超分辨率重建模型(EDSR).

解决超分辨率问题的另一种深度网络模型, 是Ledig等^[16]提出的基于生成对抗网络的图像超分辨率重建模型(SRGAN), 通过添加对抗损失来生成更逼真的图像. 生成对抗网络(GAN)^[17]自提出以来, 已被广泛应用于各种图像处理领域, 包括图像生成^[18]、图像风格转换^[19]、图像修复^[20]以及图像超分辨率重建^[16]. 在GAN的训练中, 生成器生成样本以欺骗判别器, 同时判别器又试图区分来自生成器的样本和真实样本的差别. GAN成功的关键在于引入了对抗的思想, 通过生成器和判别器的相互博弈, 使得判别器最终无法区分由生成器生成的样本和真实的样本. 在解决图像超分辨率重建问题上, SRGAN将判别损失加入到传统的感知损失中, 利用判别器区分重建后的超分辨率图像和真实超分辨率图像的差异, 可以很好地生成更接近超分辨率图像的纹理细节, 使图像更加的真实. 然而, SRGAN在一些具体纹理细节重构方面仍然不够清晰, 而且会伴随模糊的伪影, 同时基于VGG分类网络的感知损失无法捕获图像超分辨率重建任务中所需的高频细节.

针对上述问题, 提出一种基于自注意力生成对抗网络的图像超分辨率重建算法(image super-resolution reconstruction based on self-attention GAN, SRAGAN), 通过设计带有自注意力机制残差模块的深层生成器结构, 利用自注意力层提取的全局特征信息进行超分辨率图像的重建. 生成器采用不同

尺寸的卷积核进行输入图像的特征提取和超分辨率图像的特征重构, 不同大小的卷积核可以多方面地捕获图像的特征信息. 生成器中采用实例归一化层替代传统的批量归一化层, 对单个图像实例进行归一化处理, 更有助于解决图像重建任务. 在损失函数方面, 采用Charbonnier损失作为内容损失函数来评估所生成的图像与真实图像之间的相似性, 提供更强大的监督能力, 进而提高图像的重建精度. 同时, SRAGAN使用预训练好的VGG-19网络激活前的特征值来计算感知损失, 以实现超分辨率图像的精确纹理细节重构. 最后, 训练过程中使用Wasserstein距离来优化GAN网络的对抗训练, 以保证模型训练的收敛性.

1 基于自注意力生成对抗网络的图像超分辨率重建

1.1 超分辨率图像重建模型

给定低分辨率图像 I_{LR} , 图像超分辨率重建问题的任务就是根据 I_{LR} 估计出超分辨率图像 \hat{I}_{HR} , 同时使生成的 \hat{I}_{HR} 尽可能地接近原来真实的超分辨率图像 I_{HR} . 为了解决这个任务, 需要训练一个生成器网络 G , 网络的参数用 θ 表示, G_θ 的输入和输出分别是低分辨率图像和重建后的超分辨率图像, 则

$$\hat{I}_{HR} = G_\theta(I_{LR}). \quad (1)$$

通过优化训练样本上的 G_θ , 使参数 θ 满足

$$\theta = \operatorname{argmin} \sum L(\hat{I}_{HR}, I_{HR}). \quad (2)$$

其中: $L(\hat{I}_{HR}, I_{HR})$ 为重构误差, (I_{LR}, I_{HR}) 表示配对的训练集.

为解决图像超分辨率重建问题, 需要训练出生成器网络 G 和判别器网络 D . 系统总框架由生成器和判别器组成, 生成器试图将低分辨率图像变换为真实的超分辨率图像, 判别器试图区分重建后的超分辨率图像和真实超分辨率图像的差异, 总体模型框图如图1所示, 具体生成器和判别器的网络参数结构会在1.3节中介绍.

在这项工作中, 最终目标是训练一个生成器网络, 使其生成与真实超分辨率图像尽可能相似的超分辨率图像. 为了实现这一点, 引入自注意力层, 设计自注意力机制残差模块, 它可以更好地利用全局特征信息进行超分辨率图像的重建. 在深度网络结构基础上, SRAGAN在生成器中将传统的批量归一化层替换为实例归一化层. 另外, 通过优化内容损失, 进一步评估所生成的图像与真实图像之间的相似性, 同时将感知损失进行优化, 使用VGG-19网络激活前的特征值来计算感知损失, 后面的章节会分别介绍损失函数.

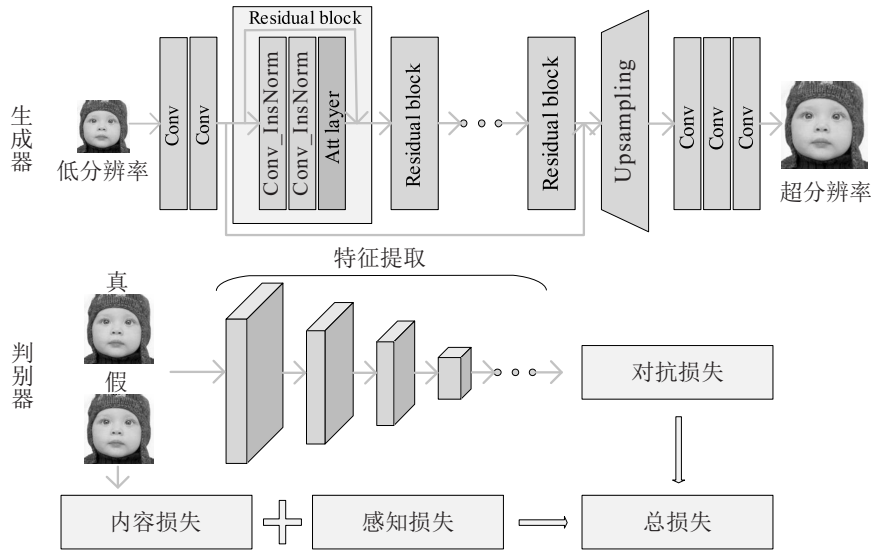


图1 SRGAN模型

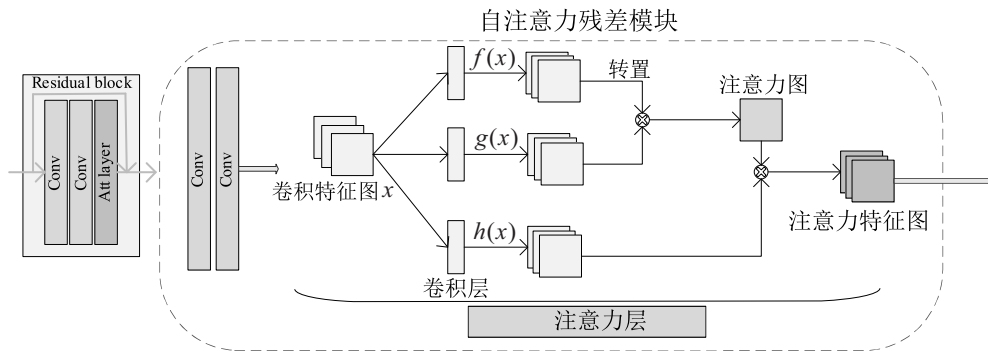


图2 自注意力层原理图

最近几年,注意力模型在深度学习各个领域得到了广泛的应用^[21]. 由于卷积操作主要用于处理邻域之间的信息,单独的卷积层对图像中的全局依赖性进行建模在计算上是非常低效的. 自注意力机制^[22]是从全局信息中选择出对当前任务目标更关键的信息,可以很好地利用图像全部的特征信息. 另外,自注意力机制还可以更好地处理图像中长范围、多层次的依赖关系,协调生成图像中每一个位置的远近细节. 为此,本文考虑在残差模块中引入自注意力机制层,利用全局特征信息更好地重建出超分辨率图像的纹理细节.

如图2所示,从残差块第2层卷积层中提取的特征图 $x \in R^{C \times N}$ 分别经过两个核为1的卷积层变换到了两个特征空间 $f(x)$ 和 $g(x)$,其中 $f(x) = W_f x$ 用于提取像素特征, $g(x) = W_g x$ 用于提取全局特征. 通过对 $f(x)$ 和 $g(x)$ 进行变换来计算注意力图 β_{ji} ,即

$$\beta_{ji} = \frac{\exp(s_{ij})}{\sum_{i=1}^N \exp(s_{ij})}. \quad (3)$$

其中: $s_{ij} = f(x_i)^T g(x_j)$, β_{ji} 表示在合成第 j 个区域

时模型对第 i 个位置的关注度. 输出的注意力层表示为 $o = (o_1, o_2, \dots, o_i, \dots, o_N) \in R^{C \times N}$,其中

$$o_i = \sum_{i=1}^N \beta_{ji} h(x_i), \quad (4)$$

$h(x_i) = W_h x_i$. 进一步将注意力层的输出乘以比例参数并添加回输入特征图,最终输出为

$$y_i = \gamma o_i + x_i. \quad (5)$$

其中: γ 为预先定义的系数, y_i 表示最终的输出. 输出的注意力特征图会进入下一个注意力机制网络中继续特征提取与学习的过程.

虽然批量归一化已被证明在许多图像分类任务中是有效的,但是这种归一化方法在一定程度上会降低图像生成任务的性能. 在图像风格转换任务中,使用实例归一化代替批量归一化,可以很好地实现图像转换的效果. 为此,将生成器中的批处理规范化层替换为实例规范化层,以提高图像超分辨率重建任务的性能. 实例归一化是将归一化应用到单个图像上,有利于单个图像实例的生成. 设 $x \in R^{T \times C \times W \times H}$ 表示包含了 T 张图像的一个输入批次, C 表示通道数, W

和 H 表示宽和高, 设 x_{tkmn} 表示这批图像的第 $tkmn$ 个元素, 其中 m 和 n 表示空间维度, k 表示输入的特征通道, t 表示成批图像的索引, 则实例正则化公式可表示为

$$y_{tkmn} = \frac{x_{tkmn} - u_{tk}}{\sqrt{\sigma_{tk}^2 + \epsilon}}. \quad (6)$$

其中: ϵ 为可变参数, $u_{tk} = \frac{1}{HW} \sum_{l=1}^W \sum_{p=1}^H x_{tklp}$, $\sigma_{tk}^2 = \frac{1}{HW} \sum_{l=1}^W \sum_{p=1}^H (x_{tklp} - u_{tk})^2$.

1.2 损失函数

为重建出具有清晰纹理的超分辨率图像, SRAGAN 在损失函数中引入对抗损失, 通过生成器网络重建出的超分辨率图像尽可能去欺骗判别器网络, 以更加接近真实超分辨率图像. GAN 的对抗性损失为

$$L(G, D) = E_{I_{HR} \sim p(I_{HR})} [\log D_{\theta}(I_{HR})] + E_{I_{LR} \sim p(I_{LR})} [\log(1 - D_{\theta}(G_{\theta}(I_{LR})))]. \quad (7)$$

其中: $G_{\theta}(I_{LR})$ 表示重建后的超分辨率图像, $p(I_{HR})$ 为超分辨率图像训练样本, $p(I_{LR})$ 为低分辨率图像训练样本. 在对抗训练过程中, 使用 Wasserstein 距离来优化对抗损失 $L_W(G, D)$, 即

$$L_W(G, D) = E_{I_{HR} \sim p(I_{HR})} D_{\theta}(I_{HR}) - E_{I_{LR} \sim p(I_{LR})} D_{\theta}(G_{\theta}(I_{LR})). \quad (8)$$

基于所有样本的生成器损失 L_G 和判别器损失 L_D 分别定义为

$$L_G = \sum_{n=1}^N (-D_{\theta}(G_{\theta}(I_{LR}))), \quad (9)$$

$$L_D = \sum_{n=1}^N (D_{\theta}(I_{HR}) - D_{\theta}(G_{\theta}(I_{LR}))). \quad (10)$$

为衡量网络重建出的超分辨率图像与真实超分辨率图像之间的相似程度, 通常是将相应模型重建的超分辨率图像与真实超分辨率图像以像素为单位计算其均方差损失函数, 该评价方式在一定程度上削弱了模型的泛化能力, 使其仅局限于像素级信息的重建. 感知损失是在预训练的深度网络的激活层上定义的, 通过激活后的特征值来计算相应的损失函数, 可以使网络学习到更鲁棒的效果. 但是在比较深的网络之中, 激活后的特征会变得非常稀疏, 导致监督能力比较差, 所以 SRAGAN 采用激活前的特征值来计算感知损失. 激活前的特征可以更好地表示图像

的特征信息, 它可以很好地监督重建后的图像与原图像纹理的一致性.

采用训练好的 VGG-19 网络来获得相应激活层前的特征值, 再通过计算生成的超分辨率图像特征图和原图像特征图之间的欧氏距离得到感知损失 L_P , 即

$$L_P(I_{HR}, \hat{I}_{HR}) = \sum (\phi(I_{HR}) - \phi(G_{\theta}(I_{LR})))^2, \quad (11)$$

其中 ϕ 是预训练好的 VGG-19 网络.

为了保证生成的超分辨率图像和原图像在内容上不会有较大的偏差, 模型中还需要加入内容损失函数. SRAGAN 采用 Charbonnier 损失^[23] 来代替传统的 L_2 损失, 以提高网络的性能. Charbonnier 损失 L_C 定义如下:

$$L_C(I_{HR}, \hat{I}_{HR}) = \sum (\rho(I_{HR} - G_{\theta}(I_{LR}))). \quad (12)$$

其中: $\rho(x) = \sqrt{x^2 + \delta^2}$ 是 Charbonnier 惩罚函数, δ 为超参. 由于 Charbonnier 损失可以很好地处理异常值, 它具有更强大的监督能力, 同时具有更强的鲁棒性. 综合以上损失, 生成器网络模型的总损失 L_{total} 为

$$L_{total} = L_P + \lambda_1 L_G + \lambda_2 L_C, \quad (13)$$

其中 λ_1 和 λ_2 是系数, 用来平衡不同的损失.

1.3 生成器和判别器网络参数设置

训练时, 生成器主要任务是生成与真实超分辨率图像尽可能相似的超分辨率图像, 判别器则试图区分重建后的超分辨率图像和真实超分辨率图像的差异. 生成器网络主要由 3 部分组成, 分别是多卷积核下采样模块、自注意力残差模块和上采样图像重建模块. 如图 3 所示, 首先采用 3 个不同尺寸 3×3 、 5×5 和 7×7 的卷积核对输入图像进行特征提取, 通过 3 个不同尺寸的感受野对图像进行卷积, 然后把卷积后的特征图和原图像送入 Concatlayer 层, 再送入一个 1×1 的卷积层, 激活函数使用 ParametricRelu^[24]. 在单个自注意力残差模块中, 采用两个卷积层和一个自注意力层, 卷积层后面接 InstanceNormLayer(实例归一化层), 激活函数用 ParametricRelu, 同时加入跳跃连接. SRAGAN 生成器中采用 16 个相同的自注意力残差模块, 在最后一个残差块输出层也加入跳跃连接. 在图像上采样部分, 采用两个 Subpixel 卷积层^[12] 进行图像像素扩增, 通过加深深度网络层数, 以便更好地进行图像重建. 在判别器网络中, 采用深度卷积神经网络, 由于 BatchNorm(批量归一化) 在图像分类任务中的有效性, 在判别器网络中仍然采用

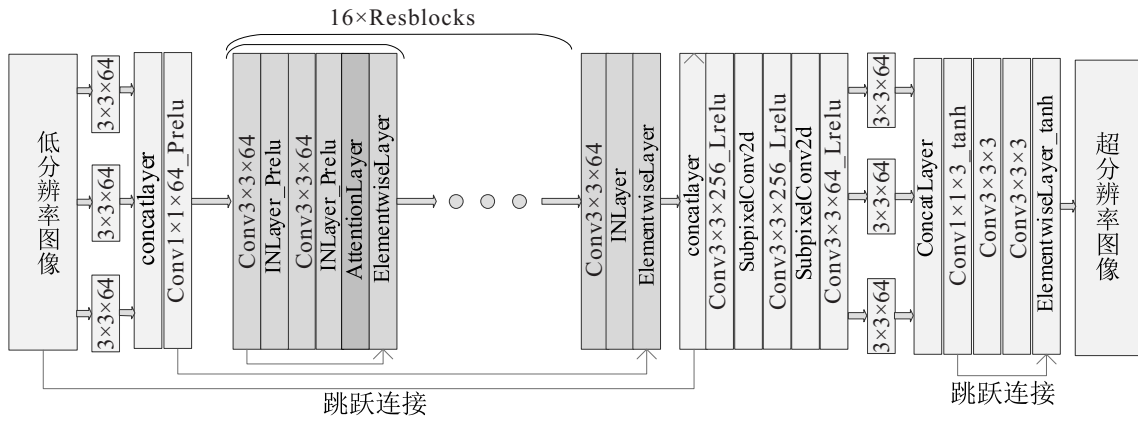


图3 生成器网络具体参数结构

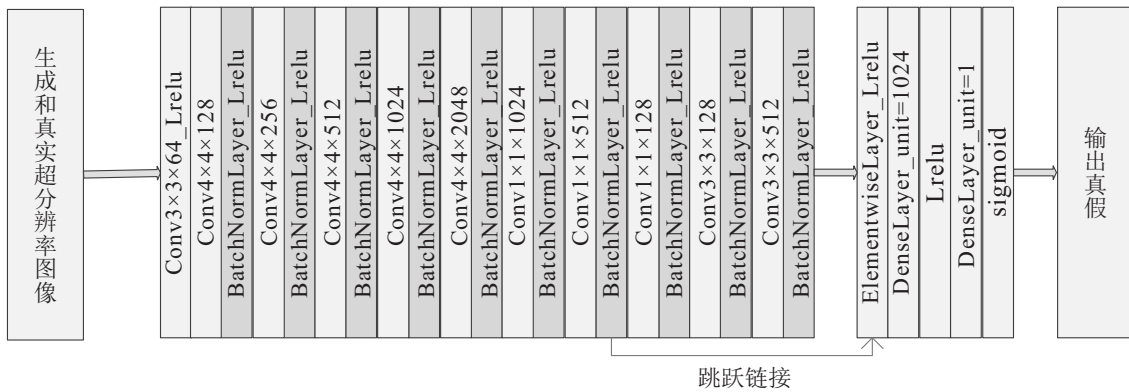


图4 判别器网络具体参数结构

BatchNorm层,使用LeakyRelu($\alpha = 0.2$)作为激活函数,每一层卷积网络的卷积核的大小和通道数如图4所示.在经过卷积层之后,将特征图输入到全连接层,然后再输入到sigmoid激活函数进行二分类,判别输出超分辨图像的真假.

1.4 算法步骤

输入:训练样本为配对的图像训练集(I_{LR}, I_{HR}),包含低分辨率图像训练集 I_{LR} 和超分辨率图像集 I_{HR} .

step 1: 将 I_{LR} 输入到生成器网络 G 中,输出得到重建的 \hat{I}_{HR1} 图像;

step 2: 计算 \hat{I}_{HR1} 图像和对应真实的 I_{HR} 图像的均方差(MSE);

step 3: 根据MSE更新生成器,迭代 T_1 次,得到预训练的生成器网络 G_1 ;

step 4: 将 I_{LR} 输入到预训练的生成器网络 G_1 中,输出得到重建的 \hat{I}_{HR2} 图像;

step 5: 将 \hat{I}_{HR2} 和 I_{HR} 输入到判别器 D 中,计算判别损失 L_D ,并更新 D ;

step 6: 将 \hat{I}_{HR2} 和 I_{HR} 分别输入到预训练的VGG-19网络中,通过激活前的特征值计算感知损失 L_P ;

step 7: 计算 \hat{I}_{HR2} 和 I_{HR} 之间的内容损失 L_C ;

step 8: 计算总损失 L_{total} ,并根据 L_{total} 更新生成器网络 G ;

step 9: 迭代step 4到step 8,总共 T_2 次.

输出: 生成器网络 G ,即为图像超分辨率重建模型.

2 实验分析

2.1 实验设置

实验使用DIV2K和RAISE作为训练数据集,其中DIV2K包含800张超分辨率图像和对应的经过双三次插值处理过的低分辨率图像,RAISE数据集包含8156张通过高质量照相机捕获的超分辨率图像和对应的低分辨率图像.测试集采用3个广泛使用的基准标记数据集set5,set14和BSD100,除此之外还使用valid数据集进行测试.所有的实验均在低分辨率图像和超分辨率图像之间以4倍的比例因子进行,这相当于图像像素增加了16倍.为公平而言,实验采用在超分辨率图像重建问题中使用最为广泛的峰值信噪比(PSNR)和图像结构相似度(SSIM)进行实验效果衡量.对比方法包括Bicubic、SRCNN^[8]、FSRCNN^[13]、SelfExSR^[25]、LapSRN^[14]和SRGAN^[16].

对于每个超分辨率图像裁剪出 96×96 大小的图像块进行训练,图像感受野的大小对图像的重建

效果有很大的影响,感受野越大图像所获得的信息越多,训练每一批图片的数量均为16. 由于生成器是完全卷积的,可以将生成器模型应用于任意大小的图像. 实验使用Adam优化器,其动量和权重衰减分别是0.9和0.000 1. 训练包含两步:第1步使用MSE对生成器网络进行模型预训练,以避免不必要的局部最优,预训练使用学习率0.000 1迭代200次;第2步对生成器和判别器进行交替更新,使用学习率0.000 1迭代20 000次,同时使用Wasserstein距离进行对抗训练

的优化. 感知损失部分,采用VGG-19网络中第2个模块第2层卷积激活前的特征值来计算感知损失. 训练中,参数 λ_1 和 λ_2 分别取0.005和0.01, ϵ 取 10^{-3} .

2.2 实验量化结果

模型训练完成后,对测试数据集中的低分辨率图像进行超分辨率重建,表1给出了量化结果比较. 由表1可以看出:与对比方法相比,SRGAN在绝大多数数据集上的PSNR和SSIM均有一定的提高,验证了SRGAN的有效性和优越性.

表1 图像超分辨率重建模型量化结果比较

算法		Bicubic	SRCNN ^[8]	FSRCNN ^[13]	SelfExSR ^[25]	LapSRN ^[14]	SRGAN ^[16]	SRGAN
set 5	PSNR	28.42	30.49	30.71	30.33	31.54	29.41	31.94
	SSIM	0.81	0.862	0.865	0.861	0.885	0.847	0.90
set 14	PSNR	26.1	27.61	27.7	27.54	28.19	26.2	28.12
	SSIM	0.704	0.754	0.756	0.756	0.772	0.74	0.793
BSD100	PSNR	25.96	26.91	26.79	26.84	27.32	25.16	26.82
	SSIM	0.669	0.712	0.714	0.712	0.728	0.669	0.744
vaild	PSNR	28.55	30.12	30.52	30.23	31.32	30.96	31.91
	SSIM	0.708	0.732	0.746	0.768	0.845	0.804	0.852

2.3 纹理细节实验结果

纹理细节的精确重建一直都是图像超分辨率重建问题的难点,因为通过峰值信噪比衡量出来的图像结果往往与人眼视觉感知的图像结果不同,过于平滑的图像无法反映真实图像的纹理细节. 在实验结果中,选取set 5、set 14和BSD100数据集的部分重建图像,并把部分纹理细节放大,对比实验结果,可以清楚地看到通过SRGAN重建后的超分辨率图像具有更清晰的纹理细节.

在图5中,将set 5儿童面部重建图像方框中的纹理细节放大,可以看出:与SRCNN、FSRCNN、SelfExSR和LapSRN相比,SRGAN具有更清晰的眼

睛睫毛纹理. 同时与SRGAN相比,SRGAN也具有更清晰的轮廓,显得更加真实. 在图6和图7中,从set 14中蝴蝶翅膀纹路和BSD100中老人额头皱纹这些局部放大的细节来看,SRGAN模型在重建图像纹理细节方面具有更优秀的性能,较好地重建出了超分辨率图像清晰的纹理和更逼真的感知细节.

在valid数据集上,从实验结果中选取一张具有代表性的重建后古堡超分辨率图像,并比较不同位置的局部纹理细节,如图8所示. 建筑图像的几何棱角结构更能显示出图像重建后纹理细节的清晰性,在4个不同的位置选出对应的局部图像进行放大(图中用方框标出的位置),可以看出在窗户和屋檐等线条

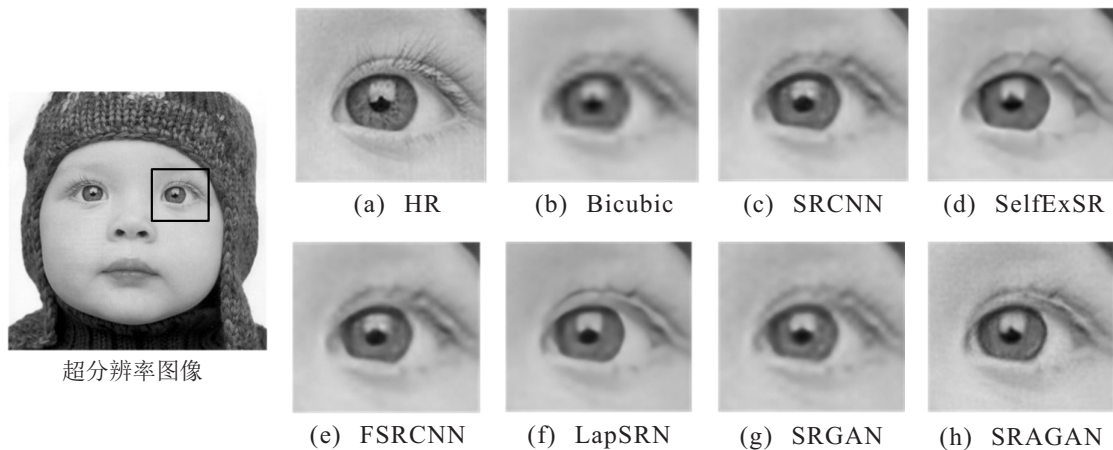


图5 局部纹理细节比较(set 5数据集)

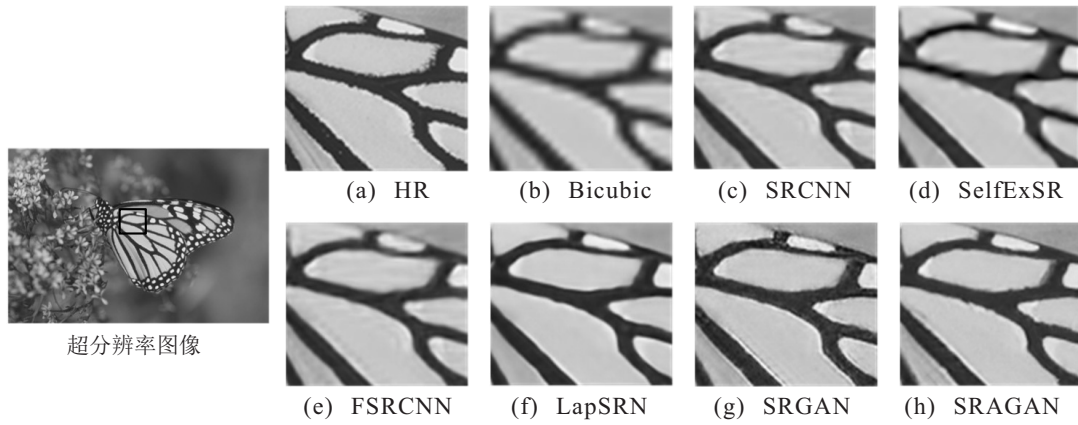


图6 局部纹理细节比较(set 14数据集)

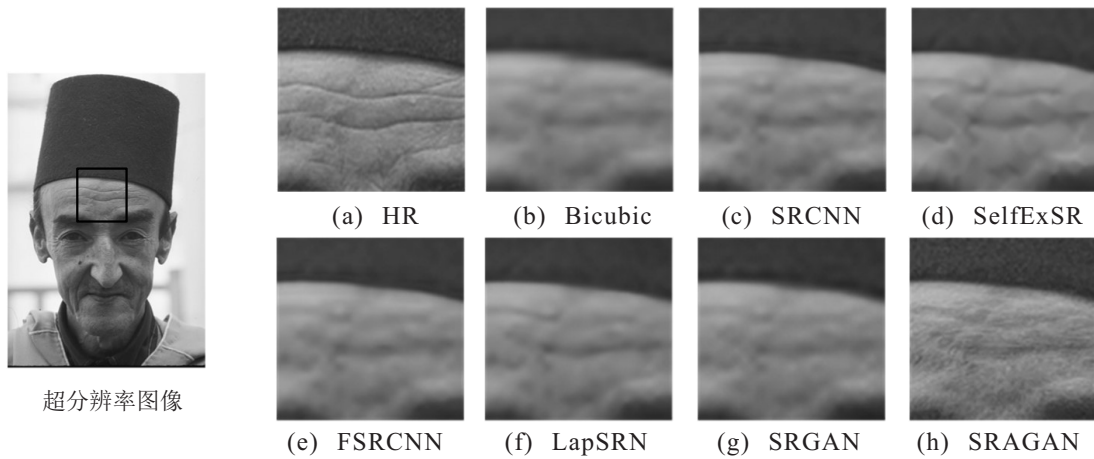


图7 局部纹理细节比较(BSD100数据集)

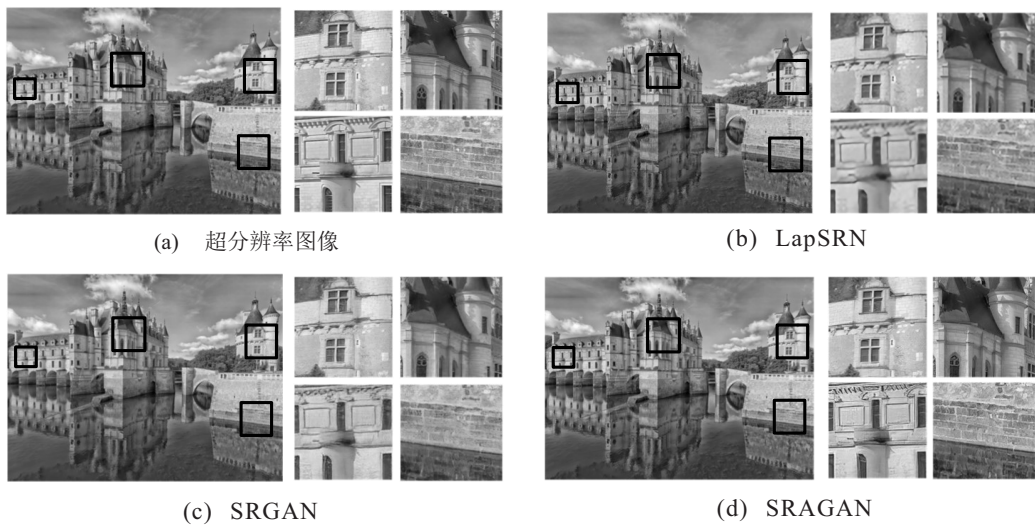


图8 局部纹理细节对比(vaild数据集)

结构上SRAGAN均优于其他方法。

与SRCNN相比,对抗损失的加入可以产生更加真实的结果,模型不再局限于单纯的像素损失,而是生成更接近于人类视觉感知的高质量图像.与SRGAN方法相比,自注意力机制模块的引入,可以很好地协调图像中每一个位置信息的关系,同时利用图像的全局特征信息来生成更清晰的超分辨率

图像.另一方面,由于实例归一化层是在单个图像实例上进行归一化的,更有利于图像的生成任务.此外,感知损失通过使用VGG-19网络激活前的特征来约束,可以更好地监督超分辨率图像的纹理恢复.实验表明,作为一种通用的图像超分辨率重建模型框架,SRAGAN可以重建出具有逼真纹理细节的超分辨率图像,适用于解决各种低分辨率图像的清晰化,

例如人脸识别、医学成像和卫星成像等。

2.4 内容损失函数的分析

最后,通过实验来进一步探索内容损失函数对算法性能的影响. 为了验证 Charbonnier 损失函数的效果,实验中将 L_2 损失和 Charbonnier 损失分别作为内容损失来训练模型. 通过训练迭代200次来分别比较通过两种损失函数训练后得到的 SSIM 分数,实验结果如表2所示. 由表2可以看出,在 set 5 和 vaild 数据集上,使用 Charbonnier 损失作为内容损失可以使模型表现出更优秀的性能,相比较 L_2 损失可以得到更高的 SSIM 分数.

表2 内容损失对比表

数据集	内容损失	SSIM	训练次数
set 5	Charbinnier	0.862	200
set 5	L_2	0.854	200
vaild	Charbinnier	0.823	200
vaild	L_2	0.805	200

3 结论

本文提出一种基于自注意力生成对抗网络的图像超分辨率重建模型,构建了自注意力机制残差模块,通过设计深层生成器网络,利用生成对抗网络进行更好的图像超分辨率重建. SRAGAN 模型中使用 Charbonnier 损失函数代替传统的 L_2 内容损失函数以提高重建精度,感知损失部分使用 VGG-19 网络激活前的特征来约束感知损失,这样可以更好地监督超分辨率图像的纹理恢复. 训练时使用 Wasserstein 距离来优化对抗训练,提高模型训练的稳定性. 实验表明,与一些先进模型相比,SRAGAN 在 PSNR 和 SSIM 指标上均有一定的提升,同时在重建后的超分辨率图像纹理细节上也有一定的改善.

参考文献(References)

[1] Shi W Z, Caballero J, Ledig C, et al. Cardiac image super-resolution with global correspondence using multi-atlas patchmatch[C]. Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer Verlag, 2013: 9-16.

[2] Thornton M W, Atkinson P M, Holland D A. Sub-pixel mapping of rural land cover objects from fine spatial resolution satellite sensor imagery using super-resolution pixel-swapping[J]. International Journal of Remote Sensing, 2006, 27(3): 473-491.

[3] Bilgazyev E, Efraty B, Shah S K, et al. Improved face recognition using super-resolution[C]. Proceedings of IEEE International Joint Conference on Biometrics.

Piscataway: IEEE Press, 2011: 1-7.

[4] Chang H, Yeung D Y, Xiong Y M. Super-resolution through neighbor embedding[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2004: 1275-1282.

[5] Timofte R, Vincent De Smet, Luc Van Gool. A+: Adjusted anchored neighborhood regression for fast super-resolution[C]. Proceedings of Asian Conference on Computer Vision. Berlin: Springer Verlag, 2014: 111-126.

[6] 吴成东, 卢紫薇, 于晓升. 基于加权随机森林的图像超分辨率算法研究[J]. 控制与决策, 2019, 34(10): 2243-2248.
(Wu C D, Lu Z W, Yu X S. Image super resolution reconstruction algorithm based on weighted random forest[J]. Control and Decision, 2019, 34(10): 2243-2248.)

[7] 吕恩辉, 王雪松, 程玉虎. 基于反卷积特征提取的深度卷积神经网络学习[J]. 控制与决策, 2018, 33(3): 447-454.
(Lv E H, Wang X S, Cheng Y H. Deep convolution neural network learning based on deconvolution feature extraction[J]. Control and Decision, 2018, 33(3): 447-454.)

[8] Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 38(2): 295-307.

[9] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 1646-1654.

[10] Kim J, Lee J K, Lee K M. Deeply-recursive convolutional network for image super-resolution[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 1637-1645.

[11] Tai Y, Yang J, Liu X. Image super-resolution via deep recursive residual network[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 2790-2798.

[12] Shi W, Caballero J, Huszár Ferenc, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 1871-1883.

[13] Chao D, Chen C L, Tang X. Accelerating the super-resolution convolutional neural network[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 391-407.

- [14] Lai W S, Huang J B, Ahuja N, et al. Deep laplacian pyramid networks for fast and accurate super-resolution[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 5835-5843.
- [15] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 1132-1140.
- [16] Ledig C, Wang Z, Shi W, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 105-114.
- [17] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]. Proceedings of Annual Conference on Neural Information Processing Systems. New York: NIPS Press, 2014: 2672-2680.
- [18] 王星, 杜伟, 陈吉, 等. 基于深度残差生成式对抗网络的样本生成方法[J]. 控制与决策, 2020, 35(8): 1887-1894.
(Wang X, Du W, Chen J, et al. Sample generation based on residual generative adversarial network[J]. Control and Decision, 2020, 35(8): 1887-1894.)
- [19] Zhu J Y, Park T, Isola P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]. Proceedings of IEEE Conference on International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2242-2251.
- [20] Yeh R A, Chen C, Lim T Y, et al. Semantic image inpainting with deep generative models[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 6882-6890.
- [21] 李旭, 姚春龙, 范丰龙, 等. 结合注意力机制的循环神经网络复述识别模型[J]. 控制与决策, 2021, 36(1): 152-158.
(Li X, Yao C L, Fan F L, et al. Recurrent neural networks based paraphrase identification model combined with attention mechanism[J]. Control and Decision, 2021, 36(1): 152-158.)
- [22] Liu X, Li K, Li K. Attentive semantic and perceptual faces completion using self-attention generative adversarial networks[J]. Neural Processing Letters, 2020: 51(1): 211-229.
- [23] Barron J T. A general and adaptive robust loss function[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 4331-4339.
- [24] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 1026-1034.
- [25] Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars[C]. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 5197-5206.

作者简介

王雪松(1974—), 女, 教授, 博士生导师, 从事机器学习及模式识别、人工智能等研究, E-mail: wangxuesongcumt@163.com;

晁杰(1995—), 男, 硕士生, 从事生成对抗网络的研究, E-mail: vae_gan123@163.com;

程玉虎(1973—), 男, 教授, 博士生导师, 从事机器学习、模式识别与智能系统等研究, E-mail: chengyuhu@163.com.

(责任编辑: 孙艺红)