

doi:10.11835/j.issn.1000-582X.2023.05.010

利用图像掩膜优化基于生成对抗网络的图像超分辨率模型

蒋琪雷, 马原曦

(1. 上海科技大学 信息学院, 上海 201210; 2. 中国科学院 上海微系统与信息技术研究所, 上海 200050;
3. 中国科学院大学, 北京 100049)

摘要:深度学习在一定程度上解决了从低分辨率图像中恢复出高分辨率图像这一图像超分辨率问题。目前基于生成对抗网络(generative adversarial network, GAN)的方法可以从超分辨率数据集中学习低/高分辨率图像映射关系,从而生成具有真实纹理细节的超分辨率图像。然而,基于GAN的图像超分辨率模型训练通常不稳定,其结果往往带有纹理扭曲和噪声等问题,提出了采用掩膜(mask)模块以辅助对抗网络训练。在网络训练过程中,掩膜模块根据生成网络输出的超分辨率结果和原始高分辨率图像,计算得到相应观感质量信息,并进一步辅助对抗网络训练。在实验中对3个最近提出的基于GAN的图像超分辨率模型进行修改,引入掩膜模块,修改后的模型在超分辨率图像输出的观感和真实感量化指标上均有明显地提升。掩膜模块的优点是可以进一步提升基于GAN的图像超分辨率网络输出的超分辨率图像观感质量,并仅需对生成对抗网络训练框架进行修改,因此适用于多数基于GAN的图像超分辨率模型的进一步优化。

关键词:深度学习;超分辨率算法;对抗生成网络

中图分类号:TP391

文献标志码:A

文章编号:1000-582X(2023)05-093-09

Optimization of generative adversarial network based image super-resolution by using image mask

JIANG Qilei, MA Yuanxi

(1. School of Information Science and Technology, ShanghaiTech University, Shanghai 201210, P. R. China; 2. Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai 200050, P. R. China; 3. University of Chinese Academy of Sciences, Beijing 100049, P. R. China)

Abstract: Inferring high resolution image from single low resolution (LR) input is ill-posed and deep learning helps to some extent. The latest algorithms take the advantage of the Generative Adversarial Network (GAN) and present photo-realistic results by learning low/high resolution mappings from super resolution datasets. However, training of GANs can be hard and traditional GAN-based architectures often exhibit noise and texture distortion in their super-resolution (SR) results. In this paper, a mask-aided adversarial training strategy for current GAN-based

收稿日期:2021-03-11

基金项目:国家重点研发资助计划(2018YFB2100500);国家自然科学基金资助项目(61976138; 61977047)。

Supported by the National Key Research and Development Program (2018YFB2100500) and National Natural Science Foundation of China (61976138, 61977047).

作者简介:蒋琪雷(1995—),男,硕士,主要从事计算机图形学与计算机视觉方向研究,(E-mail)jiangq11@shanghaitech.edu.cn。

SR frameworks is proposed. During training, mask module helps the discriminator by introducing additional perceptual quality information with generator's outputs and the ground truth images. In experiment, three current state-of-the-art GAN-based SR models are selected and the mask module is integrated into their adversarial training. The improved mask-aided models yield better results in both quantitative and qualitative benchmarks than the original ones. Mask module only modifies GAN framework and thus is suitable for many GAN-based solutions for further improving the SR perceptual quality.

Keywords: deep learning; super resolution algorithm; generative adversarial network

单张图像超分辨率任务(以下简称为图像超分辨率)是计算机视觉领域广泛研究的问题之一。目前,该问题在监控安防和社交网络等方面也有着较高价值。图像超分辨率任务目标是从分辨率受限的单张图像中生成具有较高观感质量的高分辨率图像。基于深度学习的图像超分辨率方法利用神经网络优势,通过恢复从低分辨率图像到高分辨率图像的非线性映射,获得了较好的超分辨率图像结果。但是,由于低分辨率图像本身采样不足而丢失较多高频信息,因此也无法获得与真实高分辨率图像相当的纹理细节,其仍然有着很大的优化空间。

目前,流行的大多数基于深度学习的图像超分辨率模型,大致可以分为面向峰值信噪比(peak signal-to-noise ratio, PSNR)的网络和基于生成对抗网络(generative adversarial network, GAN)的模型。峰值信噪比的计算依赖于2张图像之间的均方误差,因此可用于量化超分辨率图像的恢复水平。SRCNN^[1]是面向PSNR的网络的先驱,它使用卷积神经网络进行从低分辨率到高分辨率图像映射的学习。由于基于卷积神经网络的模型通常无法较好地学习高频信息,这种基于编码器/解码器的体系结构在较大的缩放比例下,通常无法恢复高分辨率图像当中的纹理细节。VDSR^[2]为了进一步提升网络的学习能力,其卷积层数增加到了20。由于网络的低分辨率输入图像通常较为完整地保留了场景的低频分量,为了易于网络学习,该网络模型采用了残差学习(residual learning)方法,并仅预测了高频分量。随后出现的一些网络,例如EDSR^[3]和RDN^[4],均使用了残差块^[5]的变体来构建更深的网络。因此这些面向PSNR的模型在其重建的超分辨率图像上,均表现出了较高的峰值信噪比,有着很高的重建精度。但是峰值信噪比在捕捉高频纹理的能力上非常有限,而且该值无法较好的量化图像的观感质量。大多数面向PSNR的图像超分辨率模型都存在一定的模糊、失真等问题。

基于Goodfellow等人^[6]所提出的生成对抗网络训练框架,通常能够使得图像超分辨率网络生成更加逼真的结果。生成对抗网络的生成器的输入为低分辨率图像,并返回超分辨率结果,而鉴别器则通过区分超分辨率图像和原始高分辨率图像之间的差异来保留更多图像的高频信息。这样,生成器可以将高频纹理部分添加入其超分辨率结果,并生成比面向PSNR的图像超分辨率网络更逼真的图像。超分辨率生成对抗网络SRGAN^[7]引入了由Johnson等人^[8]所提出的观感损失函数,并获得了逼真的图像超分辨率结果。该观感损失函数依赖于由Russakovsky等人^[9]所创建的ImageNet数据库,进行预先训练的分类网络所计算得到。增强型SRGAN(enhanced SRGAN, ESRGAN)^[10]则应用了一种相对性鉴别器^[11],使得训练效果和稳定性得到了提升,并进一步调整了观感损失函数来优化SRGAN网络结构,以生成在视觉上更加逼近真实图像的超分辨率的结果。Self-Attention GAN(SAGAN)^[12]则引入了自我关注机制以突出图像内不同区域之间的远距离依赖,从而使超分辨率图像看起来更加自然。然而,通过使用整个图片作为对抗网络的输入,仍很难在训练中得到稳定的收敛,由此生成的图像在观感上仍具有不可预测的噪声和纹理的扭曲。

在此基础上,提出了一种新颖的,用以辅助基于GAN的图像超分辨率模型方法。通过使用掩膜(mask)模块来辅助对抗网络训练,将更多与真实感相关的信息纳入对抗网络中,获得更好的超分辨率结果。掩膜的计算依赖于生成网络所得到的超分辨率图像和对应的原始高分辨率图像。通过将掩膜引入鉴别器可以更好地帮助生成网络,恢复低分辨率图像输入中潜在的高频分量。以图像超分辨率网络ESRGAN^[10]为例,掩膜模块在生成对抗网络训练框架当中的位置如图1所示。基于GAN的图像超分辨率网络和面向PSNR的方法除了对抗学习部分的不同,其余均大致拥有相同的结构。研究提出的掩膜优化生成对抗网络训练的方法,在很大程度上更好地结合了两类网络的优点,因此可以生成在视觉效果上相较之前的网络更具有真实感的超分辨率图像。主要贡献有以下2点:

- 1) 提出了掩膜模块用以辅助基于 GAN 的图像超分辨率网络训练。该方法仅对生成对抗网络训练框架进行修改,因此适合应用到很多现有的此类模型上,以进一步提升超分辨率效果;
- 2) 全面分析并解释了面向 PSNR 和基于 GAN 的图像超分辨率网络模型,其本身结构所带来的优缺点,以及超分辨率结果中相应的特征。

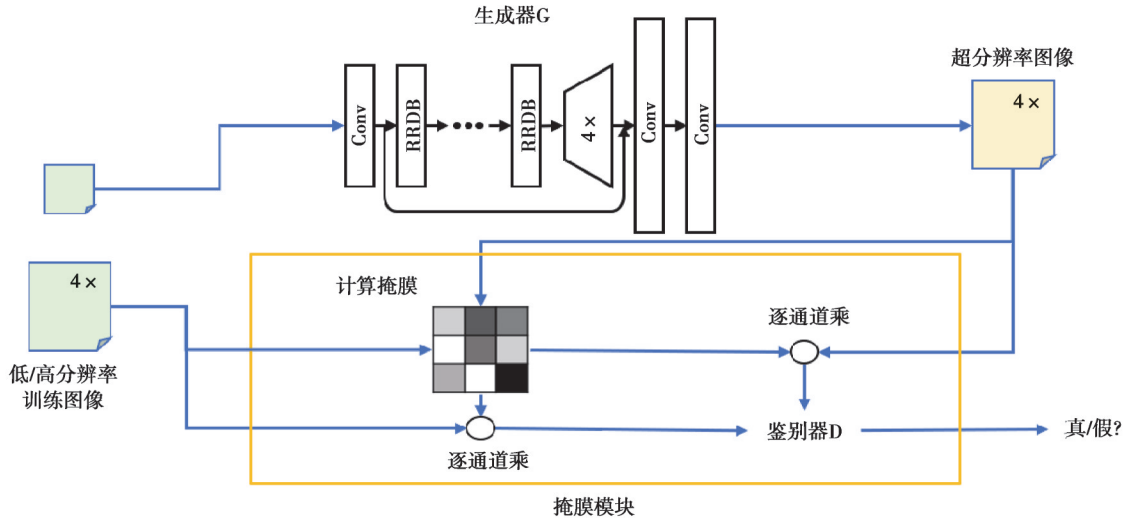


图 1 掩膜模块在生成对抗网络训练框架中的位置(以图像超分辨率网络 ESRGAN 为例,4 倍超采样)

Fig. 1 The position of the mask module in GAN framework (Taking ESRGAN as an example, with 4 times super resolution)

1 观感质量信息的估算

常见的基于 GAN 的图像超分辨率模型的损失函数通常包括均方误差损失项 $Loss_{MSE}$ 以及 GAN 损失项 $Loss_{GAN}$,并具有如下形式

$$Loss = Loss_{GAN} + Loss_{MSE} + \dots \quad (1)$$

网络从对应的低/高分辨率图像数据集中学习映射关系,生成网络的输入为低分辨率图像,输出为超分辨率图像。随后超分辨率图像传入对抗网络,通过鉴别器判断图像为真实图像的概率。生成网络的优化目标是降低超分辨率结果和对应的高分辨率图像之间的均方误差损失,以及提升超分辨率结果被鉴别器判断为真实图像的概率。

1.1 一种图像等价类的定义

一个标准的图像超分辨率神经网络通常需要一个包含了大量的真实高分辨率图像的数据集。定义该数据集为

$$\Omega = \{ I | I \text{ 为高分辨率图像} \} \quad (2)$$

定义 1. 在数据集 Ω 中的等价关系 \sim , 表示为“可以降低采样至无法区分的低分辨率图像。”

等价关系是二元关系,并且具有自反性(reflexive)、对称性(symmetric)、和传递性(transitive)。在数据集 Ω 中任意采样 2 张图像 I_1 和 I_2 ,可观察到:

- 1) $I_1 \sim I_1$ (自反性);
- 2) $I_1 \sim I_2$ 当且仅当 $I_2 \sim I_1$ (对称性);
- 3) 如果 $I_1 \sim I_2$ 而且 $I_2 \sim I_3$, 则 $I_1 \sim I_3$ (传递性)。

从数据集 Ω 中采样的图像 I 的等价类可以表示成

$$[I]_{\sim} = \{ x \in \Omega | x \sim I \} \quad (3)$$

从等价类的表达式中可以看出,图像 I 的等价类由相似的高分辨率图像所组成, $[I]_{\sim}$ 中的元素和 I 有着不可区分的降采样结果。

图2为图像等价类的一个例子。其中,网格代表像素的大小,低分辨率纹理为 1×1 像素区域,表示低分辨率图像中出现的纹理,对应4倍上采样的超分辨率纹理为 4×4 像素区域。由于低分辨率纹理采样不足,无法包含对应超分辨率纹理中的高频信息;图中4种超分辨率纹理均可以作为图像超分辨率网络的输出,对低分辨率纹理进行解释。因此,图中出现的4种超分辨率纹理均处于相同的等价类,该等价类内的图像在降采样后损失了高频信息,得到的结果均接近于图中所示的低分辨率纹理。

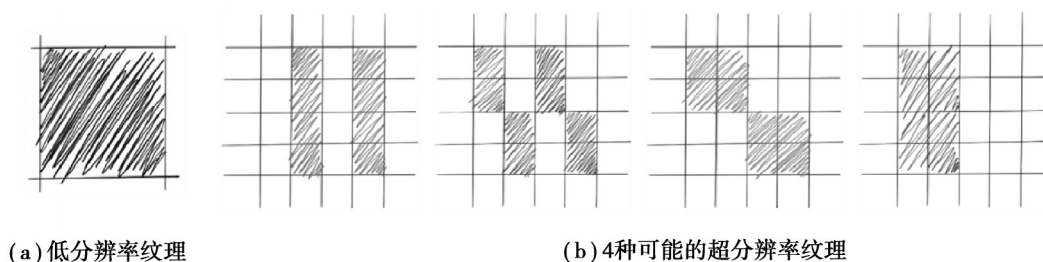


图2 等价类的概念图示

Fig. 2 Diagram of the image equivalence class

1.2 超分辨率图像误差分析

根据选择的损失函数不同,网络训练得到的超分辨率结果也不同。网络的最优参数集 θ_{Ω}^* 通常以最小化在特定数据集中的经验误差得到。在数据集 Ω 中,有

$$\theta_{\Omega}^* = \arg \min_{\theta} E_{I \in \Omega} [l_{\theta}(I)]. \quad (4)$$

这里, $l_{\theta}(\cdot)$ 是网络定义的损失函数。面向PSNR的图像超分辨率网络通常以均方误差损失项为主,在 Ω 数据集上进行训练可以得到的最优参数集为

$$\theta_{\Omega}^* = \arg \min_{\theta} \sum_{I \in \Omega} \|n(\text{DS}(I); \theta) - I\|_2^2. \quad (5)$$

其中: $n(\cdot; \theta)$ 网络; θ 是该网络本身所提供的训练参数; $\text{DS}(I)$ 对 Ω 中的高分辨率图像 I 进行降采样,得到相应的低分辨率图像。通过降低网络 $n(\text{DS}(I); \theta)$ 所输出的超分辨率图像和 I 之间的均方误差损失项,从而提升超分辨率图像的重建精度。

根据等价类 \sim 的定义,分析公式(5)中网络损失的下界

$$\begin{aligned} E_{I \in \Omega} [l_{\theta^*}(I)] &= E \left[\|n(\text{DS}(I); \theta^*) - I\|_2^2 \right] = \\ &= E \left[\|\overline{[I]_{\sim}} - I\|_2^2 \right] + \varepsilon = \\ &= \text{Var} \left[[I]_{\sim} \right] + \varepsilon. \end{aligned} \quad (6)$$

其中:网络 $n(\text{DS}(I); \theta^*)$ 的输入为 Ω 中的高分辨率图像 I 的降采样 $\text{DS}(I)$;根据等价类的定义, $[I]_{\sim}$ 中的图像或区域均可以解释该低分辨率输入。优化均方误差的过程将结果收敛至期望值,即 $\overline{[I]_{\sim}}$ 。网络所引入的误差和期望与样本之间的误差总和为 $\varepsilon > 0$ 。对于输入图像,网络所生成的超分辨率结果和原始高分辨率图像之间的误差并不能通过网络更为充分的训练来消除。其下界为输入图像对应的等价类方差。即

$$\ln f [E_{I \in \Omega} [l_{\theta^*}(I)]] = \ln f \left[\text{Var} \left[[I]_{\sim} \right] + \varepsilon \right] = \text{Var} \left[[I]_{\sim} \right]. \quad (7)$$

通过网络实际生成的结果也可以验证公式(7)。面向PSNR的图像超分辨率网络输出结果的质量和输入的低分辨率图像相关,包含较多高频信息场景的低分辨率输入图像所对应的等价类方差较大,其均值往往损失了较多的高频信息。因此,超分辨率结果往往在观感上呈现出模糊的纹理细节;然而在输入的低分辨率图像包含较少高频信息的情况下,潜在的等价类方差较小,期望可以很好地代表该等价类的元素。因此,在这

种情况下,面向 PSNR 的图像超分辨率网络往往可以生成具有较高重建精度的超分辨率图像。

2 掩膜模块辅助对抗网络训练的流程

基于 GAN 的图像超分辨率模型的损失函数通常包括均方误差损失项和 GAN 损失项。均方误差损失项提升了超分辨率结果的重建精度,但不可避免地损失了一定的高频信息;GAN 损失项可以更好地引导生成网络计算,得到具有真实感的超分辨率结果,但通常包含难以避免的纹理扭曲和噪点等特征。

公式(7)给出了进一步优化基于 GAN 的图像超分辨率网络的方向。对于生成网络得到的超分辨率结果,通过计算相应均方误差,可以得到图像不同区域的等价类方差估计;较大方差区域,通过优化均方误差损失函数得到的超分辨率图像往往无法很好地表示真实的结果,因此对抗网络需要加强此类区域的真实感。等价类方差和 GAN 损失项的系数成正比。通过公式(7)对超分辨率图像进行相应的计算,可以在对抗网络当中引入更多的超分辨率图像真实感信息。

进一步提出掩膜的概念,将对超分辨率图像的真实感估计,封装成辅助对抗网络训练的模块。结合公式(7),有

$$\text{mask} = kE_{I \in \Omega} [l_{\sigma}(I)] + b. \tag{8}$$

其中: k, b 为掩膜模块的相应参数;掩膜的数值范围为 $[0, 1]$,并对超出的该范围的数值进行截断处理。由上述讨论可知,掩膜数值较小的区域,均方误差较小,对应的等价类方差较小,因此超分辨率结果所在的等价类的期望,可以较好地解释网络输入的低分辨率图像。反之,则需要引入对抗网络训练,以增强超分辨率结果的真实感。因此掩膜数值较大的区域,GAN 损失项系数较大。在超分辨率结果输入到对抗网络之前,计算相应的掩膜,并且和超分辨率结果进行乘积,以有效约束 GAN 损失项在不同图像区域中的作用强度(具体流程见图 3)。生成网络从输入的低分辨率图像中计算得到相应的超分辨率图像,该结果和对应的高分辨率图像一起输入至掩膜模块,以计算图像掩膜。对抗网络的输入为超分辨率图像和掩膜的逐通道乘积结果,以添加额外的超分辨率图像真实感信息,并自动调节不同超分辨率图像区域的 GAN 损失项权重。



图 3 超分辨率 GAN 框架中集成掩膜模块的流程图

Fig. 3 Pipeline for integrating mask module into super resolution GAN framework

3 实验

笔者挑选了目前效果较好的3个基于GAN的图像超分辨率模型。通过比较应用掩膜模块前后的超分辨率结果,以验证所提出的掩膜模块的有效性;选取的模型为ESRGAN, PESR^[13], 和EDSR, 它们将作为基础模型,并在各自的生成对抗网络训练框架当中,添加掩膜模块(添加方法见图1)。选取了4个超分辨率数据集:Urban100^[14]由真实建筑图像组成,包含窗户,栏杆,砖石等结构性较强的纹理;PIRM数据集^[15]本身即作为图像超分辨率任务的验证数据集提出,因此包含各类检测超分辨率效果的场景;Berkeley数据集(BSD100)^[16]包含各类自然场景,最初作为检验图像分割任务所提出;DIVerse 2K(DIV2K)数据集^[17]也是一个包含各种高分辨率图像的数据集。每一类型均包含100张用于检验超分辨率效果的低/高分辨率图像。超分辨率结果量化方法采用观感系数(perceptual index, PI)^[18]

$$PI = \frac{(10 - Ma) + NIQE}{2} \quad (9)$$

其中, Ma^[19]和 NIQE^[20]均为无参考(no-reference)的图像质量量化方法。观感系数根据预先定义的一组图像真实特征的计算方法,对图像进行真实感估计,因此并不需要提供对应的真实图像。越低的观感系数值表示越高的图像观感质量。

3.1 网络训练参数设定

ESRGAN最开始的10个训练周期的学习速率为 2×10^{-4} ,并仅通过均方误差损失项训练生成网络;随后的50个训练周期为基于GAN的训练,并保持 2×10^{-4} 的学习速率,训练50个训练周期;在60个训练周期之后,每隔10个训练周期,减少一半的学习速率,此过程进行80个训练周期。训练过程中生成网络输入的低分辨率图像的大小为 64×64 像素单位,4倍上采样。损失函数为

$$Loss = 1 * VGG54 + 5 \times 10^{-3} * GAN + 1 \times 10^{-2} * L_1 \quad (10)$$

PESR按照 5×10^{-5} 的学习速率训练20个周期,随后学习速率减至 2.5×10^{-5} 。训练周期共计40个。网络输入的LR图像尺寸为 64×64 像素单位,4倍上采样,所采用的损失函数为

$$Loss = 50 * VGG54 + 1 * GAN + 1 * L_1 + 1 \times 10^{-6} * TV \quad (11)$$

EDSR对于超过2倍的上采样,均需要2倍上采样的训练网络。因此,此处训练基于EDSR原始的2倍上采样模型,并进行4倍上采样模型的训练。训练周期共计200个。学习速率设置为 1×10^{-4} 。网络输入的低分辨率图像尺寸为 48×48 像素单位,所采用的损失函数为

$$Loss = 5 * VGG54 + 0.15 * GAN + 0.1 * L_1 \quad (12)$$

上述网络,使用DIV2K所提供的训练数据集进行训练,共计800张高分辨率图像。网络均采用ADAM优化器,参数为 $\beta = (0.9, 0.999)$,并且均为4倍的上采样。

3.2 量化结果的比较

在ESRGAN, PESR, 和EDSR的基础上应用掩膜模块(掩膜模块的应用方法见图1),并对比前后的观感质量变化。4个所采用的测试数据集均在上文有相应的介绍。量化的结果见表1。其中,选取的掩膜模块参数为 $k=0.2, b=0.5$ 。可以观察到应用掩膜模块之后,3个网络的超分辨率结果,均显示出更低的观感系数,即更高的观感质量。观感质量提升的水平和模型有关。PESR在应用掩膜模块后,其超分辨率图像观感质量有着显著的提升,另外2个图像超分辨率模型在应用掩膜模块后,观感质量量化结果也有着不同的提升,并且在4个验证数据集中,均没有出现观感质量指标下降的情况。

表 1 各模型应用掩膜模块前后超分辨率结果观感系数的对比 ($k=0.2, b=0.5$)。

Table 1 Perceptual index comparison before/after applying mask module ($k=0.2, b=0.5$)

模型	DIV2K	Urban100	BSD100	PIRM
ESRGAN	3.02	3.37	2.19	2.1
ESRGAN+Mask	3.02	3.34	2.19	2.07
PESR	3.66	3.81	3.02	2.59
PESR+Mask	3.24	3.58	2.38	2.23
EDSR	3.35	3.54	2.52	2.33
EDSR+Mask	3.24	3.46	2.49	2.33

3.3 观感上的比较

图 4 为 3 个图像超分辨率模型在应用掩膜模块前后的超分辨率结果对比。总体上可以看出,应用掩膜模块后的各个模型的超分辨率结果,在图像观感质量和重建精度上均得到了有效提升。如第一行窗户的纹理,

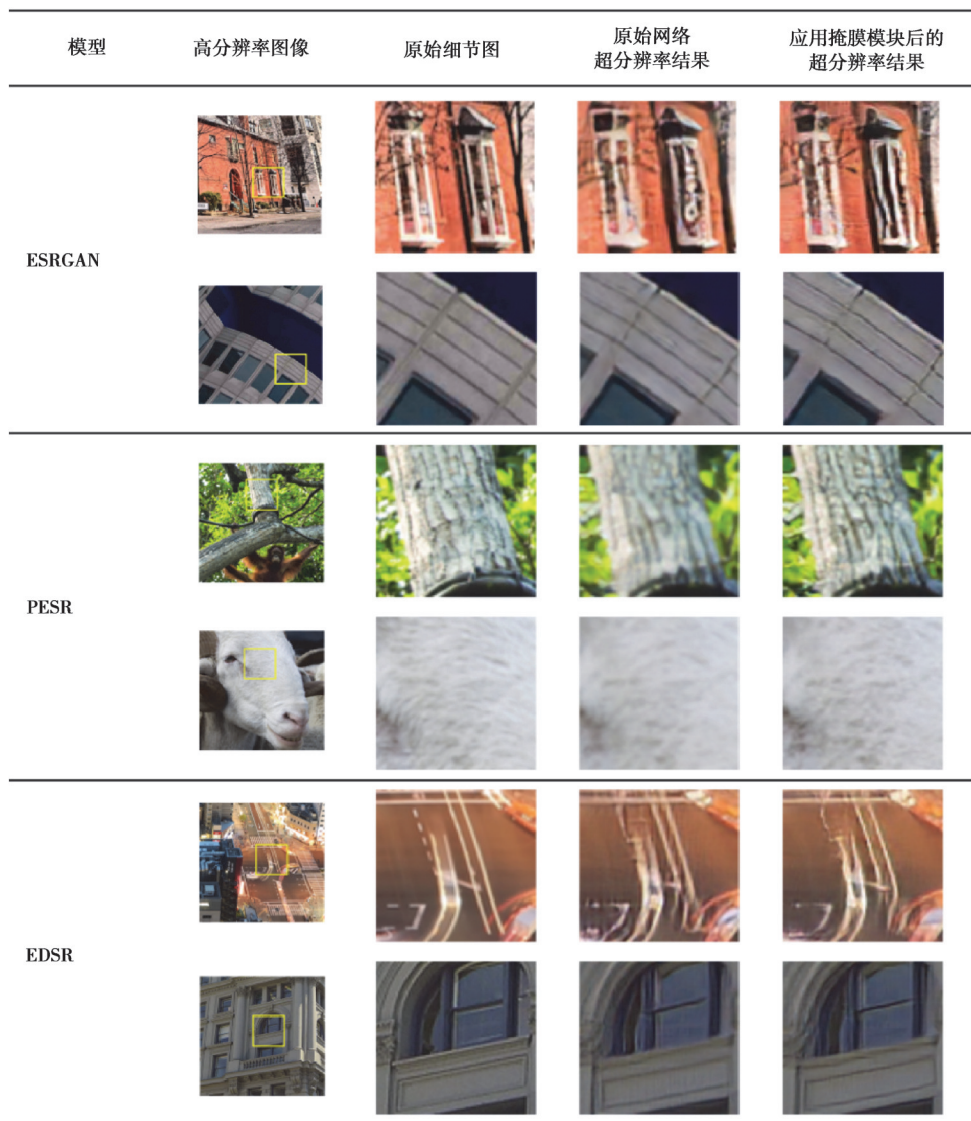


图 4 应用掩膜模块前后各超分辨率网络的超分辨率图像结果对比

Fig. 4 Comparisons of super resolution models' image results before/after applying mask module

原始ESRGAN网络输出图像有着一定的纹理扭曲,应用掩膜模块后,该现象得到了显著的消除;PESR对应的2个超分辨率结果中可以观察到,原图树木纹理和毛发纹理包含较多的高频信息,由于均方误差损失项的作用,原始网络生成的超分辨率结果在观感上呈现模糊的效果,而应用掩膜模块的对应模型,则可以明显地看出图像纹理的清晰度得到了提升。

3.4 掩膜模块参数对超分辨率结果的影响

生成对抗网络的训练结果和初始参数的设定相关,且不容易收敛。通过实验观察掩膜模块参数对整体网络效果的影响,以ESRGAN作为基础网络应用掩膜模块,并测试应用不同掩膜模块参数下的超分辨率结果差异。掩膜模块参数由公式(8)给出。原始ESRGAN以及应用不同掩膜模块参数后的对应超分辨率图像观感质量的量化结果见图5。可以看到,应用掩膜模块的ESRGAN模型,均得到了比原始网络更优的超分辨率图像观感系数;因此掩膜模块具有良好的鲁棒性。应用合理参数的掩膜模块,通常可以得到相较原始网络更优的超分辨率结果。

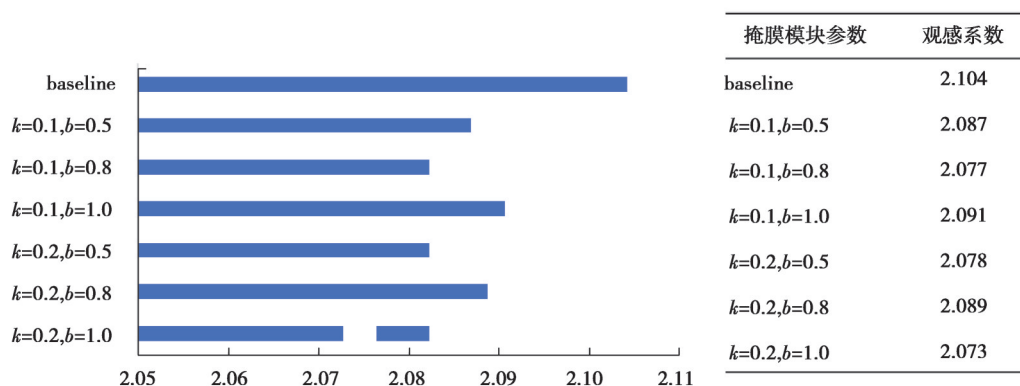


图5 ESRGAN在应用不同掩膜模块参数下的超分辨率图像观感质量对比

Fig. 5 Comparisons of super resolution image perceptual index between different mask module parameters on ESRGAN.

4 结 语

研究提出了掩膜模块用以辅助基于GAN的单张图像超分辨率模型训练,可以实现更好的超分辨率图像观感质量。掩膜由超分辨率结果和原始高分辨率图像计算得到,提供超分辨率图像区域相关的观感质量信息。对抗网络,则根据掩膜对超分辨率图像各区域独立地调整GAN损失权重。实验选取了3个基于GAN的超分辨率模型,通过比较掩膜模块应用前后的超分辨率结果对方法有效性验证。实验结果表明,掩膜模块可以较好地提升基于GAN的超分辨率模型的超分辨率结果观感质量。掩膜模块在辅助对抗网络训练过程当中仅依赖生成网络的超分辨率结果和原始高分辨率图像,因此可添加至现有的很多基于GAN的超分辨率模型当中,以进一步提升相应网络的超分辨率图像结果观感质量。目前,对超分辨率图像区域相关的观感质量信息由相应公式估算得到,如何通过深度学习更有效地计算相关信息,将是下一步研究的方向。

参考文献

- [1] Dong C, Loy C C, He K M, et al. Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(2): 295-307.
- [2] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016, Las Vegas, NV, USA. IEEE, 2016: 1646-1654.
- [3] 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). July 21-26, 2017. Honolulu, HI, USA: IEEE, 2017: 1132-1140.
- [4] Zhang Y L, Tian Y P, Kong Y, et al. Residual dense network for image super-resolution[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018. Salt Lake City, UT, USA: IEEE, 2018: 2472-2481.

- [5] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 27-30, 2016. Las Vegas, NV, USA:IEEE, 2016: 770-778.
- [6] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [7] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). July 21-26, 2017. Honolulu, HI, USA:IEEE, 2017: 105-114.
- [8] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[EB/OL]. 2016: arXiv: 1603.08155. <https://arxiv.org/abs/1603.08155>
- [9] Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [10] Wang X T, Yu K, Wu S X, et al. ESRGAN: enhanced super-resolution generative adversarial networks[C]//European Conference on Computer Vision. Cham: Springer, 2019: 63-79.
- [11] Jolicoeur-Martineau A. The relativistic discriminator: a key element missing from standard GAN[EB/OL]. 2018: arXiv: 1807.00734. <https://arxiv.org/abs/1807.00734>
- [12] Zhang H, Goodfellow I, Metaxas D, et al. Self-attention generative adversarial networks[EB/OL]. 2018: arXiv: 1805.08318. <https://arxiv.org/abs/1805.08318>
- [13] Vu T, Luu T M, Yoo C D. Perception-enhanced image super-resolution via relativistic generative adversarial networks[C]// European Conference on Computer Vision. Cham: Springer, 2019: 98-113.
- [14] Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 7-12, 2015. Boston, MA, USA:IEEE, 2015: 5197-5206.
- [15] Blau Y, Mechrez R, Timofte R, et al. The 2018 PIRM challenge on perceptual image super-resolution[C]//European Conference on Computer Vision. Cham: Springer, 2019: 334-355.
- [16] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]//Proceedings Eighth IEEE International Conference on Computer Vision. ICCV. July 7-14, 2001, Vancouver, BC, Canada. IEEE, 2002: 416-423.
- [17] Agustsson E, Timofte R. NTIRE 2017 challenge on single image super-resolution: dataset and study[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). July 21-26, 2017. Honolulu, HI, USA: IEEE, 2017: 1122-1131.
- [18] Ignatov A, Timofte R, Vu T V, et al. PIRM challenge on perceptual image enhancement on smartphones: report[C]// European Conference on Computer Vision. IEEE, 2018.
- [19] Ma C, Yang C Y, Yang X K, et al. Learning a no-reference quality metric for single-image super-resolution[J]. Computer Vision and Image Understanding, 2017, 158: 1-16.
- [20] Mittal A, Soundararajan R, Bovik A C. Making a “completely blind” image quality analyzer[J]. IEEE Signal Processing Letters, 2013, 20(3): 209-212.

(编辑 侯 湘)