**计算机视觉实践-练习3**

1. **SRCNN**

**一、简介**

SRCNN的网络结构仅包含三个卷积层，如图1所示。

图示, 工程绘图

描述已自动生成

图1 SRCNN的网络结构

SRCNN首先使用双三次插值将低分辨率图像放大成目标尺寸，接着通过三层卷积网络拟合非线性映射，最后输出高分辨率图像结果。

具体分为三个步骤：

Patch extraction: 提取图像Patch，进行卷积提取特征，类似于稀疏编码中的将图像patch映射到低分辨率字典中

Non-linear mapping: 将低分辨率的特征映射为高分辨率特征，类似于字典学习中的找到图像patch对应的高分辨字典

Reconstruction：根据高分辨率特征进行图像重建。类似于字典学习中的根据高分辨率字典进行图像重建

**二、实验目标**

实现SRCNN在Set5数据集上的测试，得到超分辨率图像，并进行分析

**三、测试结果**

首先将图像插值一定倍数作为输入图像，然后通过超分辨率网络输出结果。按照输入图像放大的倍数可以分为×2，×3，×4三个网络模型。分别在Set5数据集上对三个网络模型进行测试。

电脑萤幕的截图

描述已自动生成

图2 部分测试数据结果

从图2中可以看出，放大倍数为2倍时，重建效果最好，随着放大倍数的增大，重建效果变差。

原论文中的测试结果和我的测试结果如表1所示，其中好的结果用加粗表示。

表1 PSNR from paper VS My PSNR

|  |  |  |  |
| --- | --- | --- | --- |
| Test Image | Scale | PSNR from paper | My PSNR |
| Baby | ×2 | 38.30 | **41.20** |
| Bird | ×2 | **40.64** | 38.32 |
| Butterfly | ×2 | **32.20** | 29.55 |
| Head | ×2 | 35.64 | **41.27** |
| Woman | ×2 | **34.94** | 34.62 |
| AVE |  | 36.34 | **36.99** |
| Baby | ×3 | 35.01 | **37.89** |
| Bird | ×3 | 34.91 | **35.31** |
| Butterfly | ×3 | **27.58** | 27.53 |
| Head | ×3 | 33.55 | **39.70** |
| Woman | ×3 | 30.90 | **31.72** |
| AVE |  | 32.39 | **34.43** |
| Baby | ×4 | 32.98 | **35.51** |
| Bird | ×4 | 32.98 | **33.52** |
| Butterfly | ×4 | 25.07 | **25.93** |
| Head | ×4 | 32.19 | **38.87** |
| Woman | ×4 | 28.21 | **30.21** |
| AVE |  | 30.09 | **32.81** |

**二、SRGAN**

**一、简介**

SRGAN网络模型以SRResNet作为生成网络用于空间超分辨。和以往的文章以PSNR为导向不同，这是首篇以人类视觉感知为导向进行超分的文章。使用SRGAN进行超分辨，可以获得让人眼感知到高分辨率的图形。

SRGAN本质上提供了一种新的损失函数—感知损失。以往的SR都是由MSE教会网络如何实现LR-HR。但是MSE有着缺陷，即会对图像的细节进行平滑处理，使得重建的图像具有高PSNR，但是失去了人肉眼能够感知的高分辨率感。

图表, 条形图

描述已自动生成

图3 SRGAN的网络结构

SRGAN的网络结构如图1所示，分为生成器和鉴别器。生成器由3个卷积层，16个res block和两个上采样模块组成；鉴别器由9个卷积层组成。

文本

描述已自动生成

图4 鉴别器的网络结构

**二、实验目标**

实现SRGAN在Set5数据集上的测试，得到超分辨图像，并进行分析。

**三、测试结果**

将经过裁剪后的56×56的HR图像进行×4下采样，然后通过网络进行超分辨。部分测试对比结果如图5所示。

SRGAN使用了MOS图像评价指标，实验证明SRGAN在PSNR上达到了SOTA，但在MOS指标上比SRResNet更好。



树枝上有彩色的鸟

中度可信度描述已自动生成

图5 SRGAN在Set5数据集上部分测试结果对比

**四、实验中遇到的问题**

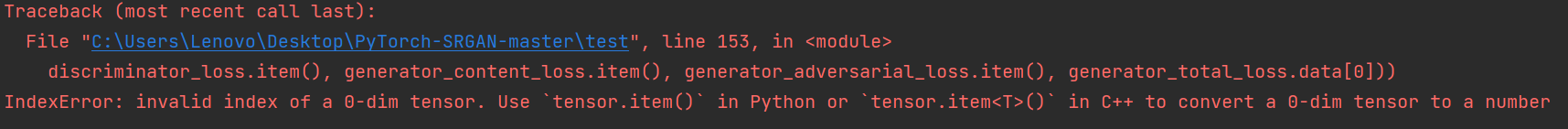
1. 在range中使用“/”，导致参数变成小数

文本

描述已自动生成

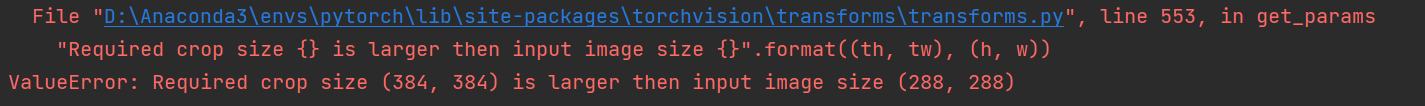
解决方法：这是因为python2中“/”只保留整数部分，而在python3中得到的是一个浮点数。将“/”改成“//”。

2. train\_loss += [loss.data](https://link.zhihu.com/?target=http%3A//loss.data)[0] 是pytorch0.3.1版本代码,在0.4-0.5版本的pytorch会出现警告,不会报错,但是0.5版本以上的pytorch就会报错,总的来说是版本更新问题。



解决方法：将loss.data[0]改为loss.item()

3.SRGAN要求的裁剪图像大小×缩放倍数大于图像原本的大小



解决方法：将裁剪图像的大小改小。

**三、对比分析**







男人张着嘴

描述已自动生成

女人的脸

描述已自动生成

图6 SRCNN和SRGAN在Set5数据集上的超分结果比较

SRCNN使用的损失函数是均方误差，而SRGAN在MSE的基础上添加了对抗损失和感知损失。SRCNN以PSNR为指标进行优化，而SRGAN在优化PSNR的基础上，更加注重人眼对高分辨率图像的感知，使用人为打分作为评价指标。SRCNN的输入是整张原图像，经过bicubic插值放大后放进网络进行重建；而SRGAN是先将图像进行裁剪，放大固定倍数后通过生成器进行重建，再通过鉴别器进行real和fake的判别。

从图6中可以看出，虽然SRCNN的PSNR指标更高，但是图像的模糊程度上来说，SRGAN的效果更好，色彩的对比，区分度更加明显。由于CNN网络优化中对平滑度的优化，图像的细节会趋于模糊，这导致SRCNN重建的图像在视觉效果上不如SRGAN网络重建的图像。