

# 计算机视觉实践报告

## 目录

- 一、SRCNN 算法 ..... 1
  - 1.前言 ..... 1
  - 2.模型结构..... 1
  - 3.损失函数..... 2
  - 4.对比实验..... 2
- 二、SRGAN 算法 ..... 4
  - 1.前言 ..... 4
  - 2.模型结构..... 4
  - 3.损失函数..... 4
  - 4.对比实验..... 5

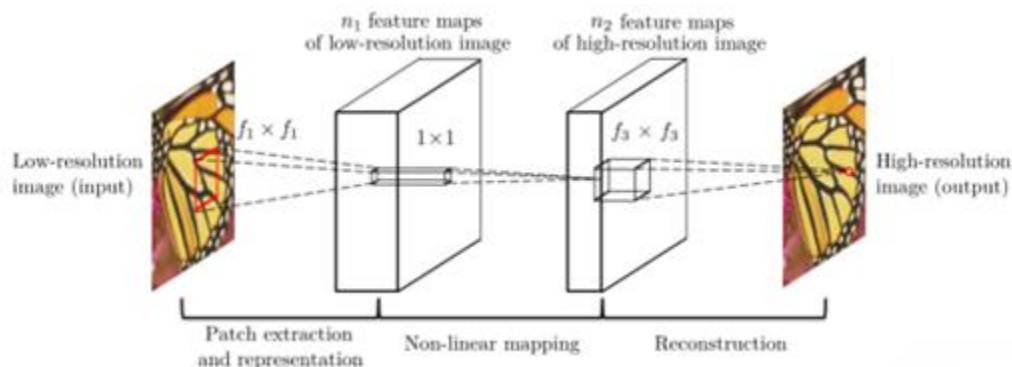
## 一、SRCNN 算法

### 1.前言

超分辨率技术 (Super Resolution, SR)是指从观测到的低分辨率图像重建出相应的高分辨率图像，在监控设备、卫星图像和医学影像等领域都拥有着重要的应用价值。

### 2.模型结构

**SRCNN** 是深度学习用在超分辨率重建上的开山之作。其结构十分简单，仅仅只用了三个卷积层，结构如下：



首先说明以下符号的含义：

- $Y$ : 输入图像经过预处理(双三次插值)得到的图像, 我们仍将  $Y$  当作是低分辨率图像, 但它的 size 要比输入图像要大。
- $F(Y)$ : 网络最后输出的图像, 我们的目标就是通过优化  $F(Y)$  和 Ground-Truth 之间的 loss 来学会这个函数  $F(\cdot)$ 。
- $X$ : 高分辨率图像, 即 Ground-Truth, 它和  $Y$  的 size 是相同的。

如上图所示是 SRCNN 的网络模型, 其分为三部分, 分别是:

- Patch extraction and representation(其实就是图像特征提取层)。通过 CNN 将图像  $Y$  的特征提取出来存到向量中, 这个向量里包含了多张 feature map, 即一张图所含的一些特征。
- 非线性映射层。将上一层的 feature map 进一步做非线性映射处理, 使得网络深度加大, 更有利于学到东西。
- 网络重建层。重建用于将 feature map 进行还原成高分辨率图像  $F(Y)$ , 其与  $X$  做 loss 并通过反传来学习整个模型的参数。

### 3.损失函数

设 batchsize 为  $n$ , SRCNN 网络参数集为  $\Theta = \{W1, W2, W3, B1, B2, B3\}$ , 则 Loss function 可定义为:

$$L(\Theta) = \frac{1}{n} \sum_{i=1}^n \|F(Y_i; \Theta) - X_i\|^2.$$

### 4.对比实验

模型经过 5 次训练后得到模型, 并在 Set5 数据集上进行验证。  
原始图像:



**Bicubic 下采样: (PSNR:30.45)**



**SRCNN 处理后: (PSNR:35.33)**



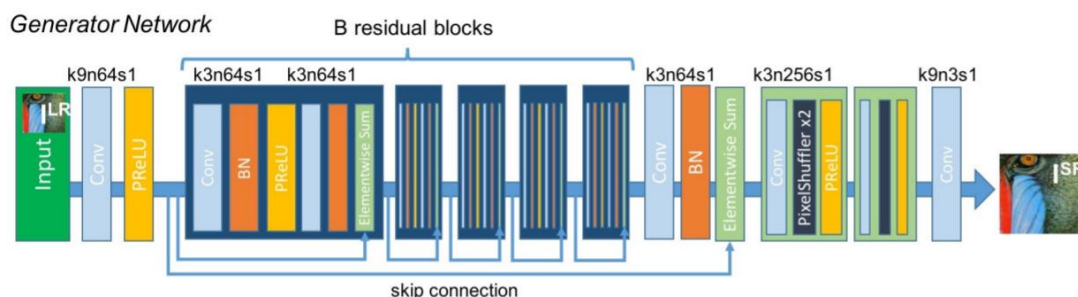
## 二、SRGAN 算法

### 1.前言

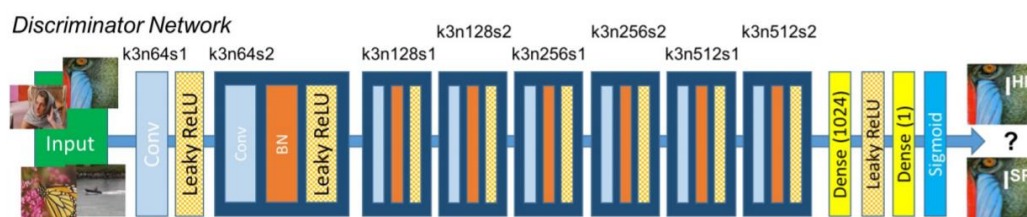
生成对抗网络(GAN)是一类非常有趣的神经网络。借助 GAN，计算机能够生成逼真的图片。近年来有许多“AI 绘画”的新闻，这些应用大多是通过 GAN 实现的。实际上，GAN 不仅能做图像生成，还能辅助其他输入信息不足的视觉任务。比如 SRGAN，就是把 GAN 应用在超分辨率(SR)任务上的代表之作。

### 2.模型结构

和其他对抗生成网络一样，SRGAN 有生成网络和判别网络，我们先看生成网络：



其中  $k$  代表卷积核的尺寸， $n$  代表卷积输出的通道数， $s$  代表步长，不同指向的箭头表示残差结构，Elementwise Sum 就是残差中相加的操作。相同颜色表示相同的操作，低分辨率图片 (LR) 输入网络后输出高分辨率图片 (HR)。下面来看判别网络：



判别网络没有残差结构，其中的符号表示的意思和上面解释的一样，判别网络输入一张图片，判断这张图片是原始高分辨率的图片还是生成网络输出的高分辨率图片。

### 3.损失函数

SRGAN 生成的网络损失函数为感知损失，由两部分组成 **content loss**，和 **adversarial loss**

$$l^{SR} = \underbrace{l_X^{SR}}_{\text{content loss}} + \underbrace{10^{-3} l_{Gen}^{SR}}_{\text{adversarial loss}}$$

perceptual loss (for VGG based content losses)

**content loss** 是生成的 HR 和真实的 HR 通过 VGG 网络前 16 层得到的特征之间的 MSE 损失，可以表示为：

$$\text{Content loss} = \text{MSE}(\text{VGG}(G(LR)), \text{VGG}(HR)).$$

**adversarial loss:**

$$l_{Gen}^{SR} = \sum_{n=1}^N -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

$$D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

其中的  $D_{\theta_D}(G_{\theta_G}(I^{LR}))$  表示判别器判断生成图片为真实的高分辨率图片的概率。

## 4.对比实验

原始图像：



Bicubic 下采样: (PSNR:21.59)



SRGAN 处理后: (PSNR:21.15)

