

学号：2021522077 | 班级：2021669101 | 课程：智能信息网络实验

组长：巩羽飞

组员：廉涟，孟祥超

题目：FinalLab 图像超分辨率

源代码见文末，算法参考自：<https://arxiv.org/pdf/2107.10833>

一、问题描述

1.1 待解决问题的解释

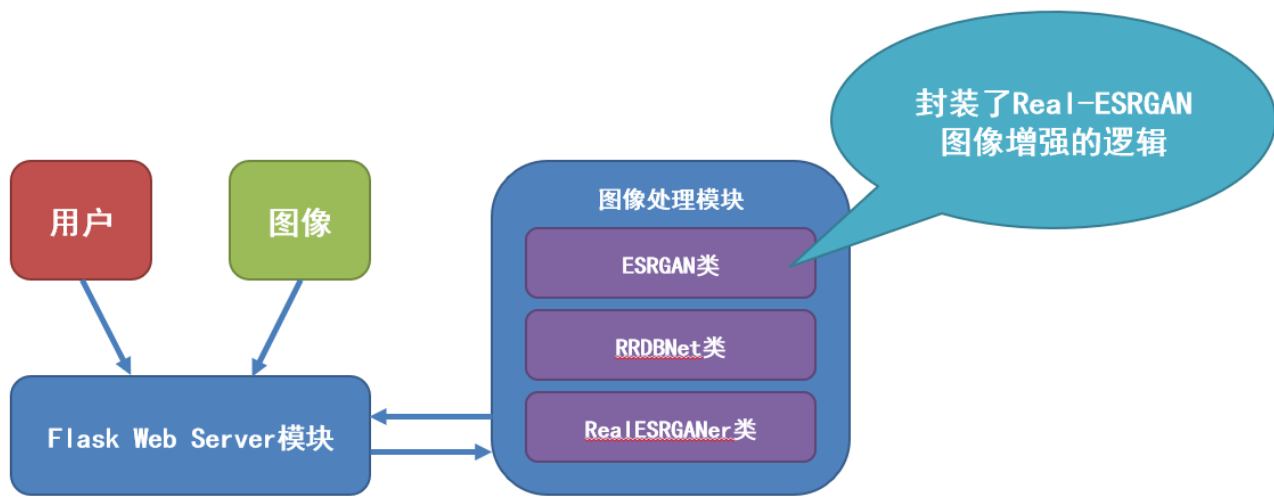
单幅图像超分辨率（Image Super Resolution）是一种图像处理技术，旨在从低分辨率图像生成高分辨率图像。这一过程在图像增强、监控视频分析、医学图像诊断等多个领域有广泛的应用。其主要目的是通过提升图像的细节和质量，提供更准确的分析结果。

1.2 问题的形式化描述

给定一张低分辨率图像  $I_{LR}$ ，目标是通过某种算法生成高分辨率图像  $H_{LR}$ ，使得在视觉效果和质量上尽可能接近真实高分辨率图像。

二、系统

2.1 系统架构



系统架构主要包括以下几个模块：

- 用户界面：用户通过Web界面上传图像并查看处理结果。
- Flask Web服务器模块：负责处理用户请求和响应。
- 图像处理模块：包含具体的图像超分辨率算法，实现图像的超分辨率处理。
- ESRGAN类：使用增强的超分辨率生成对抗网络（ESRGAN）进行处理。
- RRDBNet类：残差密集块网络，用于图像特征提取和重建。
- RealESRGANer类：包含Real-ESRGAN的具体实现，用于图像增强。

## 2.2 各部分介绍

- **用户界面**: 用户可以通过一个简洁的Web界面上传低分辨率图像, 并查看高分辨率图像的生成结果。
- **Flask Web服务器模块**: 负责处理前端请求, 调用图像处理模块, 并返回处理后的高分辨率图像。
- **图像处理模块**: 封装了图像超分辨率的逻辑, 包括图像的预处理、超分辨率算法的调用以及结果的后处理。
- **ESRGAN类**: 使用ESRGAN算法进行图像超分辨率处理, 该算法通过生成对抗网络提高图像的细节和质量。
- **RRDBNet类**: 包含多个残差密集块, 用于有效地提取和重建图像特征。
- **RealESRGANer类**: 基于Real-ESRGAN算法, 利用二阶退化过程和sinc滤波器模拟更真实的退化现象, 生成高质量的图像。

### 2.2.1 二阶退化过程

二阶退化过程是一种用于模拟图像退化的复杂过程, 以更真实地反映现实世界中图像质量下降的情况。具体而言, Real-ESRGAN采用二阶退化模型, 该模型包含多个经典退化步骤, 如模糊、噪声添加和下采样。这些步骤可以组合应用, 以生成更加逼真的低分辨率图像。此外, 使用sinc滤波器来合成公共信号振铃和超调伪影, 从而更准确地模拟实际拍摄条件下图像的劣化现象。这种处理方法能够在训练过程中帮助模型更好地学习处理各种复杂的退化模式, 提高生成的高分辨率图像的质量。

### 2.2.2 Generator

Generator 是 GAN 中的一个核心组件, 其主要任务是从输入的低分辨率图像生成高分辨率图像。在Real-ESRGAN中, Generator采用了与ESRGAN相同的结构, 即使用了RRDBNet。这种网络结构通过多层残差密集块, 能够有效地捕获图像的细节信息和特征, 并进行重建。具体而言, 对于较高的放大比例, Generator会利用像素拆分 (Pixel Unshuffle) 操作, 将空间分辨率降低, 并将信息重新排列到通道维度, 以便更好地处理和生成高分辨率图像。

### 2.2.3 Discriminator

Discriminator 是 GAN 的另一个关键组件, 其主要任务是区分生成的高分辨率图像和真实的高分辨率图像。在Real-ESRGAN中, 判别器采用了U-Net结构, 而不是传统的VGG网络结构。这一改进使得判别器能够更好地捕捉全局样式和局部纹理信息。为了提高输出的稳定性, 判别器中还加入了谱归一化 (Spectral Normalization) 技术。通过这种方式, 判别器不仅能够提供更精确的反馈, 帮助生成器生成更逼真的图像, 还能有效防止模型训练过程中的不稳定性。

## 2.3 算法的伪代码

```
def super_resolution(image):  
    # Step 1: Pre-process the input image  
    lr_image = preprocess(image)  
  
    # Step 2: Use the Real-ESRGAN model to enhance the image  
    sr_image = real_esrgan_enhance(lr_image)
```

```
# Step 3: Post-process the enhanced image
hr_image = postprocess(sr_image)

return hr_image

def preprocess(image):
    ...
    return processed_image

def real_esrgan_enhance(image):
    ...
    rrdnet = RRDBNet()
    enhanced_image = rrdnet.forward(image)
    return enhanced_image

def postprocess(image):
    ...
    return final_image
```

## 三、实验

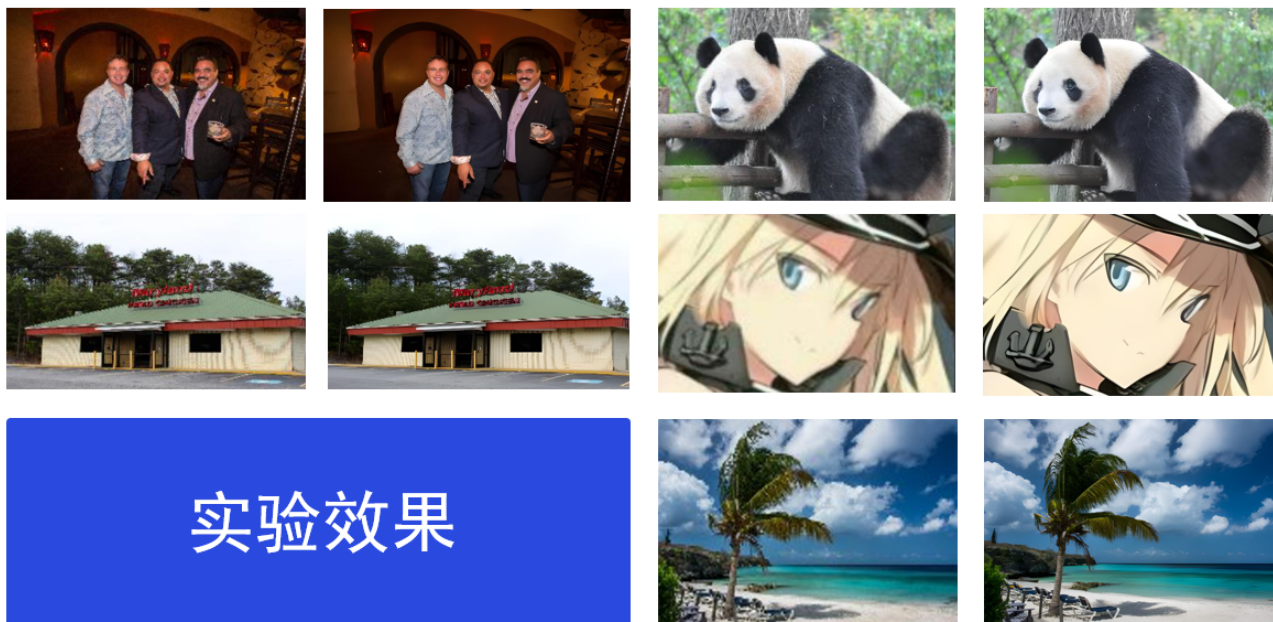
### 3.1 实验环境

- 操作系统：Windows / Linux
- 编程语言：Python 3.10
- 硬件：NVIDIA GPU
- 深度学习框架：PyTorch 2.0.1
- 图像处理库：OpenCV 4.8.0.74, PIL 10.0.0, NumPy 1.24.4
- Web框架：Flask 2.3.2

### 3.2 数据

- **DIV2K**: 包含1000张不同场景的高分辨率图像，数据划分为训练集、验证集和测试集，比例为8:1:1。
- **Flickr2K**: 包含2650对图像，具有较高的多样性。
- **OutdoorSceneTraining**: 针对户外场景，包含2687对图像，模拟真实世界中的各种情况。

### 3.3 实验结果



## 实验效果

实验结果表明，采用Real-ESRGAN算法能够在大多数情况下显著提升图像的分辨率和质量。具体效果可以通过以下方式评估：

- **主观评价：**通过肉眼观察图像细节的提升和噪声的减少。
- **数学指标：**使用峰值信噪比（PSNR）和结构相似性指数（SSIM）等指标进行定量评价。

该方法在处理复杂纹理和真实场景时表现出色，但在某些细微纹理和特定室内场景的超分辨率效果上还有待提升。

## 四、总结与展望

本实验采用的Real-ESRGAN算法通过引入高阶退化建模过程，能够更好地模拟复杂的现实世界退化现象，并使用sinc滤波器模拟常见的振铃和伪影。尽管在大多数场景中效果显著，但对于一些特定场景和细微纹理的处理仍有提升空间。未来工作将重点在于优化模型架构和丰富训练数据集，以进一步提升模型的泛化能力和细节重建效果。

## 五、参考文献

1. X. Wang, L. Xie, C. Dong and Y. Shan, "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data," 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 2021, pp. 1905-1914, doi: 10.1109/ICCVW54120.2021.00217.  
keywords: {Training;Degradation;Visualization;Computer vision;Conferences;Computational modeling;Superresolution}
2. Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., and Change Loy, C. (2018). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 63-79. DOI: [10.1007/978-3-030-01234-2\\_5](https://doi.org/10.1007/978-3-030-01234-2_5).
3. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., and Shi, W. (2017). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In

*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4681–4690. DOI: [10.1109/CVPR.2017.19](https://doi.org/10.1109/CVPR.2017.19).

4. Lim, B., Son, S., Kim, H., Nah, S., and Mu Lee, K. (2017). Enhanced Deep Residual Networks for Single Image Super-Resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 136–144. DOI: [10.1109/CVPRW.2017.151](https://doi.org/10.1109/CVPRW.2017.151).
5. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., and Fu, Y. (2018). Residual Dense Network for Image Super-Resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2472–2481. DOI: [10.1109/CVPR.2018.00262](https://doi.org/10.1109/CVPR.2018.00262).
6. Haris, M., Shakhnarovich, G., and Ukita, N. (2018). Deep Back-Projection Networks for Super-Resolution. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1664–1673. DOI: [10.1109/CVPR.2018.00179](https://doi.org/10.1109/CVPR.2018.00179).