# SRCNN实现图像超分辨分析

## 介绍

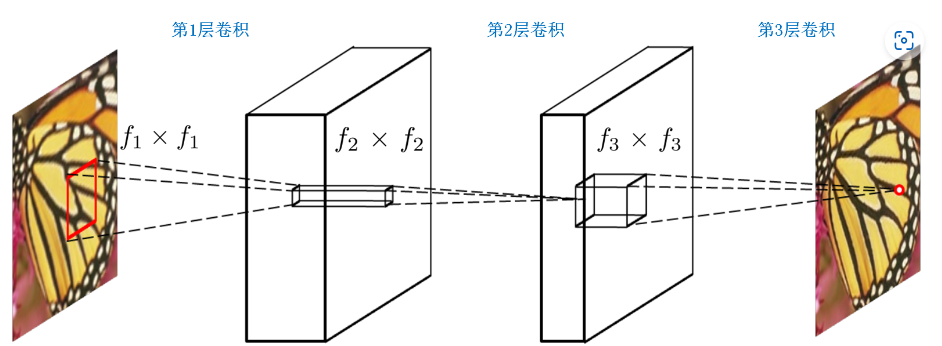
SRCNN采用了插值的方式先将低分辨率图像进行放大，再通过模型进行复原。Shi等人则认为这种预先采用近邻插值的方式本身已经影响了性能，如果从源头出发，应该从样本中去学习如何进行放大，他们基于这个原理提出了ESPCN (Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network)算法。该算法在将低分辨率图像送入神经网络之前，无需对给定的低分辨率图像进行一个上采样过程，而是引入一个亚像素卷积层（Sub-pixel convolution layer），来间接实现图像的放大过程。这种做法极大降低了SRCNN的计算量，提高了重建效率。

这里需要注意到，不管是SRCNN还是ESPCN，它们均使用了MSE作为目标函数来训练模型。2017年，Christian Ledig等人从照片感知角度出发，通过对抗网络来进行超分重建（论文题目：Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network）。他们认为，大部分深度学习超分算法采用的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。他们改用生成对抗网络（Generative Adversarial Networks, GAN）来进行重建，并且定义了新的感知目标函数，算法被命名为SRGAN，由一个生成器和一个判别器组成。生成器负责合成高分辨率图像，判别器用于判断给定的图像是来自生成器还是真实样本。通过一个博弈的对抗过程，使得生成器能够将给定的低分辨率图像重建为高分辨率图像。在SRGAN这篇论文中，作者同时提出了一个对比算法，名为SRResNet。SRResNet依然采用了MSE作为最终的损失函数，与以往不同的是，SRResNet采用了足够深的残差卷积网络模型，相比于其它的残差学习重建算法，SRResNet本身也能够取得较好的效果。

## 原理

### 超分重建基本处理流程

早的采用深度学习进行超分重建的算法是SRCNN算法，其原理很简单，对于输入的一张低分辨率图像，SRCNN首先使用双立方插值将其放大至目标尺寸，然后利用一个三层的卷积神经网络去拟合低分辨率图像与高分辨率图像之间的非线性映射，最后将网络输出的结果作为重建后的高分辨率图像。尽管原理简单，但是依托深度学习模型以及大样本数据的学习，在性能上超过了当时一众传统的图像处理算法，开启了深度学习在超分辨率领域的研究征程。SRCNN的网络结构如图1所示。



其中f_1、f_2、f_3分别表示1、2、3层卷积对应的核大小。

SRCNN作为早期开创性的研究论文，也为后面的工作奠定了处理超分问题的基本流程：

(1) 寻找大量真实场景下图像样本；

(2) 对每张图像进行下采样处理降低图像分辨率，一般有2倍下采样、3倍下采样、4倍下采样等。如果是2倍下采样，则图像长宽均变成原来的1/2.。下采样前的图像作为高分辨率图像H，下采样后的图像作为低分辨率图像L，L和H构成一个有效的图像对用于后期模型训练；

(3) 训练模型时，对低分辨率图像L进行放大还原为高分辨率图像SR，然后与原始的高分辨率图像H进行比较，其差异用来调整模型的参数，通过迭代训练，使得差异最小。实际情况下，研究学者提出了多种损失函数用来定义这种差异，不同的定义方式也会直接影响最终的重建效果；

(4) 训练完的模型可以用来对新的低分辨率图像进行重建，得到高分辨率图像。

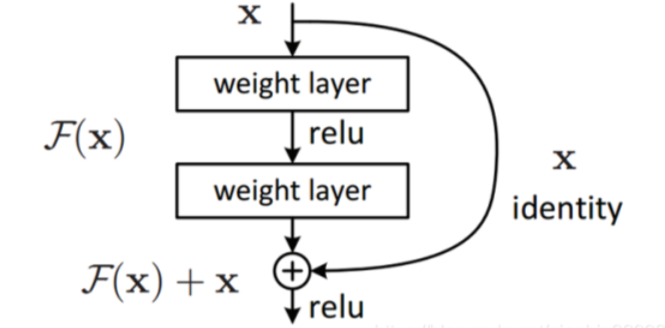
从实际操作上来看，整个超分重建分为两步：图像放大和修复。所谓放大就是采用某种方式（SRCNN采用了插值上采样）将图像放大到指定倍数，然后再根据图像修复原理，将放大后的图像映射为目标图像。超分辨率重建不仅能够放大图像尺寸，在某种意义上具备了图像修复的作用，可以在一定程度上削弱图像中的噪声、模糊等。因此，超分辨率重建的很多算法也被学者迁移到图像修复领域中，完成一些诸如jpep压缩去燥、去模糊等任务。

### 构建深度网络模型提高超分重建性能

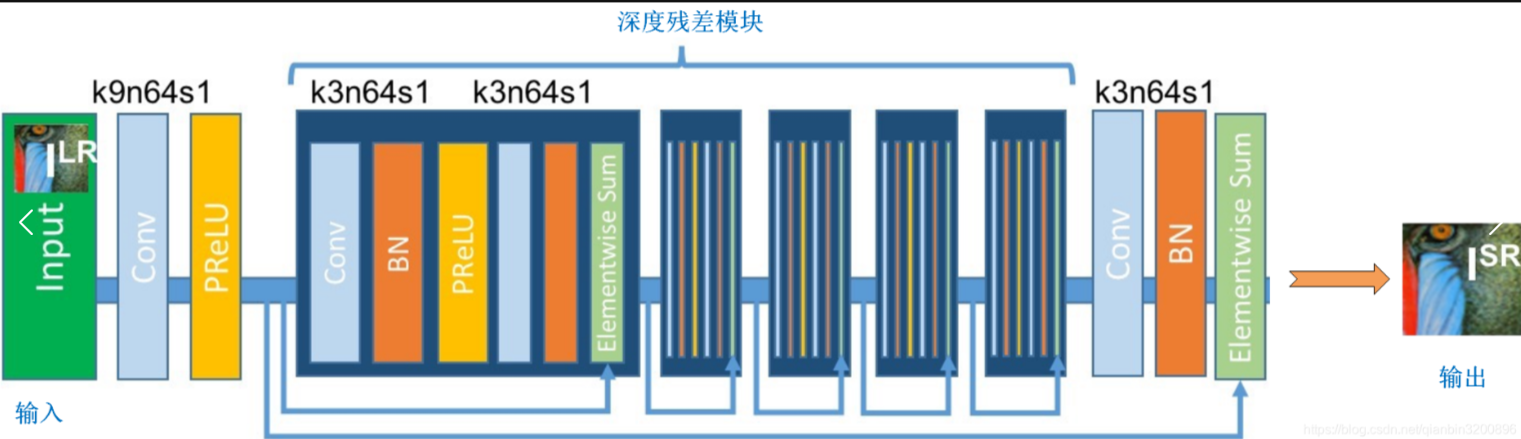
SRCNN只采用了3个卷积层来实现超分重建，有文献指出如果采用更深的网络结构模型，那么可以重建出更高质量的图像，因为更深的网络模型可以抽取出更高级的图像特征，这种深层模型对图像可以更好的进行表达。在SRCNN之后，有不少研究人员尝试加深网络结构以期取得更佳的重建性能，但是越深的模型越不能很好的收敛，无法得到期望的结果。部分研究学者通过迁移学习来逐步的增加模型深度，但这种方式加深程度有限。因此，亟需一种有效的模型，使得构建深层网络模型变得容易并且有效。这个问题直到2015年由何凯明团队提出ResNet网络才得以有效解决。

ResNet中文名字叫作深度残差网络，主要作用是图像分类。现在在图像分割、目标检测等领域都有很广泛的运用。ResNet在传统卷积神经网络中加入了残差学习（residual learning），解决了深层网络中梯度弥散和精度下降（训练集）的问题，使网络能够越来越深，既保证了精度，又控制了速度。

ResNet可以直观的来理解其背后的意义。以往的神经网络模型每一层学习的是一个 y = f(x) 的映射，可以想象的到，随着层数不断加深，每个函数映射出来的y误差逐渐累计，误差越来越大，梯度在反向传播的过程中越来越发散。这时候，如果改变一下每层的映射关系，改为 y = f(x) + x,也就是在每层的结束加上原始输入，此时输入是x，输出是f(x)+x，那么自然的f(x)趋向于0，或者说f(x)是一个相对较小的值，这样，即便层数不断加大，这个误差f(x)依然控制在一个较小值，整个模型训练时不容易发散。



上图为残差网络的原理图，可以看到一根线直接跨越两层网络（跳链），将原始数据x带入到了输出中，此时F(x)预测的是一个差值。有了残差学习这种强大的网络结构，就可以按照SRCNN的思路构建用于超分重建的深度神经网络。SRResNet算法主干部分就采用了这种网络结构，如下图所示：



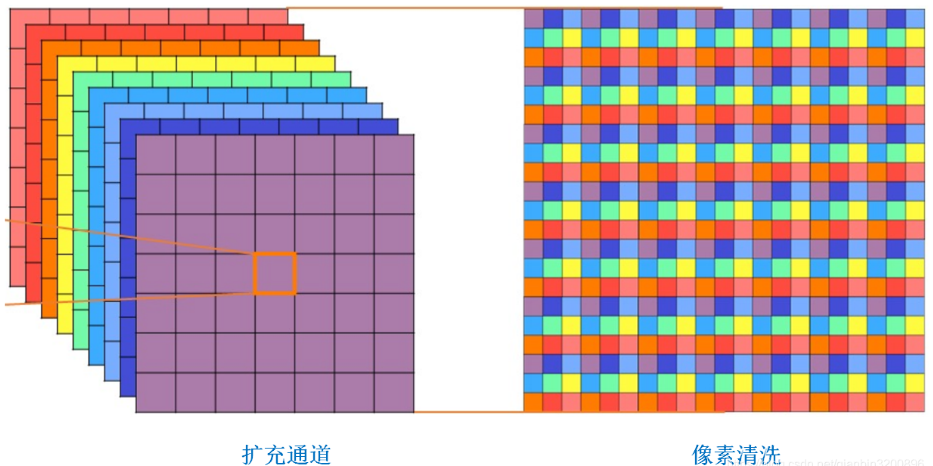
上述模型采用了多个深度残差模块进行图像的特征抽取，多次运用跳链技术将输入连接到网络输出，这种结构能够保证整个网络的稳定性。由于采用了深度模型，相比浅层模型能够更有效的挖掘图像特征，在性能上可以超越浅层模型算法（SRResNet使用了16个残差模块）。注意到，上述模型每层仅仅改变了图像的通道数，并没有改变图像的尺寸大小，从这个意义上来说这个网络可以认为是前面提到的修复模型。下面会介绍如何在这个模型基础上再增加一个子模块用来放大图像，从而构建一个完整的超分重建模型。

### 基于子像素卷积放大图像尺寸

子像素卷积（Sub-pixel convolution）是一种巧妙的图像及特征图放大方法，又叫做pixel shuffle（像素清洗）。在深度学习超分辨率重建中，常见的扩尺度方法有直接上采样，双线性插值，反卷积等等。ESPCN算法中提出了一种超分辨率扩尺度方法，即为子像素卷积方法，该方法后续也被应用在了SRResNet和SRGAN算法中。因此，这里需要先介绍子像素卷积的原理及实现方式。

采用CNN对特征图进行放大一般会采用deconvolution等方法，这种方法通常会带入过多人工因素，而子像素卷积会大大降低这个风险。因为子像素卷积放大使用的参数是需要学习的，相比那些手工设定的方式，这种通过样本学习的方式其放大性能更加准确。

具体实现原理如下图所示：



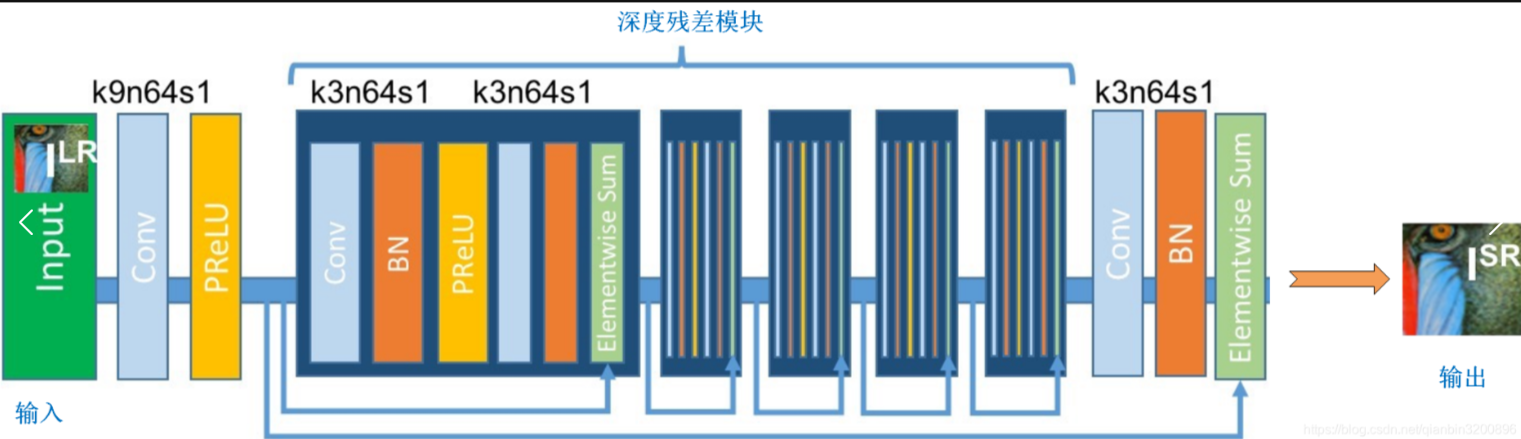
上图很直观得表达了子像素卷积的流程。假设，如果想对原图放大3倍，那么需要生成出3^2=9个同等大小的特征图，也就是通道数扩充了9倍（这个通过普通的卷积操作即可实现）。然后将九个同等大小的特征图拼成一个放大3倍的大图，这就是子像素卷积操作了。

实现时先将原始特征图通过卷积扩展其通道数，如果是想放大4倍，那么就需要将通道数扩展为原来的16倍。特征图做完卷积后再按照特定的格式进行排列，即可得到一张大图，这就是所谓的像素清洗。通过像素清洗，特征的通道数重新恢复为原来输入时的大小，但是每个特征图的尺寸变大了。这里注意到每个像素的扩展方式由对应的卷积来决定，此时卷积的参数是需要学习的，因此，相比于手工设计的放大方式，这种基于学习的放大方式能够更好的去拟合像素之间的关系。

SRResNet模型也利用子像素卷积来放大图像，具体的，在图5所示模型后面添加两个子像素卷积模块，每个子像素卷积模块使得输入图像放大2倍，因此这个模型最终可以将图像放大4倍。

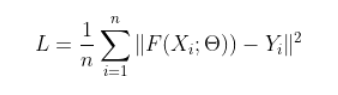
### SRResNet结构剖析

SRResNet使用深度残差网络来构建超分重建模型，主要包含两部分：深度残差模型、子像素卷积模型。深度残差模型用来进行高效的特征提取，可以在一定程度上削弱图像噪点。子像素卷积模型主要用来放大图像尺寸。完整的SRResNet网络结果如下图所示：



上图中，k表示卷积核大小，n表示输出通道数，s表示步长。除了深度残差模块和子像素卷积模块以外，在整个模型输入和输出部分均添加了一个卷积模块用于数据调整和增强。

需要注意的是，SRResNet模型使用MSE作为目标函数，也就是通过模型还原出来的高分辨率图像与原始高分辨率图像的均方误差，公式如下：



MSE也是目前大部分超分重建算法采用的目标函数。后面我们会看到，使用该目标函数重建的超分图像并不能很好的符合人眼主观感受，SRGAN算法正是基于此进行的改进。

## 优缺点与改进

SRResNet作为一种用于单幅图像超分辨率的深度学习模型，具有以下优点和缺点：

### 优点：

**高质量的超分辨率重建**：SRResNet能够生成高质量、细节丰富的高分辨率图像，提升了图像的视觉质量。

**利用残差学习：**采用残差网络结构，有助于训练深层网络，减轻了梯度消失问题，提高了模型的学习效率和性能。

**有效地捕获图像特征：**SRResNet通过多个残差块来有效地捕获图像中的特征信息，提升了超分辨率重建的准确性。

### 缺点：

**计算复杂度较高：**由于SRResNet是一个深度神经网络模型，需要较大的计算资源进行训练和推理，导致计算复杂度较高。

**对训练数据需求高**：SRResNet在训练时需要大量的高分辨率图像与对应的低分辨率图像配对，对训练数据的要求较高。

### 可改进的方向：

**减小模型规模**：可以尝试简化SRResNet的结构，减小模型规模以降低计算复杂度，例如通过剪枝技术或模型压缩方法等。

**引入注意力机制**：考虑在SRResNet中引入注意力机制，使模型能够更加关注重要的像素区域，提高重建图像的质量。

**多尺度融合**：探索多尺度特征融合的方法，结合不同尺度的信息来增强特征表示，改善超分辨率重建效果。

**数据增强策略**：设计更加有效的数据增强策略，以扩充训练数据集，提高模型泛化能力。

通过在这些方面进行改进，可以进一步提升SRResNet模型的性能和效率，使其在图像超分辨率任务中表现更加优异。