## 基于SRCNN的图像超分辨率实验报告

**姓名：卓威 学号：123106222860**

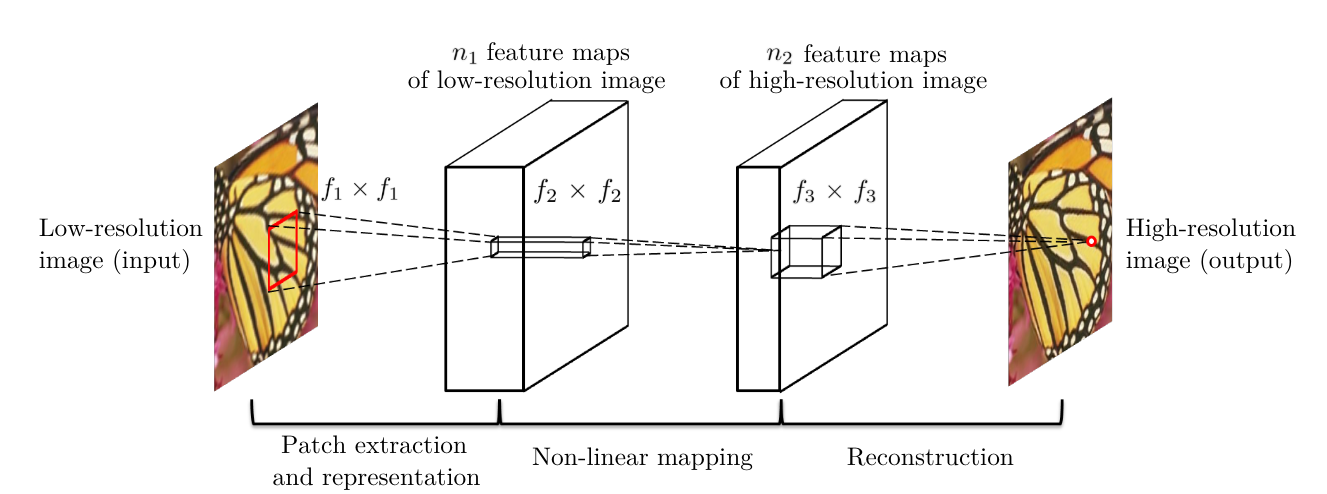
#### **一、**超分辨率介绍

首先介绍图像超分辨率重建技术，图像超分辨率重建技术分为两种，一种是从多张低分辨率图像合成一张高分辨率图像，另外一种是从单张低分辨率图像获取高分辨率图像，在本专栏中，我们使用单幅图像超分辨率重建技术(SISR)。

在这些方法中，可以分为三类，基于插值，基于重建，基于学习。基于插值的方法实现简单，已经广泛应用，但是这些线性的模型限制住了它们恢复高频能力的细节。基于稀疏表示的技术[1]通过使用先验知识增强了线性模型的能力。这类技术假设任意的自然图像可以被字典的元素稀疏表示，这种字典可以形成一个数据库且从数据库中学习到低分辨率图像到高分辨率图像的映射，但是这类方法计算复杂，需要大量计算资源。

基于CNN(卷积神经网络)的模型SRCNN首先将CNN引入SISR中，它仅仅使用三层网络，就取得了先进的结果。随后，各种基于深度学习的模型，进入SISR领域，大致分为以下两个大的方向。一种是追求细节的恢复，以PSNR,SSIM等评价标准的算法，其中以SRCNN模型为代表。另外一种以降低感知损失为目标，不注重细节，看重大局观，以SRGAN为代表的一系列算法。两种不同方向的算法，应用的领域也不相同。

#### 二、SRCNN网络架构



###### 2.1 三层的作用

在进入网络之前会有将input图像用双三次插值放大至目标尺寸的预处理。

1）**LR 特征提取（Patch extraction and representation）**，这个阶段主要是对LR进行特征提取，并将其特征表征为一些feature maps。

2）**特征的非线性映射（Non-linear mapping）**，这个阶段主要是将第一阶段提取的特征映射至HR所需的feature maps。

3）**HR重建（Reconstruction）**，这个阶段是将第二阶段映射后的特征恢复为HR图像。

###### 2.2网络结构细节

LR特征提取可表征为“卷积层（c\*f1\*f1卷积核）+RELU"，c是通道数，f1是卷积核的大小。非线性映射可表征为“全连接层+RELU”，而全连接层又可表征为卷积核为1x1的卷积层，因此，本层最终形式为“卷积层（n1\*1\*1卷积核）+RELU";n1是第一层卷积核的个数。  
HR重建可直接表征为“卷积层（n2\*f3\*f3）”；n2是第二层卷积核的个数。

#### 三、SRCNN设计思路

###### 3.1 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。基于样例的算法目的是找到一组可以表达之前预处理后所得到图像块的一组“基”，这些基是沿着不同方向的边缘，稀疏系数就是分配给各个基的权重。**作者认为这部分可以转化为用一定数量的滤波器（卷积核）来代替。**



###### 3.2 Non-linear mapping

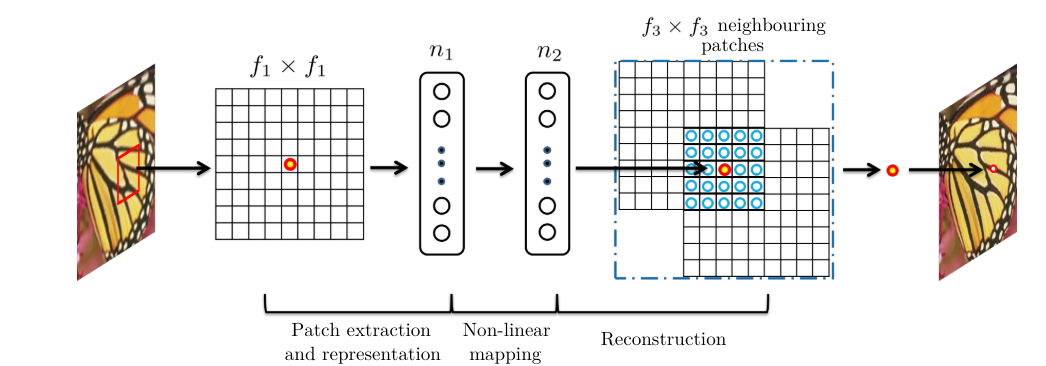
Non-linear mapping: 将低分辨率的特征映射为高分辨率特征，类似于字典学习中的找到图像patch对应的高分辨字典。基于样例的算法将第一步得到的表达图像块的高维向量映射到另外一个高维向量中，通过这个高维向量表达高分辨率图像块，用于最后的重建。**作者认为这一步骤可以使用1\*1的卷积来实现向量维数的变换。**



###### 3.3 Reconstruction

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。基于样例的算法将最后得到的高分辨率图像块进行聚合（重合的位置取平均）形成最后的高分辨率图像。**作者认为这一部分可以看成是一种线性运算，可以构造一个线性函数（不加激活函数）来实现。**





#### 训练细节与效果

###### 4.1训练细节

在SRCNN的训练过程中，作者采用平均均方差（MSE：mean squared error）来作为损失函数，并且为了在训练过程中避免边界效应，并未对图像进行填充。在训练过程中前两层的学习率设为0.0001，最后一层学习率为0.00001。在训练过程中，图片采用的是YCbCr格式的图像，并只对Y通道的图像进行超分辨处理，因此输入的通道数c=1。通过对比不同的数据集以及网络结构，作者得到如下结论。

1）大量的数据集会对训练的结果有一定提升作用，但提升作用不像其他视觉任务那样明显。

2）增加网络的宽度（filter数量）会显著的提升训练结果。

3）增加filter的size可以收集更丰富的结构信息，从而得到更好的结果。

4）增加网络层数不会对结果有明显的提升作用，反而会增加训练的难度。

5）在RGB图像上实现最好的效果，与单通道（Y-only）相比提升并不明显。加入CrCb反而不利于训练。

6）图像被转化为 YCbCr 色彩空间，尽管该网络只使用亮度通道(Y)。然后，网络的输出合并已插值的 CbCr 通道，输出最终彩色图像。**我们选择这一步骤是因为我们感兴趣的不是颜色变化(存储在 CbCr 通道中的信息)而只是其亮度(Y 通道);根本原因在于相较于色差，人类视觉对亮度变化更为敏感。**

###### 4.2训练效果

**本次作业在BSD100数据集进行训练并在Set5上进行测试（模型训练过程见jupyter notebook）。**

