# Set5数据集对SWinIR方法的测试

## 一、图像超分辨率方法

图像超分辨率是一种通过使用计算机视觉和图像处理技术，将低分辨率图像转换为高分辨率图像的方法。在图像超分辨率中，我们旨在从给定的低分辨率图像中恢复出相应的高分辨率图像，以改善图像的清晰度、细节和质量。以下是几种常见的图像超分辨率方法：

1. 插值方法（Interpolation-based methods）：

最简单的超分辨率方法之一是插值方法，如双线性插值、双三次插值等。这些方法基于对低分辨率图像进行插值来生成高分辨率图像。该方法简单易实现，计算速度快。但是生成的高分辨率图像缺乏真实细节，容易出现伪影和模糊。

1. 基于重建的方法（Reconstruction-based methods）：

基于重建的方法试图恢复出高分辨率图像的细节信息。常见的方法包括最小二乘法、正则化方法等。这种方法能够利用图像统计信息进行更准确的重建，可以获得较好的视觉效果。但是计算复杂度较高，对参数设置敏感，容易过拟合。

3. 基于学习的方法（Learning-based methods）：

基于深度学习的方法在图像超分辨率任务中取得了很大成功。通过训练深度神经网络来学习从低分辨率到高分辨率的映射关系，如SRCNN、ESPCN等。深度学习方法能够学习到更复杂的特征映射，生成质量更高的高分辨率图像。

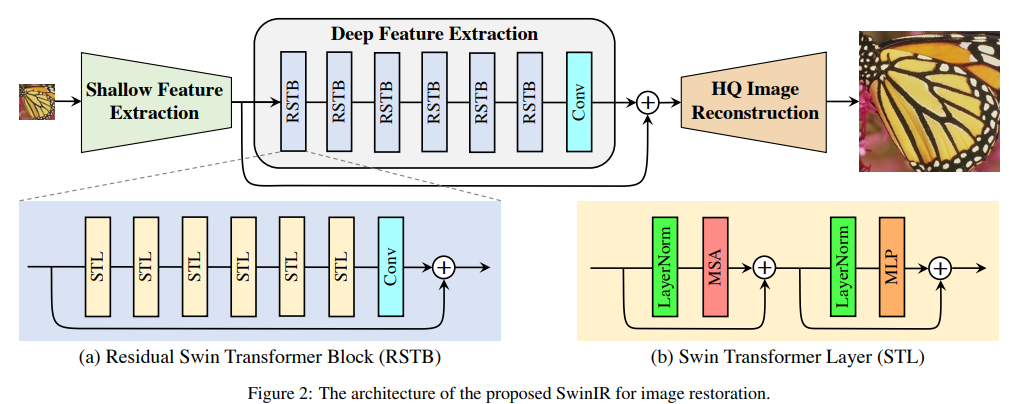
4. 基于GAN的方法（GAN-based methods）：

基于生成对抗网络（Generative Adversarial Network, GAN）的方法结合了生成器和判别器网络，能够生成更加逼真的高分辨率图像。对抗损失能够帮助生成器产生更加真实的结果。同时，GAN训练过程不稳定，容易发生模式崩溃、训练不收敛等问题，需要精细调参。

基于上述各种方法的特点，本次实验使用基于学习的方法**SWinIR** ，SWinIR是一种基于弱监督学习的图像恢复方法，在没有大规模带注释数据的情况下，通过设计有效的损失函数和训练策略，利用少量的标记数据或者具有较低质量标签的数据进行模型训练。

## 二、SWinIR 网络结构

### 2.1 整体框架

SwinIR 的网络结构主要分为 3 个部分，分别是浅层[特征提取](https://so.csdn.net/so/search?q=%E7%89%B9%E5%BE%81%E6%8F%90%E5%8F%96&spm=1001.2101.3001.7020" \t "https://blog.csdn.net/Wenyuanbo/article/details/_blank)模块，深层特征提取模块和高质量图像重建模块。

### 2.2 浅层特征提取

浅层特征提取只使用一层卷积进行提取。

### 2.3 深层特征提取

深层特征提取模块由若干个残差 SwInTransformer 块 (RSTB) 和卷积块构成。

(1) 首先将来自浅层特征提取模块的特征图分割成多个不重叠的 patch embeddings；

(2) 再通过多个串联的残差 SWin Transformer 块 (RSTB);

(3) 将多个不重叠的 patch embeddings 重新组合成与输入特征图分辨率一样；

(4) 最后通过一个卷积层 (1 层或3 层卷积) 输出；

(5) 在每个 RSTB 中都引入残差连接。

|  |  |
| --- | --- |
|  |  |

残差 SwInTransformer 块 (RSTB) 中的 STL 就是 SwIn Transformer Layer 的意思，具体结构如下图。

(1) 首先通过一个归一化层 LayerNorm；

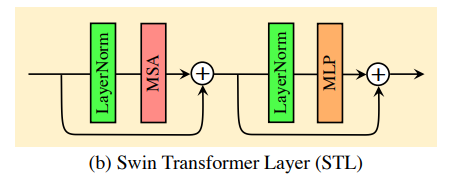
(2) 再通过多头自注意力 (MSA) 模块；

(3) 在多头自注意力结尾引入残差；

(4) 再通过一个归一化层 LayerNorm；

(5) 最后通过一个多层感知机 MLP；

(6) 结尾同样引入残差。



### 2.4 图像重建模块

图像重建模块其实就是卷积+上采样的组合，包含 4 种结构。

1. 经典超分 (卷积 + pixelshuffle 上采样 + 卷积):

优点：结构简单，易于实现和训练。使用卷积和pixelshuffle上采样可以有效提高图像分辨率。

缺点：对于复杂纹理和细节的重建效果可能不理想，在处理超高分辨率图像时可能受限。

2. 轻量超分 (卷积 + pixelshuffle 上采样):

优点：简化了经典超分模型，减少了计算和参数量。使用卷积和pixelshuffle上采样可以获得一定的图像增强效果。

缺点：相对于更复杂的模型，重建质量可能有所降低。对于一些复杂场景或要求更高的任务可能不够强大。

3. 真实图像超分 (卷积 + 卷积插值上采样 + 卷积插值上采样 + 卷积):

优点：通过多次卷积插值上采样可以有效地提高图像分辨率。可以应对一些复杂的场景和真实图像的重建需求。

缺点：模型相对较复杂，需要更多的计算资源和时间来训练和推理。可能存在过拟合问题，需要适当的正则化和控制。

4. 像去噪和JPEG压缩去伪影 (卷积 + 引入残差):

优点：结合了去噪和去JPEG压缩伪影的功能，可以提供更好的视觉质量，引入残差结构可以帮助保留更多的图像细节。

缺点：需要更多的训练数据和计算资源来学习和推理，对于其他类型的失真或噪声可能不适用。

## 三、实验

实验使用经典超分结构，用[DIV2K](https://cv.snu.ac.kr/research/EDSR/DIV2K.tar) (800 training images) 数据集训练的模型来对Set5数据集进行测试。

测试方式:先将图像用Bicubic插值进行4倍下采样，再使用超分辨率算法处理，将得到的超分辨率图像与真实的原始图像进行对比，使用PSNR以及SSIM进行评估。

Set5数据集：Bicubic插值处理后经过算法处理的图像与原图如下所示：

|  |  |
| --- | --- |
|  |  |
|  |  |
|  | |

测试得到的评估结果如下表：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Set5 | baby | bird | butterfly | head | woman | Averge |
| PSNR | 32.58 | 33.33 | 28.31 | 30.12 | 30.10 | 30.89 |
| SSIM | 0.88 | 0.94 | 0.92 | 0.72 | 0.92 | 0.87 |

## 四、总结

根据实验结果可以发现，即使使用较小的训练集，对Set5数据集进行测试得到的结果依旧是较好的，PSNR基本在30以上，SSIM也是0.9左右。这说明SWinIR作为一种弱监督学习方法，可以在数据有限的情况下仍然能够取得不错的性能表现。对比原图以及上采样重建后的图像，我们也能发现SWinIR具有较好的重建效果，在保持图像细节和纹理方面能够取得较好的性能。

同时，本次实验的各种参数设置等等都是采用默认的原设置，且只使用了经典超分结构，实际上共有四种结构，根据具体实验任务需要自己选择，这需要耗费一定的时间和精力，这说明了是否可以通过某个结构让模型自已判断从而进行合适的选择，这对于未来用户的使用可以提供很大便利。