

基于 SRCNN 模型的图像压缩方法研究

曾诗悦*

ZENG Shi-yue

摘要

传统的图像压缩方法常采用小波变换,该方法不利于保留图像的原始细节并且清晰度也得不到满足,而 SRCNN 网络问世后,它在图像超分辨率重建的应用使得深度学习在图像处理领域得到扩展。2014 年,深度学习模型 SRCNN 一经提出,为深度学习解决图像压缩领域的像素问题开创了新纪元。本文利用深度学习技术,完成了图像的超分辨率重建,通过实验,将原始图像输入网络后不仅可以保留图像的细节信息还能将图像进行有效的压缩,最后的结果显示改进后的 SRCNN 模型在处理高亮度低纹理度图像上具有优势,并且具有良好的峰值信噪比。

关键词

图像的超分辨率重建; SRCNN 模型; 深度学习; 图像压缩

doi: 10.3969/j.issn.1672-9528.2020.09.031

0 引言

在图像压缩中,常用的图像压缩方法是 DWT 变换(离散小波变换),可是对于一些岩心图像或者医疗诊断图像,其对图像的要求较高,若是想在达到理想压缩比的同时也能获得较高的清晰度和信噪比则需采用另一种图像压缩方法。在 SRCNN 模型出现之前,应用在图像重建上的方法大多是样例学习方法,稀疏编码是其中一个比较常用的样例学习方法,想要使用这种方法进行图像重建需要收集很多的数据,数据过多又导致难以有效建模。2012 年,深度卷积神经网络^[1]引发了广大研究者的关注,DONG 等人^[2-3]将稀疏编码的训练过程与 CNN 的学习过程结合起来提出了 SRCNN 模型,2015 年 DONG 等人^[4]在 SRCNN 模型的基础上,引入了特征增强层形成一个更深层的网络来达到减少压缩伪像的目的。另外一种与 SRCNN 模型不同的模型是 ESPCN 模型^[5],使用 SRCNN 模型时要对图像进行预处理,即把图像输入训练网络之前要将原图放大到一定的倍数,而 ESPCN 模型不用将原图进行放大而是直接对原图进行卷积操作,在进行最后一步重构图像时,再利用卷积放大操作将图像进行放大。2017 年 LEDIG 等人^[6]根据生成对抗性网络(GAN)提出了 SRGAN 模型,该模型可以对大尺寸图像进行有效的重构。本文在 SRCNN 模型的基础上同时训练 YCbCr 色域中的三个通道,通过端到端的训练形式完成了低分辨率图像到高分辨率图像的非线性映射。

1 改进的 SRCNN 模型压缩图像算法

1.1 原有的 SRCNN 模型

2014 年提出的 SRCNN 模型中的 CNN 网络结构是基于样例学习的方法所构造的,但是它跟稀疏编码不同的是,CNN 网络在训练参数时,所有的参数都可以在训练过程中得到优化找到最适合的参数值,并且该方法可以利用现有的图片集(Image Net)进行训练,最后完成的重构图像比较理想。下面是 SRCNN 模型^[2]的实现思路:

图像输入网络进行运算之前要把它放大到所需大小,将原始图像通过插值(双三次插值)扩展成插值图像 Y,目标就是将 Y 经过 F(Y) 变换,实现 Y 与 X 的高度相似,因此要学习一个映射 F,该映射在概念上由图像块提取和表示、非线性映射和重建三个部分组成。

第一个部分是图像块提取和表示,图像块提取和表示类似于通过一组过滤器对图像进行卷积操作,它的第一层表示为操作 F_1 :

$$F_1(Y) = \max(0, W_1 * Y + B_1) \quad (1)$$

第二个部分为非线性映射,上一步得到的每一个 n_1 维向量对应于原始图像中的一个图像块,将提取出来的这些特征映射到 n_2 维向量中。第二层的操作是:

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2) \quad (2)$$

每个输出的 n_2 维向量都可以用来表示用于重构的高分辨率补丁。

对预测的高分辨率色块进行均匀分配^[7]是进行图像重建的常用方法,整合第二部得到的高分辨率图像块,得到最终

* 长江大学电子信息学院 湖北荆州 434023

的高分辨率重建图像:

$$F(Y) = W_3 * F_2(Y) + B_3 \tag{3}$$

将以上三个操作放在一起形成了一个卷积神经网络，如图 1 所示。

1.2 改进后的 SRCNN 模型

为了实现较好的图像重构效果，在以往的 SRCNN 模型的基础上进行了相应的改进，图 2 是改进后的 SRCNN 模型。

将大小为 $M \times N$ 的图像输入网络，网络结构分为三层，图像首先经过卷积核大小为 9×9 、数目为 64 的卷积层，经过 relu 激活函数处理后再通过第二层；第二层经过卷积核大小为 1×1 、数目为 32 的卷积层，经过 relu 激活函数处理通过第三；第三层是一个尺寸为 5×5 的卷积核，卷积核数目为 1。经过这样的网络结构后，第一层可以输出 64 张特征图，第二层可以输出 32 个特征图，第三层仅输出一张特征图。

原有的 SRCNN 模型的作者在进行实验时仅考虑亮度通道（在 YCrCb 颜色空间中），即文献 [2] 仅设置了 $C=1$ ，改进后的模型将此方法扩展为直接在彩色图像上进行训练，把 Y、Cb、Cr 三个通道的数据输入到卷积神经网络进行运算，将获得的结果合成新的 YCbCr 色域空间图像，然后转换回 RGB 三通道，最终使用生成的 RGB 图像与输入网络的原始 RGB 图像进行损失计算。

2 实验结果与数据分析

实验使用 500 张图像作为数据集，每张图像的大小为 487×321 ，选取 Adam 算法对参数进行训练优化，迭代 20 个 epoch，学习率设置为 0.1，batch_size 设置为 1，最后得到在一定压缩比下重建的图像，如图 3 所示。

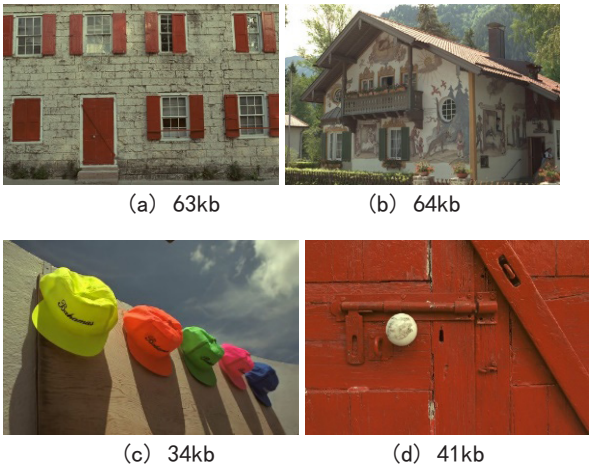


图 3 重构后的图像所占用的字节数

表 1 图像重构后各指标结果

	a	b	c	d
原始图像大小	200KB	208KB	138KB	185KB
重构图像大小	63KB	64KB	34KB	41KB
压缩率	3.1746	3.2399	4.1317	4.5121
PSNR	28.607	29.049	30.781	33.041

由图 3 和表 1 我们可以观察到，图像 c 和图像 d 相对于图像 a 和图像 b 具有高亮度低纹理的特性，因此图像 c 和 d 得到的压缩率均高于 4.1。再来观察 PSNR 值，发现亮度条件差不多的情况下，纹理复杂度越低的图像其 PSNR 值越高。

3 结论

改进的后 SRCNN 图像压缩模型与 2014 年提出的原型相比，能够得到 RGB 三通道图像，而不仅仅是灰度图像，同时使得重构的图像在一定的压缩率上具有较高的 PSNR 值（高亮度、低纹理度情况下），而且重构出来的图像也具有良好的视觉效果。

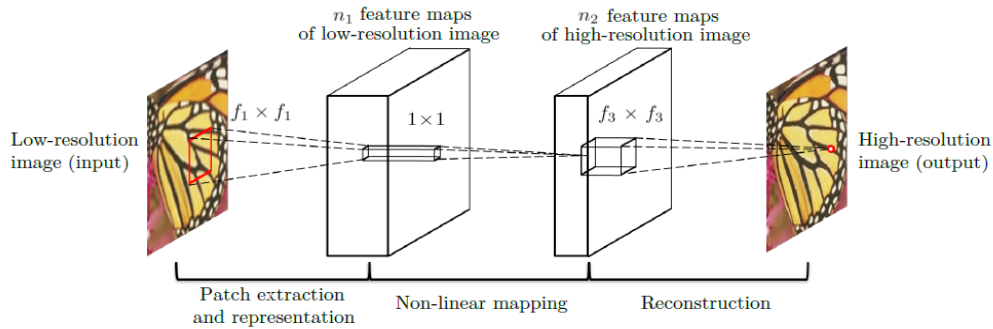


图 1 SRCNN 模型

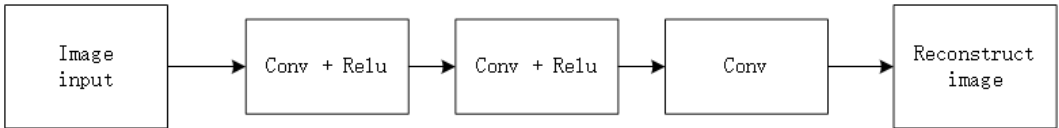


图 2 改进的 SRCNN 模型

深度学习中激活函数权重初始值的选取

孙洲鑫* 王文松

SUN Zhou-xin WANG Wen-song

摘要

在深度学习的过程中, 激活函数的选取是非常重要的, 对于不同的激活函数如何选择合适的权重初始值是本文研究的重点。本文主要就激活函数权重初始值的选取提出问题, 根据相关资料的查询选择几个较为特殊的权重初始值在不同激活函数上进行验证实验, 得出不同激活函数对应的最佳的权重初始值的组合。

关键词

深度学习; 激活函数; 权重初始值; MNIST 数据集; sigmoid; ReLu; tanh

doi: 10.3969/j.issn.1672-9528.2020.09.032

0 引言

众所周知, 在最近几年里, 深度学习 (Deep Learning, DL)^[1] 的研究是最热的方向之一。深度学习是机器学习研究的一个分支, 在机器学习中, 参数的设定是通过大量事例进行类比得出, 深度学习则是通过人工神经网络 (Artificial Neural Networks, ANNs) 通过不同隐藏层^[2] 的学习进而得出最优的参数组合。

在深度学习的过程中, 重点就是激活函数^[3] 和权重^[4]

的选取, 合理的组合会使得学习速率加快, 同时可以降低损失函数的值。激活函数的选取一般都是根据实际的情况来做一定的改变。权重的参数是根据大量的自学习进而得到有效的参数组合, 对于不同的激活函数, 如何选择合适的权重初始值就成了本文所要研究的重点。

1 权重初始值的设定

在深度学习的过程中, 常用的激活函数主要有 sigmoid() 函数^[5] 如式 (1) tanh() 函数^[6] 如式 (2) 线性整流函数^[7] (Rectified Linear Unit, ReLU) 如式 (3) 等, 其数学表达式分别如下:

* 长江大学 湖北荆州 434023

参考文献:

- [1]KRIZHEVSKY A,SUTSKEVER I,HINTON G E. Image net classification with deep convolution neural networks[C]. International Conference on Neural Information Processing Systems,2012:1097-1105.
- [2]DONG C,CHEN C L,HE K, et al. Learning a deep convolution network for image super-resolution[C].European Conference on Computer Vision,2014:184-199.
- [3]DONG C,LOY C C,HE K, et al. Image super-resolution using deep convolution networks[J].IEEE Transactions on Pattern Analysis and Machine Intelligence,2016,38(2):295-307.
- [4]DONG C,LOY C,YUBIN D et al. Compression Artifacts Reduction by a Deep Convolutional Network[C].Proceeding of the IEEE International Conference on Computer Vision,2015:576-584.

- [5]WENZHE S,JOSE CABALLERO,ROB B et al. Real-Time Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolution Neural Network[C].2016 IEEE Conference on Computer Vision and Pattern Recognition(CV PR),2016:1874-1883.
- [6]LEDIG C,THEIS L,HUSZAR F et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[C].2017 IEEE Conference on Computer Vision and Pattern Recognition,2017:105-114.
- [7]Aharon M. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. TSP 2006,54(11):4311-4322.

(收稿日期: 2020-07-19 修回日期: 2020-08-11)