**计算机视觉与应用实践**

课程名称： 计算机视觉与应用实践

学院： 电子工程与光电技术学院 专业： 光学工程

姓名： 徐溢馨  **学号:**  823104010021

南京理工大学

## 一、实验要求

1. 实现一种图像超分辨率方法在Set5数据集上的测试，得到超分辨率图像，测量其与真实图像之间的PSNR、SSIM指标值。对所选择方法细节进行介绍，并试着讨论该方法可能存在的优缺点，以及可能的改进方向。
2. 测试方式：先将图像用Bicubic插值进行下采样，再使用超分辨算法处理，将得到的超分辨率图像与真实的原始图像进行对比。

## 二、方法介绍

本次实验采用的图像超分辨率通过ESPCN模型实现，以下为ESPCN模型详细介绍。

1. 基于深度学习的图像超分辨技术

图像的超分辨率重建技术指的是将给定的低分辨率图像通过特定的算法恢复成相应的高分辨率图像。简单来理解超分辨率重建就是将小尺寸图像变为大尺寸图像，使图像更加“清晰”，但放大时通过了深度学习的技术补充了更多细节。主要包括以下几个步骤：

(1) 特征提取：首先对输入的低分辨率图像进行去噪、上采样等预处理，然后将处理后的图像送入神经网络，拟合图像中的非线性特征，提取代表图像细节的高频信息；

(2) 设计网络结构及损失函数：组合卷积神经网络及多个残差块，搭建网络模型，并根据先验知识设计损失函数；

(3) 训练模型：确定优化器及学习参数，使用反向传播算法更新网络参数，通过最小化损失函数提升模型的学习能力；

(4) 验证模型：根据训练后的模型在验证集上的表现，对现有网络模型做出评估，并据此对模型做出相应的调整。

1. ESPCN模型

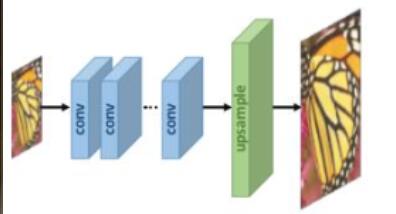
ESPCN 的核心思想是通过深度卷积神经网络学习到一个映射函数，该函数将低分辨率图像的特征转换为高分辨率图像的特征。具体而言，ESPCN 使用多个卷积层进行特征提取，并使用像素重组技术（sub-pixel convolution）来增加图像的分辨率。

在 SRCNN 和 DRCN 中，低分辨率图像都是先通过上采样插值得到与高分辨率图像同样的大小再作为网络输入，这意味着卷积操作在较高的分辨率上进行，相比于在低分辨率的图像上计算卷积会降低效率。ESPCN 提出一种在低分辨率图像上直接计算卷积得到高分辨率图像的高效率方法。它提出了一种新的CNN体系结构，在LR空间中提取特征映射。此外，该模型还引入了一个有效的亚像素卷积层，它学习一组尺度扩展滤波器，以便将最终的LR特征映射放大到HR输出中。

该算法在将低分辨率图像送入神经网络之前，无需对给定的低分辨率图像进行一个上采样过程，而是引入一个亚像素卷积层（Sub-pixel convolution layer），来间接实现图像的放大过程。这种做法极大降低了计算量，提高了重建效率。

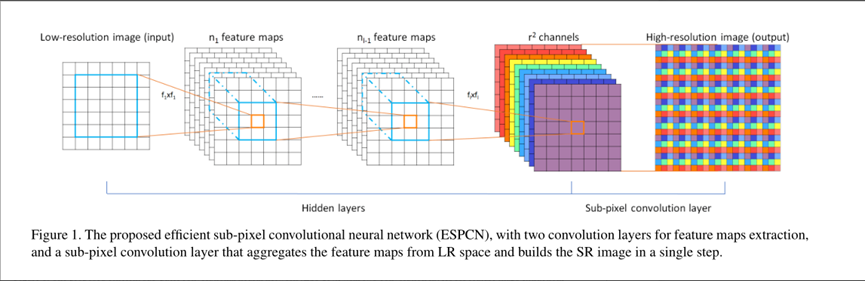
* 1. 新的CNN体系结构

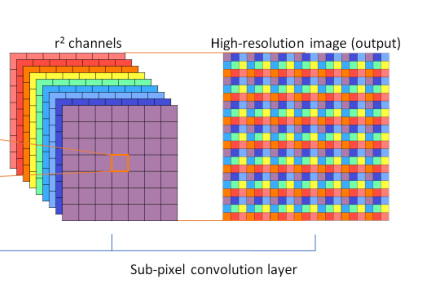
这种新的CNN体系结构在LR空间中提取特征映射，是能够在单个K2 GPU上实时1080p视频SR的卷积神经网络（CNN）。下图为ESPCN的Post-upsampling SR结构。



* 1. 亚像素卷积层(sub-pixel convolutional layer)

亚像素卷积层是ESPCN的核心概念。如下图所示，网络的输入是原始低分辨率图像，通过若干卷积层以后，得到的特征图像大小与输入图像一样，但是特征通道为 。将每个像素的个通道重新排列成一个的区域，对应于高分辨率图像中的一个大小的子块，从而大小为的特征图像被重新排列成大小的高分辨率图像。这个变换虽然被称作sub-pixel convolution , 但实际上并没有卷积操作。亚像素卷积层，并没有做卷积操作，而是直接将个通道的特征图，重新进行排列，这样就完成了LR到HR的恢复。主要学习映射的步骤已经隐含在之前的隐藏层网络中。





总之亚像素卷积层包含两个过程，分别是普通的卷积层和后面的排列像素的步骤。就是说，最后一层卷积层输出的特征个数需要设置成固定值，即放大倍数r的平方，这样总的像素个数就与要得到的高分辨率图像一致，将像素进行重新排列就能得到高分辨率图。

ESPCN提出在网络的最末端实现LR到HR的分辨率，并从LR特征图中获得高分辨率的HR数据。这样就不需要在更大的HR分辨率下执行大部分SR操作。也就是输入是低分辨率图像（未经过插值处理，尺寸和HR图像不一样），在隐藏层中都是低分辨率图像进行处理，也就是在隐藏层没有添加任何关于高分辨率的先验信息。到了亚像素卷积层，才映射到高分辨率空间。将的特征图像，重新排列成高分辨率图像。

亚像素卷积层能够将LR数据映射到HR空间中，与传统的卷积层相比，所需的计算量很小。在ESPCN网络中，图像尺寸放大过程的插值函数被隐含地包含在前面的卷积层中，可以自动学习到。由于卷积运算都是在低分辨率图像尺寸大小上进行，因此效率会较高。

* 1. 其他

训练时，可以将输入的训练数据，预处理成重新排列操作前的格式，比如将21×21的单通道图，预处理成9个通道，7×7的图，这样在训练时，就不需要做重新排列的操作。另外，ESPCN激活函数采用tanh替代了ReLU。损失函数为均方误差。

1. 超分辨率效果评价
   1. PSNR（Peak Signal to Noise Ratio）

峰值信噪比，单位是dB，数值越大表示失真越小。PSNR是最普遍和使用最为广泛的一种图像客观评价指标，然而它是基于对应像素点间的误差，并未考虑到人眼的视觉特性。因为人眼对空间频率较低的对比差异敏感度较高，对亮度对比差异的敏感度较色度高，人眼对一个区域的感知结果会受到其周围邻近区域的影响，因而经常出现评价结果与人的主观感觉不一致的情况。

* 1. SSIM（structural similarity）

结构相似，结构相似性，SSIM从亮度、对比度和结构这三个方面来评估两幅图像的相似性，是一种衡量两幅图像相似度的指标。SSIM使用的两张图像中，一张为未经压缩的无失真图像，另一张为失真后的图像。结构相似性理论认为，自然图像具有极高的结构性，表现在图像的像素间存在着很强的相关性，尤其是在空间相似的情况下。当两张图像一模一样时，SSIM的值等于1。

## 三、实现说明与结果

1. 实现说明

读入Set5数据集图像，对图像进行预处理，使用Bicubic插值进行下采样，调用ESPCN模型对图像进行超分辨处理并保存处理后的图像。计算对比原始图像和超分辨后图像的PSNR和SSIM，评估模型效果。

1. 实验结果与分析
   1. PSNR和SSIM

|  |  |
| --- | --- |
| Baby |  |
| Bird |  |
| Butterfly |  |
| Head |  |
| Woman |  |

1. 图像对比

|  |  |  |
| --- | --- | --- |
| 原始图像 | Bicubic处理 | ESPCN处理 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

1. 分析

由实验结果可得，ESPCN模型超分辨图像的PSNR和SSIM都体现较好的结果，SSIM基本能达到0.9以上，表明超分辨后的图像与原始图像相似度很高。其中，Baby组的PSNR最高，Bird组的SSIM最高。

而从对比图中可以看出，ESPCN能够一定程度上使经过Bicubic插值下采样后的图像恢复到近似与原始图像，其中Baby组肉眼效果恢复最好。

## 四、方法讨论

1. ESPCN优点：
2. 采用亚像素对高分辨率信息进行恢复，与传统卷积层相比计算量小。
3. ESPCN可以适用于各种类型的低分辨率图像，如灰度图像、彩色图像等。同时，它也可以适用于不同的超分辨率应用场景，如单帧超分辨率、视频超分辨率等。
4. ESPCN的网络结构相对简单，易于实现和使用。同时，它的参数数量也相对较少，因此可以更快地进行训练和推断。
5. ESPCN缺点：
   1. 使用MSE作为目标函数来训练模型，MSE损失函数会导致重建的图像过于平滑，缺乏感官上的照片真实感。
   2. 当输入图像的分辨率非常低时，ESPCN可能会出现重构质量下降的问题。
   3. ESPCN在处理纹理和细节复杂的图像时也可能会出现一些困难。
6. 改进思路：
7. 模型结构的改进：可以考虑设计更加复杂和高效的网络结构来提高ESPCN的性能。例如，可以使用更深的卷积神经网络或引入注意力机制等方法来提取更加丰富的特征信息。
8. 数据增强技术：可以使用数据增强技术来扩充训练数据集，从而提高模型的泛化能力。例如，可以通过旋转、翻转、裁剪等方式来增加训练样本的多样性。
9. 多模态融合：可以考虑将ESPCN与其他图像处理技术进行融合，从而提高图像超分辨率的效果。例如，可以将ESPCN与生成对抗网络（GAN）相结合，利用GAN的生成能力来提高重构图像的质量和多样性。
10. 实时超分辨率：可以考虑研究实时超分辨率技术，即在视频流中实时地对低分辨率图像进行超分辨率处理。这需要设计高效的算法和硬件系统来保证实时性和处理速度。