

[首页](#)[活动](#)[未来网络学院](#)[实验平台](#)[特色](#)[知识社区](#)[网络平台优化](#)[SDN](#)[NFV](#)[NetDevOps](#)[5G](#)[物联网](#)[网络安全](#)[搜索](#)[登录](#) | [注册](#)

基于深度学习的图像超分辨率重建技术的研究

SDNLAB君 • 18-10-09 • 649 人围观

作者：刘永信，段添添，内蒙古大学，博士，教授，博导。

摘要：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。随着人工智能的不断发展，超分辨率重建技术在视频图像压缩传输、医学成像、遥感成像、视频感知与监控等领域得到了广泛的应用与研究。本文简要介绍了图像超分辨率技术的研究背景与意义，同时概述了其基本原理及评估指标，然后着重介绍了基于深度学习的超分辨率重建技术的处理流程及几种具有代表性的超分辨率深度学习模型。

关键词：人工智能，深度学习，超分辨率



SDNLAB君 发表于18-10-09

0

0



1 超分辨率重建技术的研究背景与意义

图像分辨率是一组用于评估图像中蕴含细节信息丰富程度的性能参数，包括时间分辨率、空间分辨率及色阶分辨率等，体现了成像系统实际所能反映物体细节信息的能力。相较于低分辨率图像，高分辨率图像通常包含更大的像素密度、更丰富的纹理细节及更高的可信赖度。

但在实际应用中，受采集设备与环境、网络传输介质与带宽、图像退化模型本身等诸多因素的约束，我们通常并不能直接得到具有边缘锐化、无成块模糊的理想高分辨率图像。

提升图像分辨率的最直接的做法是对采集系统中的光学硬件进行改进，但这种做法受限于制造工艺难以大幅改进、制造成本十分高昂等约束。由此，从软件和算法的角度着手，实现图像超分辨率重建的技术成为了图像处理和计算机视觉等多个领域的热点研究课题。

1955年，Torvaldo di Francia在光学成像领域首次明确定义了超分辨率这一概念，主要是指利用光学相关的知识，恢复出行射极限以外的数据信息的过程。1964年左右，Harris和Goodman则首次提出了图像超分辨率这一概念，主要是指利用外推频谱的方法合成出细节信息更丰富的单帧图像的过程。1984年，在前人的基础上，Tsai和



图像的方法后, 超分辨率重建技术开始受到了学术界和工业界广泛的关注和研究。

具体来说, 图像超分辨率重建技术指的是利用数字图像处理、计算机视觉等领域的相关知识, 借由特定的算法和流程, 从给定的低分辨率图像中复原出高分辨率图像的过程。其旨在克服或补偿由于图像采集系统或采集环境本身的限制, 导致的成像图像模糊、质量低下、感兴趣区域不显著等问题。

图像超分辨率重建技术在多个领域都有着广泛的应用范围和研究意义。主要包括:

(1) 图像压缩领域

在视频会议等实时性要求较高的场合, 可以在传输前预先对图片进行压缩, 等待传输完毕, 再由接收端解码后通过超分辨率重建技术复原出原始图像序列, 极大减少存储所需的空间及传输所需的带宽。

(2) 医学成像领域

对医学图像进行超分辨率重建, 可以在不增加高分辨率成像技术成本的基础上, 降低对成像环境的要求, 通过复原出的清晰医学影像, 实现对病变细胞的精准探测, 有助于医生对患者病情做出更好的诊断。

(3) 遥感成像领域

高分辨率遥感卫星的研制具有耗时长、价格高、流程复杂等特点, 由此研究者将图像超分辨率重建技术引入了该领域, 试图解决高分辨率的遥感成像难以获取这一挑战, 使得能够在不改变探测系统本身的前提下提高观测图像的分辨率。

(4) 公共安防领域

公共场合的监控设备采集到的视频往往受到天气、距离等因素的影响, 存在图像模糊、分辨率低等问题。通过对采



牌号码、清晰人脸等重要信息，为案件侦破提供必要线索。

(5) 视频感知领域

通过图像超分辨率重建技术，可以起到增强视频画质、改善视频的质量，提升用户的视觉体验的作用。

2 图像超分辨率重建技术概述

2.1 降质退化模型

低分辨率图像在成像的过程中受到很多退化因素的影响，运动变换、成像模糊和降采样是其中最主要的三个因素。如图1所示，整个过程可以通过使图示的线性变换模型来表征。

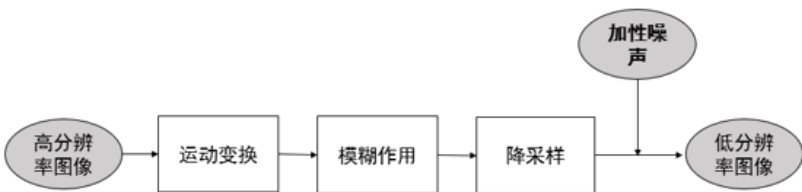


图1 图像的降质退化模型

上述退化模型可以由以下线性变换表示；

$$L = DBFH + N$$

(1)

式中，L表示观测图像，H表示输入的高分辨率图像，F表示运动变换矩阵，通常由运动、平移等因素造成，B表示模糊作用矩阵，通常由环境或成像系统本身引起，D表示降采样矩阵，通常由成像系统的分辨率决定，N表示加性噪声，通常来自于成像环境或成像过程。

图像降质退化模型描述了自然界中的高分辨率图像转换成人眼观测到的低分辨率图像的整个过程，即高分辨率图像成像过程，为图像超分辨率技术提供了坚实的理论基

2.2 重建图像的评估

为了衡量重建算法优劣，需要引入一种评估指标来对重建后的图像进行评估。重建图像的评价方式一般分为两大类，一是主观评价，二是客观评价。

主观评价以人为评价主体，对重建后图像的视觉效果做出主观和定性的评估。为保证图像的主观评价具有一定的统计意义，此种评估方法需要选择足够多的评价主体，并保证评价主体中未受训练的普通人和受过训练的专业人员数量大致均衡。

客观评价中，峰值信噪比（Peak signal-to-noise ratio，PSNR）和结构相似性（Structural Similarity，SSIM）是最常用的两种图像质量评估指标。其中PSNR通过比较两幅图像对应像素点的灰度值差异来评估图像的好坏，SSIM则从亮度、对比度和结构这三个方面来评估两幅图像的相似性。具体计算公式如下：

$$PSNR = 10 \log_{10} \frac{255^2 \times W \times H}{\sum_{i=1}^W \sum_{j=1}^H [\tilde{x}(i,j) - x(i,j)]^2} \quad (2)$$

$$SSIM = \frac{(2\mu_x \mu_{\tilde{x}} + C_1)(2\sigma_{x\tilde{x}} + C_2)}{(\mu_x^2 + \mu_{\tilde{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\tilde{x}}^2 + C_2)} \quad (3)$$

2.3 图像分辨率重建技术分类

根据分类准则的不同，可以将图像超分辨率重建技术划分为不同的类别。从输入的低分辨率图像数量角度来看，可以分为单帧图像的超分辨率重建和多帧图像（视频）的超分辨率重建；从变换空间角度来看，可以分为频域超分辨率重建、时域超分辨率重建、色阶超分辨率重建等；从重建算法角度来看，可以分为基于插值的重建、基于重构的重建和基于学习的超分辨率重建。

本节主要从算法内容出发，介绍几类常见的超分辨率重建



(1) 基于插值的超分辨率重建

基于插值的方法将每一张图像都看做是图像平面上的一个点，那么对超分辨率图像的估计可以看做是利用已知的像素信息为平面上未知的像素信息进行拟合的过程，这通常由一个预定义的变换函数或者插值核来完成。基于插值的方法计算简单、易于理解，但是也存在一些明显的缺陷。

首先，它假设像素灰度值的变化是一个连续的、平滑的过程，但实际上这种假设并不完全成立。其次，在重建过程中，仅根据一个事先定义的转换函数来计算超分辨率图像，不考虑图像的降质退化模型，往往会导致复原出的图像出现模糊、锯齿等现象。常见的基于插值的方法包括最近邻插值法、双线性插值法和双立方插值法等。

(2) 基于重构的超分辨率重建

基于重构的方法则是从图像的降质退化模型出发，假定高分辨率图像是经过适当的运动变换、模糊及噪声才得到低分辨率图像。这种方法通过提取低分辨率图像中的关键信息，并结合对未知的超分辨率图像的先验知识来约束超分辨率图像的生成。常见的基于重构的方法包括迭代反投影法、凸集投影法和最大后验概率法等。

(3) 基于学习的超分辨率重建

基于学习的方法则是利用大量的训练数据，从中学习低分辨率图像和高分辨率图像之间某种对应关系，然后根据学习到的映射关系来预测低分辨率图像所对应的高分辨率图像，从而实现图像的超分辨率重建过程。常见的基于学习的方法包括流形学习、稀疏编码和深度学习方法。

3 基于深度学习的图像超分辨率重建技术

机器学习是人工智能的一个重要分支，而深度学习则是机器学习中最主要的一个算法，其旨在通过多层非线性变换，提取数据的高层抽象特征，学习数据潜在的分布规



随着人工智能和计算机硬件的不断发展，Hinton等人在2006年提出了深度学习这一概念，其旨在利用多层非线性变换提取数据的高层抽象特征。凭借着强大的拟合能力，深度学习开始在各个领域崭露头角，特别是在图像与视觉领域，卷积神经网络大放异，这也使得越来越多的研究者开始尝试将深度学习引入到超分辨率重建领域。2014年，Dong等人首次将深度学习应用到图像超分辨率重建领域，他们使用一个三层的卷积神经网络学习低分辨率图像与高分辨率图像之间映射关系，自此，在超分辨率重建率领域掀起了深度学习的浪潮。

基于深度学习的图像超分辨率技术的重建流程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。

以下是几种常见的基于深度学习的超分辨率重建技术及其对比。

(1) SRCNN

SRCNN(Super-Resolution Convolutional Neural Network)是首次在超分辨率重建领域应用卷积神经网络的深度学习模型。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

八分辨率图像 SRCNN的网络结构如图255-2



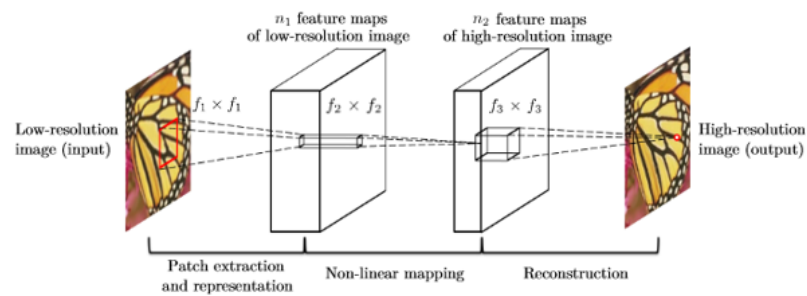


图2 SRCNN的网络结构

(2) ESPCN

与SRCNN不同，ESPCN (Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network)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，得到与目标高分率图像相同大小的低分辨率图像。如图3所示，ESPCN中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。

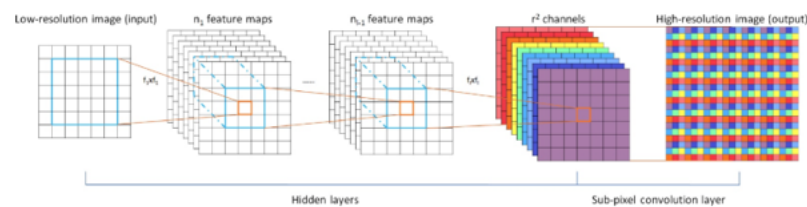


图3 ESPCN的网络结构

(3) SRGAN

与上述两种方法类似，大部分基于深度学习的图像超分辨率重建技术使用均方误差作为其网络训练过程中使用的损失函数，但是由于均方差本身的性质，往往会导致复原出的图像出现高频信息丢失的问题。而生成对抗网络（Generative Adversarial Networks, GAN）则通过其中的鉴别器网络很好的解决了这个问题，GAN的优势就是生成符合视觉习惯的逼真图像，所以SRGAN (Photo-

Generative Adversarial Network)的作者就将GAN引入了图像超分辨率重建领域。

如图4所示, SRGAN也是由一个生成器和一个鉴别器组成。生成器负责合成高分辨率图像, 鉴别器用于判断给定的图像是来自生成器还是真实样本。通过一个二元零和博弈的对抗过程, 使得生成器能够将给定的低分辨率图像复原为高分辨率图像。

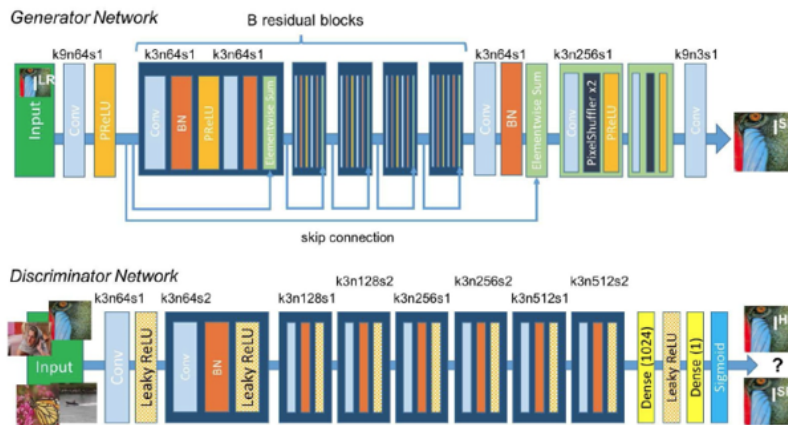


图4 SRGAN的网络结构

4 总结与展望

深度学习在图像超分辨率重建领域已经展现出了巨大的潜力, 极大的推动了该领域的蓬勃发展。但距离重建出既保留原始图像各种细节信息、又符合人的主观评价的高分辨率图像这一目标, 深度学习的图像超分辨率重建技术仍有很长的一段路要走。主要存在着以下几个问题:

(1) 深度学习的固有性的约束。深度学习存在着需要海量训练数据、高计算性能的处理器以及过深的网络容易导致过拟合等问题。

(2) 类似传统的基于人工智能的学习方法, 深度学习预先假定测试样本与训练样本来自同一分布, 但现实中二者的分布并不一定相同, 甚至可能没有相交的部分。



(3) 尽管当前基于深度学习的重建技术使得重建图像在主观评价指标上取得了优异的成绩，但重建后的图像通常过于平滑，丢失了高频细节信息。因此进一步研究基于深度学习的图像超分辨率技术仍有较大的现实意义和发展空间。

参考文献

1. Park S C, Park M K, Kang M G. Super-resolution image reconstruction: a technical overview[J]. IEEE signal processing magazine, 2003, 20(3): 21-36.
2. Kim J, Kwon Lee J, Mu Lee K. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1646-1654.
3. Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 38(2): 295-307.
4. Shi W, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1874-1883.
5. Ledig C, Theis L, Huszár F, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[C]//CVPR. 2017, 2(3): 4.



本站原创文章仅代表作者观点，不代表SDNLAB立场。所有原创内容版权均属SDNLAB，欢迎大家转发分享。但未经授权，严禁任何媒体（平面媒体、网络媒体、自媒体等）以及微信公众号复制、转载、摘编或以其他方式进行使用，转载须注明来自 SDNLAB并附上

