

学号	237214590	姓名	郎志远		
评分项					
完整性(20%)	准确性(20%)	逻辑性(20%)	深刻性(20%)	规范性(20%)	总分

## 卷积神经网络在图像超分中的应用

### 1. 引言

图像超分的目标是从低分辨率图像中恢复出高分辨率图像，传统的图像超分方法主要依赖于图像的稀疏表示和插值技术，这些方法虽然在一定程度上能够提高图像的分辨率，但在细节恢复和视觉质量上往往存在局限性。卷积神经网络以其在图像特征提取和模式识别方面的强大能力为图像超分提供了新的视角。本文从卷积神经网络的原理出发，以 SRCNN<sup>[1]</sup>为例介绍了其在图像超分中的应用，并对卷积神经网络在图像超分领域能取得较好效果的原因进行了探索。

### 2. 卷积神经网络基本概念

卷积神经网络主要由卷积层、池化层、激活函数以及全连接层组成，典型的卷积神经网络结构如图 1 所示<sup>[2]</sup>，其中卷积层和池化层是网络的核心结构。

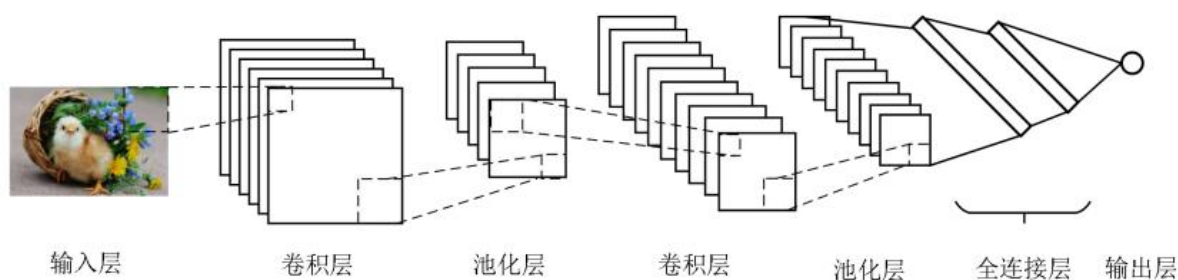


图 1 典型的卷积神经网络结构

#### 2.1 卷积层与激活函数

卷积层的设计模仿了生物通过局部感受野来感知图像的特点，负责从输入数据中提取边缘、角点、纹理局部特征。在卷积层中，卷积操作是通过一个或多个卷积核（也称为滤波器或权重矩阵）在输入数据上滑动并进行计算的过程。每个卷积核包含一组可学习的权重，这些权重定义了卷积核在输入数据上滑动时的计算方式。卷积核的大小决定了它在输入数据上覆盖的区域。例如，一个  $3 \times 3$  的卷积核在每次滑动时会覆盖输入数据的一个  $3 \times 3$  的局部区域。卷积核的深度与输入数据的通道数相匹配。步长是卷积核在输入数据上滑动的步长。步长为 1 时，卷积核在输入数据上逐像素滑动；步长大于 1 时，卷积核在输入数据上跳跃式滑动，这会导致输出特征图的尺寸减小。在某些情况下，

为了保持输出特征图的空间尺寸，或者为了引入边界信息，可以在输入数据的边缘添加额外的零值，这个过程称为填充。填充可以是零填充，也可以是其他类型的边界填充策略。

卷积操作的结果通常需要通过一个激活函数来引入非线性，使得网络能够学习到更复杂的特征。常用的激活函数包括 ReLU、Sigmoid、Tanh 等。ReLU 激活函数因其计算效率高和能够缓解梯度消失问题而被广泛使用。

## 2.2 池化层

池化层（也称汇聚层或子采样层）位于卷积层之后，主要用于降低特征图的空间维度，减少计算量，同时保留重要特征，增强模型的泛化能力，减少模型的复杂度，从而降低过拟合的风险。目前主流的卷积神经网络中池化层仅包含下采样操作，常见的池化操作主要包括最大池化和平均池化。最大池化选择池化窗口内的最大值作为输出，这种操作能够突出显示区域内的最显著的特征。而平均池化则计算池化窗口内所有值的平均值，它方法试图捕捉区域内的整体信息，而不是仅仅关注最显著的特征，在某些情况下平均池化可以提供更平滑的特征表示，有助于减少噪声的影响。

## 3. SRCNN

SRCNN（Super-Resolution Convolutional Neural Network）是一个开创性的超分模型，由何凯明、汤晓鸥等人在 2014 年提出，它首次将卷积神经网络应用于单图像超分辨率问题的模型。SRCNN 的核心思想是利用卷积神经网络的自动特征学习能力，直接从低分辨率图像学习到高分辨率图像的映射。与传统卷积神经网络相比 SRCNN 结构非常简单，整个卷积网络包括三个卷积层，甚至没有池化和全连接层，SRCNN 首先使用双三次插值将低分辨率图像放大成目标尺寸，接着通过三层卷积网络拟合非线性映射，最后输出高分辨率图像结果，具体来说它有三个处理流程：

- （1）图块的提取和重新表达，SRCNN 的第一个卷积层使用  $9 \times 9$  的卷积核负责从低分辨率图像中提取特征，每个块被卷积操作表示为多维的向量，所有的特征向量组成特征图
- （2）非线性映射，在第一个卷积层之后，SRCNN 将  $n_1$  特征矩阵，使用  $1 \times 1$  的卷积核实现非线性映射，变成另一  $n_2$  维特征矩阵。
- （3）重构，最后通过  $5 \times 5 \times$  的卷积核对非线性映射后的特征进行重构，生成高分辨率的图像。

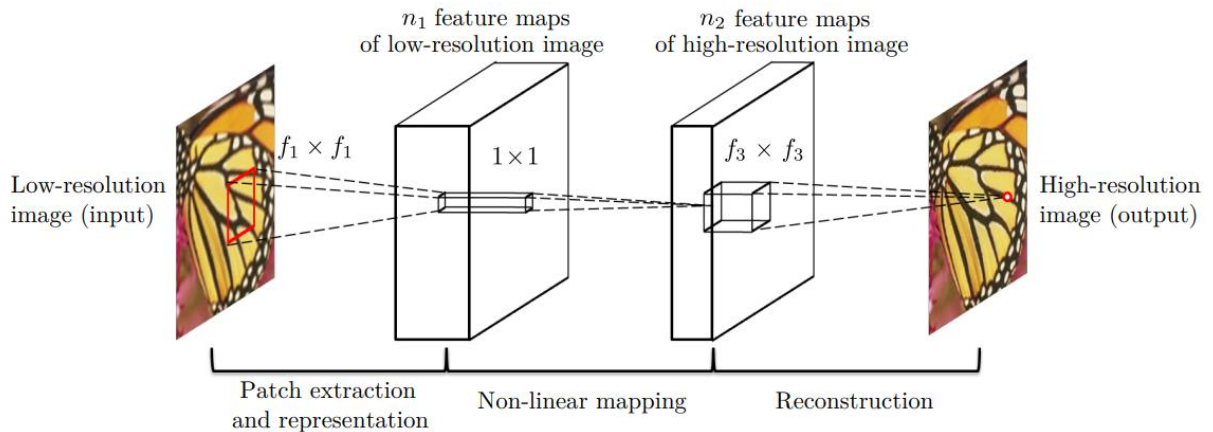


图 2 SRCNN 的网络结构

在 SRCNN 算法出现之前，进行图像重建主要是基于样例学习的方法<sup>[3]</sup>，其中包括两种主流的思路，第一种是根据同一张图片的内部相似性<sup>[4]</sup>，第二种是寻找低分辨率图片与高分辨率图片对之间的映射关系<sup>[5]</sup>。基于样例学习的方法从图像中提取大量的图像块，然后对图像块进行预处理后将其编码成低分辨率字典，将稀疏系数传到高分辨率字典中，最后用学习到的高分辨率字典产生高分辨率的图像块聚合到一起产生高分辨率图片，然而基于样例学习的方法需要大量的数据，但数据过多，难以对他们进行有效地建模。在这一过程中，研究者们往往专注于编码和字典学习技术的优化，而忽视了其他步骤的改进。即便有对这些步骤的优化，也往往是孤立进行，没有将它们视为一个整体

来考虑。

对比传统方法，SRCNN 引进了卷积，实现了局部连接和参数共享，每个神经元只与输入数据的一个局部区域相连。这种设计减少了模型的参数数量，提高了计算效率，并且通过参数共享降低了过拟合的风险。并且原有的稀疏编码过程是一个迭代过程，而 SRCNN 是一个端到端的过程，除了对输入进行双三次插值的操作外没有其他预处理操作，并且 CNN 能够自动从图像数据中学习到有用的特征，无需人工设计复杂的特征提取器。

卷积神经网络在图像超分辨率领域取得较好效果的原因可以归结为以下几个关键点：

（1）自动特征学习：CNN 能够自动从图像数据中学习到有用的特征，无需人工设计复杂的特征提取器，这大大减少了特征工程的工作量，并且能够捕捉到更加复杂和抽象的特征。

（2）局部连接和参数共享：NN 通过卷积层实现局部连接，即每个神经元只与输入数据的一个局部区域相连。这种设计减少了模型的参数数量，提高了计算效率，并且通过参数共享降低了过拟合的风险。

（3）平移不变性：CNN 的设计模仿了人类视觉系统的局部感受野，使得网络对图像的平移、旋转和缩放具有一定程度的不变性，这使得模型在处理图像时更加鲁棒。

（4）层次化特征提取：CNN 通过多层结构，能够从低级到高级逐步提取图像的特征。较低层次可能捕捉到边缘和纹理等简单特征，而高层次则能够理解更复杂的模式和对象。

（5）池化层的降维作用：池化层在保持特征的同时减少了数据的空间维度，有助于减少计算量，并且增强了模型对小的图像变化的鲁棒性。

综上所述，CNN 在图像超分辨率领域取得较好效果的原因在于其强大的特征学习能力、端到端的训练能力、参数共享和局部连接的特性、层次化的特征表示等特性，这些特性共同使得 CNN 成为图像超分辨率任务中的强大工具。SRCNN 的成功展示了 CNN 在图像超分辨率领域的潜力，为后续的研究和应用奠定了基础，它不仅推动了深度学习在图像处理领域的应用，也为后续的超分辨率模型提供了灵感，后续的 VDSR (Very Deep Super-Resolution)<sup>[6]</sup>采用了残差网络的结构，通过构建更深的网络来进一步提升超分辨率的性能，SRGAN (Super-Resolution Generative Adversarial Network)<sup>[7]</sup>则将生成对抗网络引入超分，通过对抗损失函数显著提升了生成图像的感知质量，使得超分辨率图像更加逼真，ESRGAN (Enhanced Super-Resolution Generative Adversarial Network)<sup>[8]</sup>作为 SRGAN 的增强版，通过引入感知损失函数和多尺度结构，不仅提高了图像的分辨率，还显著提升了图像的视觉效果，使得生成的图像在细节和纹理上更加丰富和真实。而 Transformer 的出现激发了对 CNN 和 Transformer 混合模型的研究，Transformer 通过自注意力机制提供了一种全新的全局信息建模方式，这在一定程度上弥补了 CNN 在处理全局依赖和长距离特征关系方面的不足，相信随着深度学习的发展和技术的不断进步，CNN 在计算机视觉领域的应用将更加广泛和深入。

## 参考文献

- [1] Yoon Y , Jeon H G , Yoo D ,et al.Learning a Deep Convolutional Network for Light-Field Image Super-Resolution[C]IEEE, 2015.DOI:10.1109/ICCVW.2015.17.
- [2] 李春平.基于卷积神经网络的图像超分辨率重建算法研究[D].华北电力大学(北京),2019.
- [3] 傅瑜.基于卷积神经网络的图像超分辨率重建算法研究[D].重庆大学,2020.
- [4] GlasnerD,BagonS,IraniM .Super-resolution from a single image[J].IEEE, 2009.DOI:10.1109/ICCV.2009.5459271.
- [5] Bevilacqua M , Roumy A , Guillemot C ,et al.Low-Complexity Single Image Super-Resolution Based on NonnegativeNeighborEmbedding[C]//BritishMachineVisionConference.BMVA Press,2012.DOI:10.5244/C.26.135
- [6] Kim J , Lee J K , Lee K M .Accurate Image Super-Resolution Using Very Deep Convolutional Networks[J].IEEE, 2016.DOI:10.1109/CVPR.2016.182.
- [7] Ledig C , Theis L , Huszar F ,et al.Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[J].IEEE Computer Society, 2016.DOI:10.1109/CVPR.2017.19.
- [8] Wang X , Yu K , Wu S ,et al.ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks[J]. 2018.DOI:10.1007/978-3-030-11021-5\_5.