

**计算机视觉工程实践实验报告**

|  |  |
| --- | --- |
| **姓 名:** | 欧阳民 |
| **学 号:** | 123106222853 |
| **学 院:** | 计算机科学与工程学院 |

**2024年 5月 4日**

**一、实验目的**

实现一种图像超分辨率方法在Set5数据集上的测试，得到超分辨率图像，测量其与真实图像之间的PSNR、SSIM指标值。对所选择方法细节进行介绍，并试着讨论该方法可能存在的优缺点，以及可能的改进方向。

**二、实现说明**

**2.1 方法介绍**

本次实验采用的方法是SRGAN实现，下面详细介绍一下SRGAN的内容。SRResNet与SRGAN一块提出，SRResNet算法是一个单模型算法，从图像输入到图像输出中间通过各个卷积模块的操作完成，整个结构比较清晰。但是SRResNet也有不可避免的缺陷，就是它采用了MSE作为最终的目标函数，而这个MSE是直接通过衡量模型输出和真值的像素差异来计算的，SRGAN算法指出，这种目标函数会使得超分重建出的图像过于平滑，尽管PSNR和SSIM值会比较高，但是重建出来的图像并不能够很好的符合人眼主观感受，丢失了细节纹理信息。

**生成对抗网络(GAN)：**

GAN的主要灵感来源于博弈论中博弈的思想，应用到深度学习上来说，就是构造两个深度学习模型：生成网络G（Generator）和判别网络D（Discriminator），然后两个模型不断博弈，进而使G生成逼真的图像，而D具有非常强的判断图像真伪的能力。生成网络和判别网络的主要功能是：

1、G是一个生成式的网络，它通过某种特定的网络结构以及目标函数来生成图像；

2、D是一个判别网络，判别一张图片是不是“真实的”，即判断输入的照片是不是由G生成；

G的作用就是尽可能的生成逼真的图像来迷惑D，使得D判断失败；而D的作用就是尽可能的挖掘G的破绽，来判断图像到底是不是由G生成的“假冒伪劣”。

以上就是GAN算法的原理。运用在图像领域，例如风格迁移，超分重建，图像补全，去噪等，运用GAN可以避免损失函数设计的困难，只要有一个基准，直接加上判别器，剩下的就交给对抗训练。相比其他所有模型, GAN可以产生更加清晰，真实的样本。

**感知损失：**

为了防止重建图像过度平滑，SRGAN重新定义了损失函数，并将其命名为感知损失(Perceptual loss)。感知损失有两部分构成：

感知损失=内容损失+对抗损失

对抗损失就是重建出来的图片被判别器正确判断的损失，这部分内容跟一般的GAN定义相同。SRGAN的一大创新点就是提出了内容损失，SRGAN希望让整个网络在学习的过程中更加关注重建图片和原始图片的语义特征差异，而不是逐个像素之间的颜色亮度差异。以往我们在计算超分重建图像和原始高清图像差异的时候是直接在像素图像上进行比较的，用的MSE准则。SRGAN算法提出者认为这种方式只会过度的让模型去学习这些像素差异，而忽略了重建图像的固有特征。实际的差异计算应该在图像的固有特征上计算。只需要把这些模型中的特征提取模块截取出来，然后去计算重建图像和原始图像的特征，这些特征就是语义特征了，然后再在特征层上进行两幅图像的MSE计算。在众多模型中，SRGAN选用了VGG19模型，其截取的模型命名为truncated\_vgg19。所谓模型截断，也就是只提取原始模型的一部分，然后作为一个新的单独的模型进行使用。

所以内容损失计算方式如下：

1、通过SRResNet模型重建出高清图像SR；

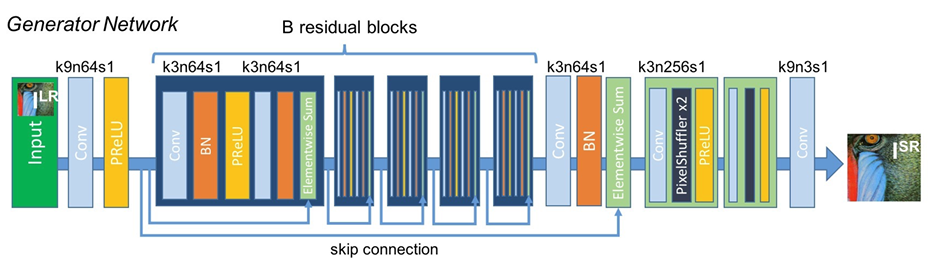
2、通过truncated\_vgg19模型对原始高清图像H和重建出的高清图像SR分别进行计算，得到两幅图像对应的特征图H\_fea和SR\_fea;

3、计算H\_fea和SR\_fea的MSE值；

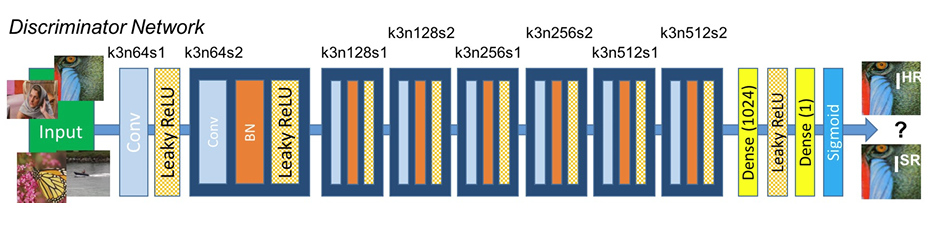
从上述计算方式上看出，原来的计算方式是直接计算H和SR的MSE值，而改用新的内容损失后只需要利用truncated\_vgg19模型对图像多作一次推理得到特征图，再在特征图上进行计算。

**SRGAN结构：**

SRGAN分为两部分：生成器模型和判别器模。



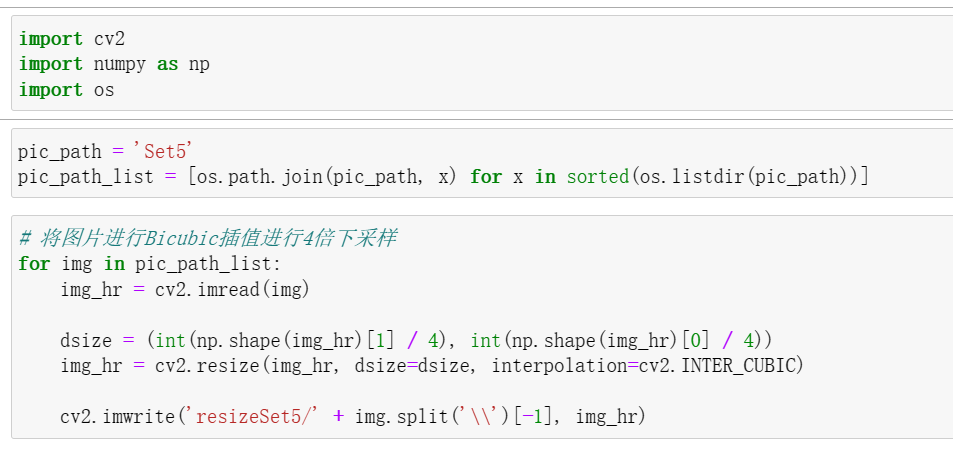
生成器模型采用了SRResNet完全一样的结构，只是在计算损失函数时需要利用截断的VGG19模型进行计算。截断的VGG19模型只是用来计算图像特征，其本身并不作为一个子模块加在生成器后面。可以将此处的VGG19模型理解为静止的（梯度不更新的），只是用它来计算一下特征而已，其使用与一般的图像滤波器sobel、canny算子等类似。



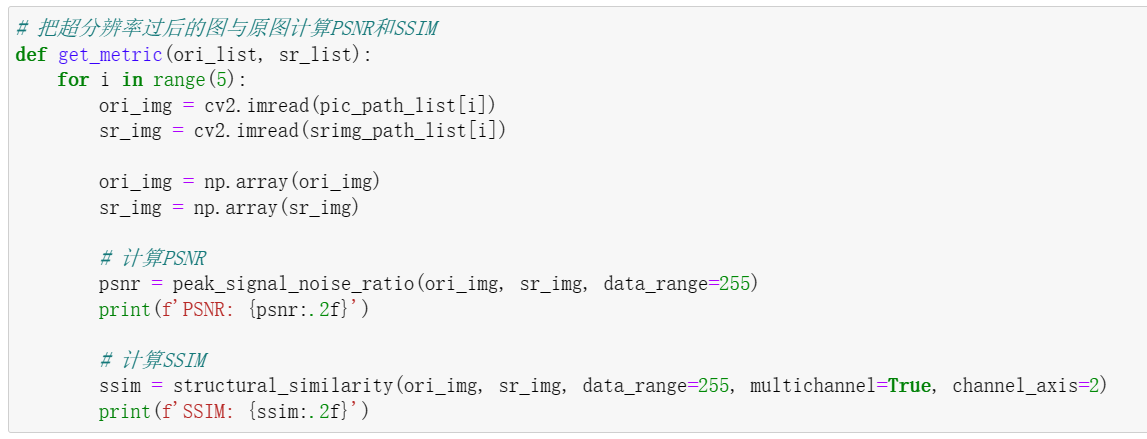
判别器模型对原始高清图像或者重建的高清图像进行判断，判断图像到底是不是生成器创造出来。本质上是一个分类模型，判别器模型中间部分使用了多个卷积模块进行特征提取，最终输出是一个1维的张量。

**2.2实验步骤**

首先将测试图片进行Bicubic插值4倍的下采样。

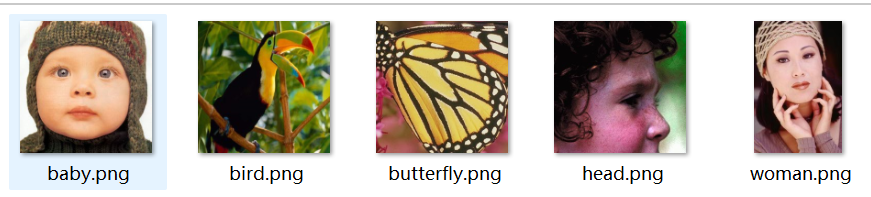


得到缩小四分之一的图像后将图像使用SRGAN超分辨率方法放大四倍，使其与原图像的分辨率相同然后测量PSNR和SSIM指标值。

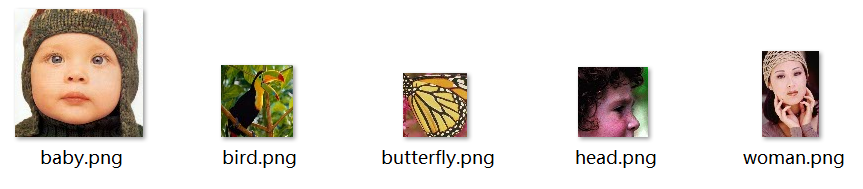


**三、结果截图**

原图：



进行Bicubic插值下采样4倍之后的图：



在进行测试时采用了两个预训练模型一个训练了2000轮一个训练了1000轮。下面展示了从左到右为原图、1000轮预训练模型和2000轮预训练模型效果。

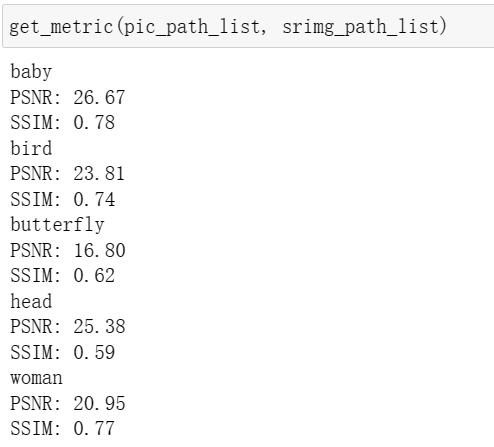
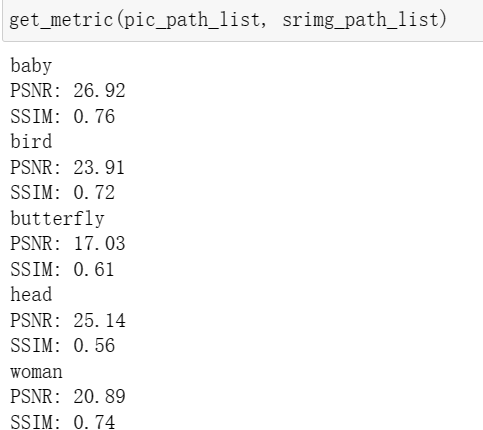
  

计算评价指标：（左图为2000轮，右图为1000轮）

**四、总结与讨论**

本次实验使用了SRGAN超分辨率方法，SRGAN通过生成对抗网络的训练方法，能够生成更加逼真和细节丰富的高分辨率图像。相比传统的插值方法，SRGAN能够产生更高质量的超分辨率图像，更好地保留细节和纹理。SRGAN通过使用SRResNet更深的网络来学习复杂的非线性映射关系。这使得模型能够更好地捕捉输入低分辨率图像和目标高分辨率图像之间的映射，从而提高超分辨率重建的质量。SRGAN的训练框架基于GAN，可以通过对抗训练的方式进行端到端的优化。这种框架具有较好的可训练性和可扩展性，可以应用于不同的超分辨率任务和数据集。

但是SRGAN的训练过程相对复杂，需要同时训练生成器和判别器，并且需要平衡二者之间的对抗性损失和感知损失。这导致了训练过程的相对困难和时间成本较高。在某些情况下，SRGAN生成的超分辨率图像可能存在轻微的模糊或伪影。可能是由于训练数据集的限制或优化目标的不足所致。

**改进方向：**

SRGAN中使用的感知损失通常是基于预训练的VGG网络提取的特征。可以探索引入其他感知损失，如基于对抗性特征学习的感知损失，以进一步提高生成图像的质量。考虑将多尺度超分辨率技术与SRGAN结合，以提高对不同尺度细节的恢复效果。通过在生成器中引入多个子网络，每个子网络负责不同尺度的超分辨率重建，可以获得更全面的超分辨率效果。