

基于 CNNC 的卷积神经网络图像的压缩方法

崔建良 李建飞 陈春晓[△] 姜睿林

(南京航空航天大学生物医学工程系 南京 211106)

摘要: 图像压缩是提高图像存储效率以及实现高速高效传输的前提。根据神经网络的基本结构和算法,设计并搭建了基于卷积神经网络的 CNNC(convolutional neural network compression ,CNNC) 图像压缩模型。该模型通过卷积层和池化层构成自编码器,反卷积层和卷积层构成自解码器,实现了图像编码压缩和解码重建的功能,并通过 Set12 数据集验证了 CNNC 图像压缩模型。实验结果表明,当压缩比较低时,JPEG 压缩方法与 CNNC 压缩方法无显著差异;当压缩比较高时,CNNC 压缩方法有明显的优势,在压缩比高达 128 时,CNNC 压缩方法重建结果仍然很好。Set12 数据集实验验证了 CNNC 压缩模型的有效性。

关键词: 图像压缩; 自编码器; 卷积神经网络; 深度学习; 图像重建

中图分类号: R318 **文献标识码:** A **文章编号:** 1672-6278 (2019)04-0415-05

Image compression method based on convolutional neural network compression

CUI Jianliang , LI Jianfei , CHEN Chunxiao , JIANG Ruilin

(Department of Biomedical Engineering , Nanjing University of Aeronautics and Astronautics , Nanjing 211106 , China)

Abstract: Image compression is the premise to improve the efficiency of image storage and realize high-speed and efficient transmission. According to the basic structure and algorithm of neural network, a convolutional neural network compression (CNNC) model was designed and built in this paper. The self-encoder was composed of convolution layer and pooling layer, deconvolution layer and convolution layer constituted self-decoder. The function of image coding compression and decoding reconstruction was realized. The CNNC image compression model was validated by Set12 data set. The experimental results showed that when the compression ratio was low, there is no significant difference between JPEG compression method and CNNC compression method; when the compression ratio was high, CNNC compression method had obvious advantages, and when the compression ratio was up to 128, the reconstruction result of CNNC compression method was still very good. Set12 data set experiment verifies the validity of CNNC compression model.

Key words: Image compression; Self-encoder; Convolutional neural network; Deep learning; Image re-construction

1 引言

图像数据间存在大量的冗余,图像压缩是减少数据冗余、提高图像存储效率以及实现高速高效传输的技术。目前常用的图像压缩方法有预测编

码^[1]、金字塔方法^[2]、矢量编码^[3]、JPEG^[4-5]和神经网络编码^[6]。

JPEG2000 是最常用的图像压缩方法,小波变换是 JPEG2000 压缩方法的核心算法。针对小波变换的编码效率问题,詹为等^[7]提出了改进的嵌入式零

DOI 10.19529/j.cnki.1672-6278.2019.04.07

[△]通信作者 Email: ccxbme@nuaa.edu.cn

树小波算法,提高了编码效率和峰值信噪比(peak signal-to-noise ratio, PSNR);针对小波变换重构质量较差的问题,王林等^[8]提出基于谱图小波变换的编码方法,PSNR 和压缩比也得到了有效提升。随着图像数据的种类和数量愈来愈大,图像压缩方法面临新的挑战。近年来,深度学习在图像处理领域取得了很多成果。任杰^[9]提出了一种卷积神经网络与 JPEG 压缩相结合的编解码框架,可以提高图像的压缩比和视觉质量。范靓^[10]结合卷积神经网络与分形压缩算法,有效缩短了图像的编码时间,提高了解码图像的质量。Jiang 等^[11]提出了一种基于端到端的压缩框架,它通过与 JPEG 等编解码器相结合,实现图像压缩的功能,取得了较好的效果。深度学习技术给图像压缩方法的研究带来了新的契机。以上深度学习方法虽然取得了较高的重建图像的质量,但均无大幅度提高压缩率。对于同时提高重建效果和压缩率的问题,本研究针对基于卷积神经网络的深度学习方法,设计和构建了 CNNC 图像压缩模型,在保证图像质量的前提下,进一步提高图像的压缩比和 PSNR 等图像压缩评价指标。

2 卷积神经网络

卷积神经网络一般包含输入层、卷积层、池化层、全连接层和输出层^[12],不同的网络所包含的层数及各种层所占的比例不同。卷积神经网络的基本模型见图 1。

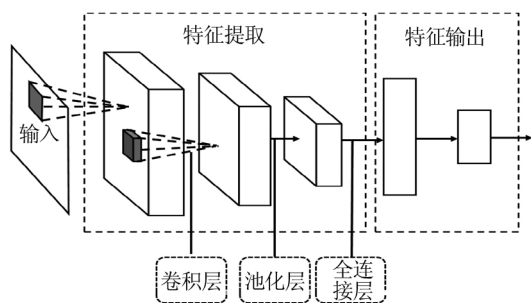


图 1 卷积神经网络基本模型

Fig. 1 Basic model of convolutional neural

卷积层可以自动提取输入信号的深层信息,具有局部感受野和权值共享的特点。池化层通过执行下采样操作,可以有效地控制过拟合。全连接层是指下层的每个神经元根据权重接受上层网络中每个神经元传递的特征值,并利用偏置与激活函数提取特征,得到预期的结构。

3 基于卷积神经网络的 CNNC 图像压缩模型

基于卷积神经网络的 CNNC 图像压缩模型的结构见图 2。网络主要由输入层、编码层、解码层和输出层组成,其中编码层由卷积层和池化层组成,解码层由卷积层和反卷积层组成。

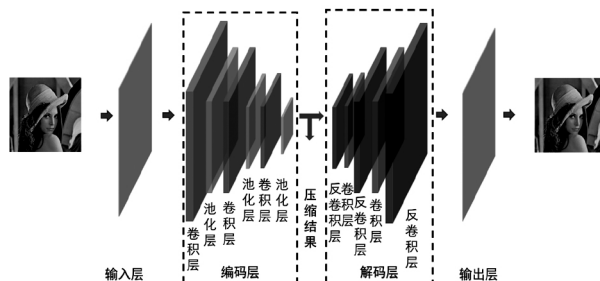


图 2 CNNC 压缩模型网络结构

Fig. 2 Network structure of CNNC compression

3.1 输入层

设置批处理量为 50。图像像素为 $M \times N$ 的数据集进入到输入层之前,数据的格式就被处理为 $[50, M, N]$ 的三维矩阵形式。三维矩阵进入输入层以后,在输入层中进行预处理,图像由原来的 $[M, N]$ 二维矩阵变为 $[M, N, 1]$ 的三维矩阵,数据格式也从 $[50, M, N]$ 的三维矩阵转变为 $[50, M, N, 1]$ 的四维张量。

3.2 编码层

编码层执行数据的编码过程。CNNC 中编码层包含了多个卷积池化操作,并定义每次卷积都提取 32 张特征图。因此第一次卷积操作后,数据格式变为 $[50, M, N, 32]$,而后进行池化操作,特征图尺寸为 $\frac{M}{2} \times \frac{N}{2}$,数据格式变为 $[50, \frac{M}{2}, \frac{N}{2}, 32]$ 。经过 n 次卷积池化操作后,数据格式为 $[50, \frac{M}{2^n}, \frac{N}{2^n}, 32]$,特征图尺寸为 $\frac{M}{2^n} \times \frac{N}{2^n}$,共有 32 张。由于特征图的数据格式是 float32 型,每个像素占 4 个字节,因此,原图被压缩了 2^{2n-7} 倍, n 为卷积池化层的数量。

3.3 解码层

解码层执行图像的还原过程,其组成为“反卷积—卷积—……—反卷积—卷积—反卷积”。反卷积执行的是图像尺寸恢复,每经过一次反卷积层,特征图尺寸放大 4 倍,特征图数量不变。由于本研究的 CNNC 网络反卷积层尺寸放大的方式是添 0 扩展,为了加快网络的训练和拟合程度,在每一次反卷积后增加了一层卷积层。最后一层反卷积层除了

放大特征图尺寸外,还会将所有的特征图整合为一张图,并传递到输出层。数据在解码层中的格式变换为: $[50, \frac{M}{2^n}, \frac{N}{2^n}, 32] \rightarrow [\dots] \rightarrow [50, \frac{M}{2}, \frac{N}{2}, 32] \rightarrow [50, M, N, 1]$ 。

3.4 输出层

输出层除了执行数据的输出外,还会对数据进行后处理。输出层后处理的方式是将数据格式从 $[50, M, N, 1]$ 的四维张量转变为 $[50, M, N]$ 的三维矩阵。至此,数据的一次前向传播结束,然后通过损失函数和优化函数进行反向传播。

4 实验结果与分析

4.1 实验环境及参数

本实验采用 NVIDIA GTX1060 6 G 显存,16 G 内存, Tensorflow - GPU1.11.0 版本在 Python3.6.6

模式下运行,编译环境为 Pycharm。

4.2 数据集选择

训练集采用自己制作的数据集,数据集中的图像包含风景、动物等,图像大小为 256×256 ; 测试集采用 Set12 数据集,数据集中包含 Cameraman、House、Pepper、Fishstar、Monarch、Airplane、Parrot、Lena、Barbara、Ship、Man、Couple 共 12 张图片。

4.3 实验结果

实验中,分别设置 CNNC 网络中卷积层、池化层的层数 $n = 4, 5, 6$ 此时 CNNC 压缩比为 2, 8, 32。JPEG 压缩比分别为 2.106 ± 0.281 , 7.993 ± 0.125 , 32.322 ± 1.691 。CNNC 训练 500 次后,对 Set12 数据集进行测试,得到的结果与 JPEG 结果进行对比,评价指标采用 PSNR 和结构相似性 (structural similarity index, SSIM), 评价结果见表 1。

表 1 JPEG 和 CNNC 对 Set12 数据集的压缩质量比较

Table 1 Compression quality comparison between JPEG and CNNC for Set12 data sets

名称	压缩比为 2				压缩比为 8				压缩比为 32			
	JPEG		CNNC		JPEG		CNNC		JPEG		CNNC	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Cameraman	54.64	0.999	29.37	0.920	31.04	0.915	28.99	0.914	22.26	0.613	29.07	0.915
House	56.78	0.999	36.57	0.954	42.30	0.985	36.06	0.951	25.02	0.739	36.42	0.952
Pepper	55.03	0.999	29.25	0.941	33.11	0.931	29.95	0.942	22.41	0.627	28.29	0.935
Fishstar	54.63	0.999	31.52	0.940	31.62	0.922	31.04	0.935	21.27	0.550	31.30	0.937
Monarch	54.75	0.999	31.10	0.946	31.58	0.935	30.68	0.942	21.11	0.625	30.94	0.943
Airplane	54.48	0.999	29.32	0.933	30.67	0.924	28.94	0.928	21.30	0.642	28.64	0.943
Parrot	53.33	0.999	29.18	0.935	31.14	0.926	28.83	0.931	22.05	0.634	29.11	0.933
Lena	56.07	0.999	33.35	0.953	33.79	0.942	32.93	0.948	22.82	0.616	33.22	0.951
Barbara	55.09	0.999	30.18	0.893	31.76	0.900	30.07	0.891	22.13	0.563	30.14	0.892
Ship	54.76	0.999	31.83	0.934	32.06	0.919	31.52	0.929	21.97	0.536	31.56	0.929
Man	54.69	0.999	32.13	0.925	32.08	0.908	31.88	0.921	22.20	0.513	32.04	0.924
Couple	54.70	0.999	31.95	0.937	32.18	0.921	31.62	0.932	22.02	0.510	31.66	0.932
平均	54.91	0.999	31.31	0.934	32.77	0.927	31.04	0.930	22.21	0.597	31.03	0.932

由表 1 可以看出,压缩比为 2 时, CNNC 压缩方法的 PSNR 和 SSIM 低于 JPEG 压缩,此时 JPEG 压缩接近无损压缩, $SSIM \approx 1$; 压缩比为 8 时, CNNC 压缩方法各指标变化不大,但 JPEG 压缩方法各个指标性能明显下降; 压缩比为 32 时, CNNC 压缩方法

的 PSNR 和 SSIM 都远高于 JPEG 压缩方法。各种压缩比情况下,图像重建结果见图 3,可以看出,压缩比为 32 时, JPEG 压缩方法出现了明显的失真,但是 CNNC 压缩方法重建的图像质量仍然较好。

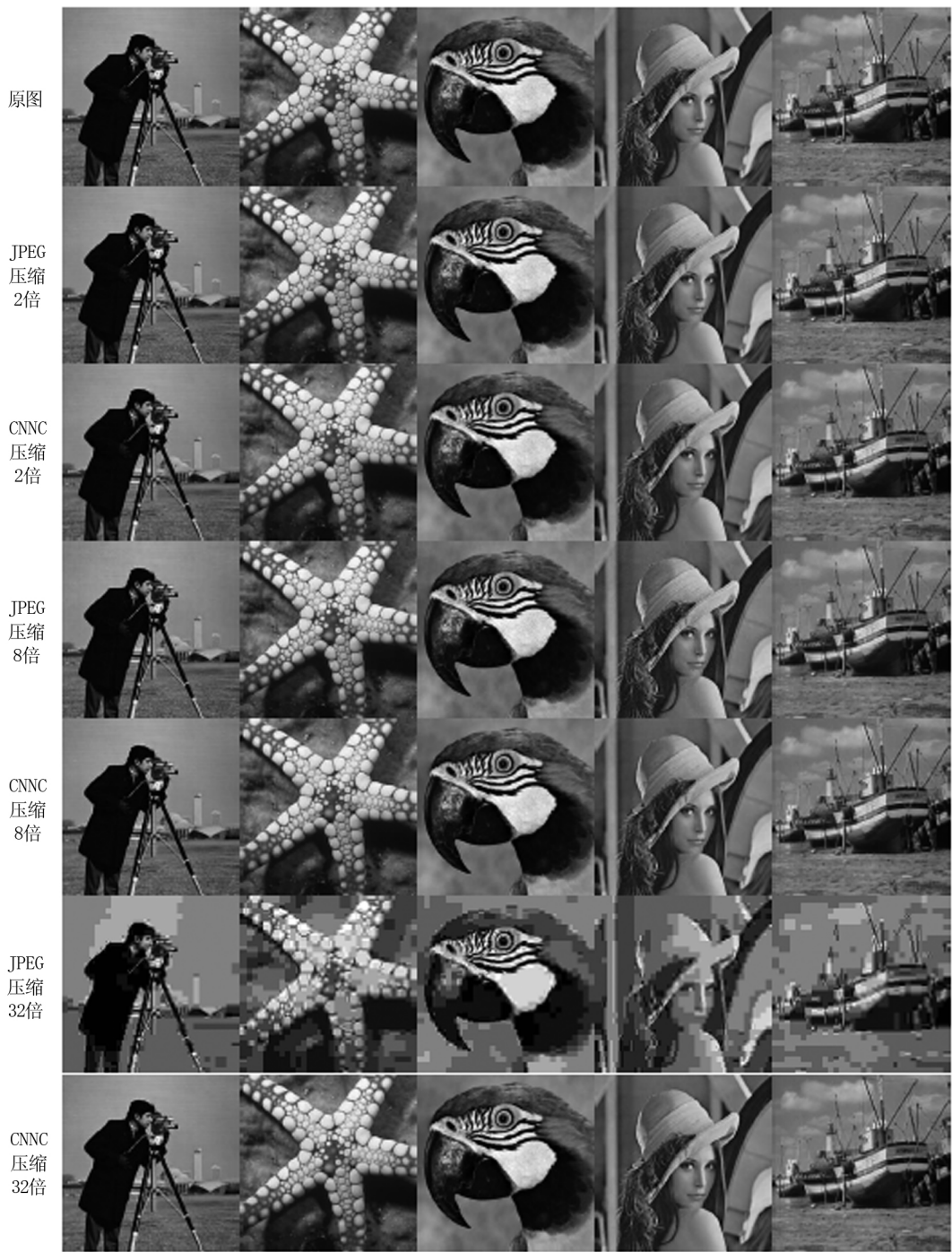


图3 CNNC 与 JPEG 在不同压缩比下的结果比较

Fig.3 Comparison of CNNC and JPEG results at different compression ratios

为进一步研究 CNNC 的压缩能力,设置卷积池化层层数 $n=7$,此时 CNNC 压缩比为 128,而 JPEG 已无法达到 128 倍的压缩比。在 CNNC 训练 500 次后,对 Set12 数据集进行测试,结果见表 2。表 2 中,Set12 数据集的 PSNR 为 31.092 ± 2.176 ,SSIM 为 0.932 ± 0.016 。由此可以看出,在压缩比为 128 时,CNNC 压缩方法的各个性能指标仍然较好。CNNC 压缩 128 倍时重建结果见图 4。

表 2 CNNC 对 Set12 数据集的压缩结果(压缩比 128)

Table 2 Compression results of CNNC on Set12 datasets (compression ratio 128)

名称	CNNC		名称	CNNC	
	PSNR	SSIM		PSNR	SSIM
Cameraman	29.134	0.917	Parrot	29.074	0.933
House	36.329	0.953	Lena	33.143	0.951
Pepper	28.793	0.937	Barbara	30.062	0.891
Fishstar	31.304	0.937	Ship	31.600	0.931
Monarch	30.977	0.944	Man	32.040	0.923
Airplane	28.938	0.930	Couple	31.705	0.934

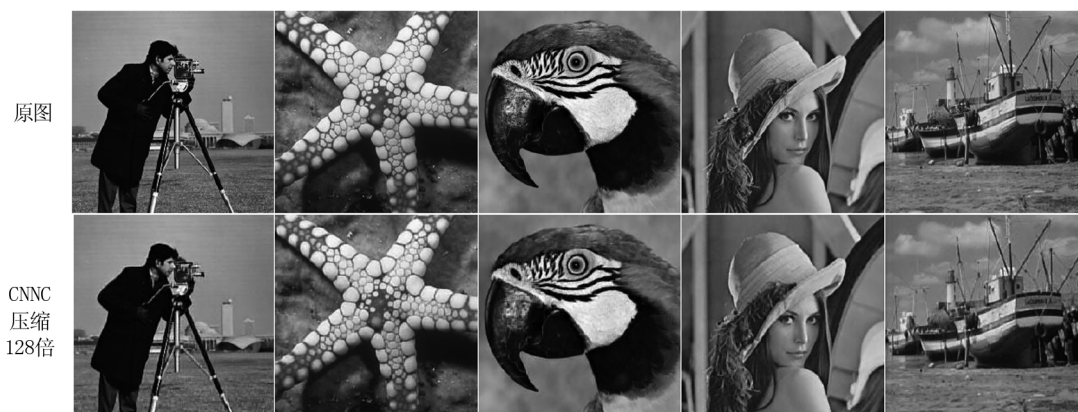


图4 CNNC 压缩结果重建图

Fig.4 Reconstruction of CNNC compression result

研究结果显示,当压缩比较低时,JPEG 压缩方法与 CNNC 压缩方法无显著差异,但当压缩比较高时,CNNC 压缩方法有明显的优势,当压缩比为 128 时,重建结果仍然较好。

5 结论

本研究设计并搭建了基于卷积神经网络的 CNNC 图像压缩模型,通过多个卷积层和池化层对图像进行压缩,并利用多个反卷积层和卷积层对压缩图像进行解码。结果表明,本研究提出的 CNNC 模型在高压比的情况下,仍可保持较高的 PSNR 和 SSIM,图像失真小,压缩性能明显高于传统的 JPEG 压缩方法。

参考文献:

- [1]高健,沈辉,郭梦琦.一种新的图像压缩预测编码[J].计算机应用与软件,2008,25(10):231-233.
- [2]Venkata P K M,Varma K C R C, Mahapatra S. Pyramid coding based rate control for constant bit rate video streaming[J]. Multimedia Tools and Applications, 2015, 75(24): 671-680.
- [3]刘丽丽,陈贺新,桑爱军,等.基于多维矢量矩阵正交变换理论的彩色图像压缩编码[J].吉林大学学报(工学版),2011,41

(6): 1754-1759.

- [4]Guo L J, Ni J Q, Shi Y Q. Uniform embedding for efficient JPEG steganography[J]. IEEE Transactions on Information Forensics & Security, 2014, 9(5): 814-825.
- [5]Acharya T, Tsai P S. JPEG - still image compression standard [M]// JPEG2000 Standard for Image Compression: Concepts, Algorithms and VLSI Architectures. 2005.
- [6]Ballé J, Laparra V, Simoncelli E P. End-to-end optimized image compression[C]// ICLR 2017. 2016.
- [7]詹为,段先华,於跃成.基于小波变换的图像压缩编码方法研究[J].计算机技术与发展,2018,28(6):21-25.
- [8]王林,宋星.基于谱图小波变换的图像压缩编码方法[J].计算机系统应用,2018,27(5):176-180.
- [9]任杰.基于深度学习的图像压缩方法研究[D].哈尔滨工业大学,2017.
- [10]范靛.基于遗传算法和深度学习的分形图像压缩算法的研究[D].内蒙古农业大学,2016.
- [11]Jiang F, Tao W, Liu S, et al. An end-to-end compression framework based on convolutional neural networks[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017: 1-1.
- [12]He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2016: 770-778.

(收稿日期:2019-05-27)