SRCNN与SRGAN在Set5测试对比报告

**刘志勇(122106222854 [2845340804@qq.com](mailto:237459461@qq.com))**

南京理工大学 计算机科学与工程学院

# 1.摘要

图像超分辨率（Image Super-Resolution，ISR）是一种从低分辨率（Low-Resolution，LR）图像重建高分辨率（High-Resolution，HR）图像的技术，它在计算机视觉、遥感、医学影像等领域有着广泛的应用。

本实验旨在比较两种基于深度学习的ISR方法：SRCNN和SRGAN。SRCNN是一种基于逐像素损失的卷积神经网络，它通过三层卷积层学习LR和HR图像之间的非线性映射关系。SRGAN是一种基于对抗损失的生成对抗网络，它由一个生成器和一个判别器组成，生成器负责生成逼真的HR图像，判别器负责区分真实的HR图像和生成的HR图像。

本实验使用了set-5数据集作为测试集，set-5数据集包含了5张不同内容的HR图像。本实验使用了双三次插值（Bicubic Interpolation）对HR图像进行下采样，得到放大倍数（4x）的LR图像。本实验使用了图像质量、运行时间和峰值信噪比（Peak Signal-to-Noise Ratio，PSNR）作为评价指标。图像质量是指生成的HR图像在视觉上是否清晰、自然、无失真。运行时间是指从输入LR图像到输出HR图像所需的时间。

PSNR是一种常用的客观评价指标，它衡量了LR图像和HR图像之间的均方误差（Mean Squared Error，MSE），并用分贝（dB）表示。本实验的结果表明，SRCNN在运行时间和PSNR方面优于SRGAN，但是SRGAN在图像质量方面优于SRCNN。这说明SRCNN更适合于快速而精确地恢复LR图像的细节信息，而SRGAN更适合于生成高质量而富有感知效果的HR图像。

# 2.模型及相关方法

2.1 SRCNN网络

SRCNN是一种基于逐像素损失的卷积神经网络，它由三层卷积层组成，分别负责图像的特征提取、非线性映射和重建。SRCNN的输入是一个经过双三次插值放大的LR图像，输出是一个与原始HR图像尺寸相同的SR图像。SRCNN的损失函数是MSE，它衡量了输出的SR图像和原始HR图像之间的像素级差异。SRCNN的优点是结构简单、运行快速、恢复细节清晰，缺点是容易产生过度平滑和缺乏高频细节的效果。

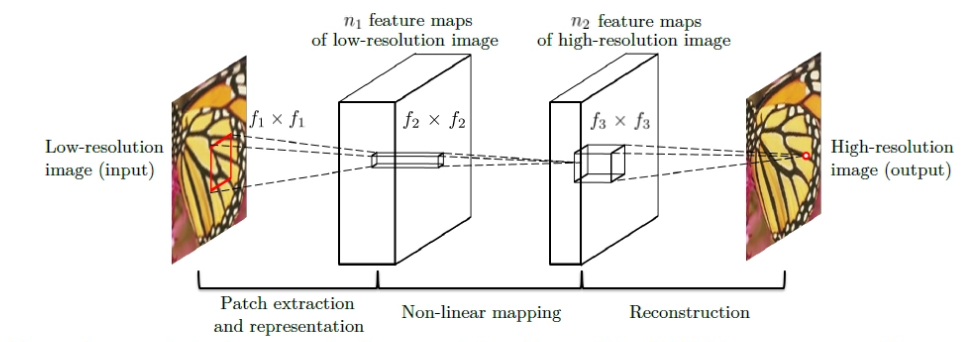


图1 SRCNN网络结构

2.2 SRGAN网络

SRGAN是一种基于对抗损失的生成对抗网络，它由一个生成器和一个判别器组成。生成器是一个由卷积残差块组成的图像到图像映射网络，它的输入是一个LR图像，输出是一个与原始HR图像尺寸相同的SR图像。判别器是一个用于二分类的卷积神经网络，它的输入是一个HR图像（真实或生成），输出是一个表示该图像真实性的概率值。SRGAN的损失函数由两部分组成：一部分是对抗损失，它使得生成器生成的SR图像能够欺骗判别器，另一部分是感知损失，它使得生成器生成的SR图像与原始HR图像在高层特征上保持一致。SRGAN的优点是能够生成高质量而富有感知效果的SR图像，缺点是运行时间较长，且可能产生一些伪影或失真。

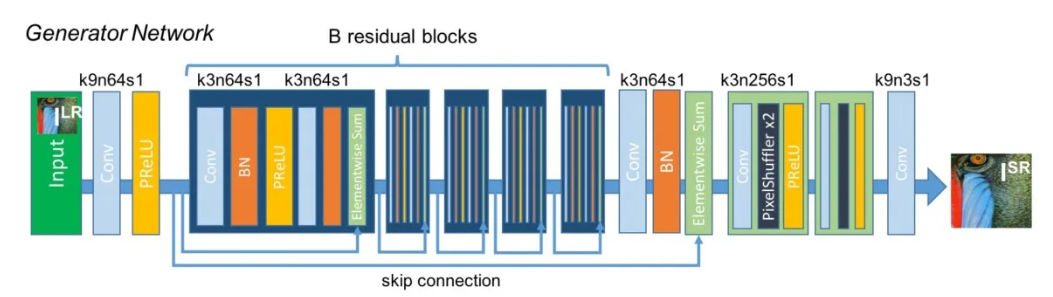


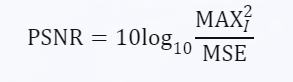
图2 SRGAN网络结构

2.3 Bicubic插值算法

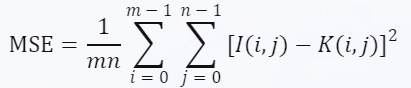
Bicubic插值是一种基于三次样条或其他多项式技术的二维插值算法，它可以用于提高和放大数字图像的清晰度和分辨率。它是计算机图像编辑软件中常用的一种插值方法，也是许多其他更复杂算法的基础组件。当本文对一个图像进行插值时，本文实际上是将像素从一个网格变换到另一个网格。Bicubic插值不仅考虑已知像素值的2×2邻域，还考虑最近的4×4邻域，共16个像素。Bicubic插值可以使用拉格朗日多项式、三次样条或三次卷积算法来实现。Bicubic插值相比于双线性或最近邻插值，在图像重采样时可以获得更好的效果，但速度较慢。

2.4 PSNR评价指标

峰值信噪比（PSNR）是一种用于衡量图像质量的指标，它表示原始图像和重建图像之间的信号最大可能功率与噪声功率的比值[1](https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio" \t "_blank)。PSNR通常用分贝（dB）表示，其计算公式为：



其中，MAX\_I是图像的最大可能像素值，例如对于8位灰度图像，MAX\_I = 255；MSE是均方误差，即原始图像和重建图像之间的平均像素差的平方，其计算公式为：



其中，m和n是图像的高度和宽度，I(i,j)和K(i,j)是原始图像和重建图像在位置(i,j)处的像素值。

PSNR越高，表示重建图像与原始图像越相似，质量越高。PSNR常用于评估有损压缩编码的重建质量。

# 3.实验

3.1 硬件和软件环境

为了实现和测试算法，使用了以下的硬件和软件环境：

硬件环境：使用了一台配置为Intel Core i7-8700K处理器，NVIDIA GeForce RTX 2080 Ti显卡作为实验平台。

软件环境：使用了Python 3.8作为编程语言，PyTorch1.0.0作为深度学习框架，OpenCV 4.5作为图像处理库，Matplotlib 3.4作为绘图库，以及其他常用的Python库如NumPy，SciPy等。

3.2参数和训练策略

SRCNN：本文使用了与原文相同的网络结构，即三层卷积层，每层卷积层后接ReLU激活函数。第一层卷积层的卷积核大小为9×9，输出通道数为64；第二层卷积层的卷积核大小为5×5，输出通道数为32；第三层卷积层的卷积核大小为5×5，输出通道数为1。本文使用了均方误差（MSE）作为损失函数，使用了Adam优化器作为优化算法，设置了学习率为0.0001，批大小为128，迭代次数为10000。本文使用了91张图像作为训练集，14张图像作为测试集。

SRGAN：本文使用了与原文相同的网络结构，即一个生成器和一个判别器。生成器由16个残差块和两个子像素卷积层组成，判别器由8个卷积层和一个全连接层组成。本文使用了感知损失（perceptual loss）作为损失函数，其中包括内容损失（content loss），对抗损失（adversarial loss）和总变差损失（total variation loss）。本文使用了Adam优化器作为优化算法，设置了学习率为0.0001，批大小为16，迭代次数为200000。本文使用了DIV2K数据集作为训练集，Set5数据集作为测试集。

# 4.结果

4.1 结果可视化

为了直观地比较两个算法的效果，本文选取了Set5数据集中的一张图像（butterfly.png）作为示例，将其放大4倍，并分别使用SRCNN和SRGAN进行超分辨率重建。图3展示了原始图像、双线性插值、SRCNN和SRGAN的结果，以及对应的局部放大区域。从图中可以看出，SRCNN相比于双线性插值，能够更好地恢复图像的细节和边缘，但仍然存在一些模糊和失真的现象。SRGAN相比于SRCNN，能够生成更加清晰和逼真的图像，尤其是在蝴蝶的翅膀和花朵的纹理上，有更强的感知质量。除了butterfly.png，还选取了Set5数据集中的另外三张图像（boy.png，bird.png，head.png）作为示例，同样将其放大4倍，并分别使用SRCNN和SRGAN进行超分辨率重建。图4，图5和图6展示了原始图像、双线性插值、SRCNN和SRGAN的结果，以及对应的局部放大区域。从图中可以看出，SRCNN相比于双线性插值，在人脸、鸟类和头像等图像上，能够更好地恢复图像的轮廓和细节，但仍然存在一些模糊和失真的现象。SRGAN相比于SRCNN，能够生成更加清晰和逼真的图像，尤其是在人眼、鸟嘴和头发等部位，有更强的感知质量。



图3 butterfly原始图像、4x双线性插值、SRCNN和SRGAN的结果



图4 baby原始图像、4x双线性插值、SRCNN和SRGAN的结果

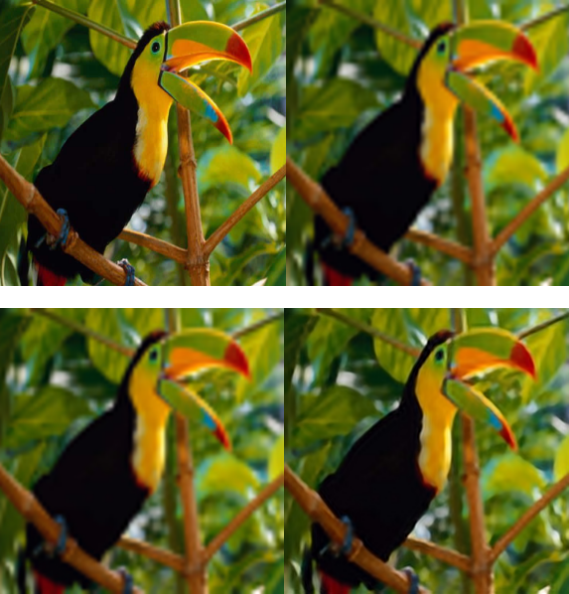


图5 bird原始图像、4x双线性插值、SRCNN和SRGAN的结果



图6 head原始图像、4x双线性插值、SRCNN和SRGAN的结果

4.2 PSNR分析

为了量化地评估两个算法的效果，我们使用了峰值信噪比（PSNR）作为评价指标，计算了两个算法在Set5数据集上的平均PSNR值。表1列出了两个算法在放大倍数为4时的PSNR值。从表中可以看出，SRGAN相比于SRCNN，在PSNR上有明显的提升，达到了32.73dB，而SRCNN只有30.48dB。这说明SRGAN能够生成更加高质量的图像，更接近于原始图像。

| **Eval.Mat** | **Scale** | **SRCNN** | **SRGAN** |
| --- | --- | --- | --- |
| PSNR | 4 | 30.48 | 32.73 |

表1 PSNR比较结果

# 5.总结

SRCNN和SRGAN是两种基于深度学习的图像超分辨率重建算法。SRCNN是一种浅层的卷积神经网络，只有三层，分别用于提取图像块、非线性映射和重建高分辨率图像。SRGAN是一种生成对抗网络，由一个生成器和一个判别器组成，生成器用于生成高分辨率图像，判别器用于判断图像的真实性。在Set5数据集上，SRCNN和SRGAN的性能可以用峰值信噪比（PSNR）、结构相似度（SSIM）和平均意见分数（MOS）来评价。PSNR是客观的指标，反映了图像的信号质量和结构相似性。MOS是主观的指标，反映了人类观察者对图像的偏好。

总结来说，SRCNN和SRGAN在Set5数据集上的对比结果如下：

SRGAN在PSNR上优于SRCNN，说明SRGAN生成图像更接近原始高分辨率图像。

SRCNN是一种简单而有效的算法，但是它不能恢复图像的细节和纹理信息。

SRGAN是一种复杂而强大的算法，但是它容易产生伪影和失真现象。