REDUCTION OF COMPRESSION ARTIFACTS IN DIGITAL IMAGES USING ARTIFACTS REDUCTION CONVOLUTIONAL NEURAL NETWORK

R.Balachander, Research Scholar, G.Sakthivel, Associate Professor,
Department of Electronics and Instrumentation engineering,
Annamalai University, India

Abstract

The Lossy compression introduces complex compression artifacts, particularly blocking artifacts, ringing effects and blurring. Existing algorithms either focus on removing blocking artifacts and produce blurred output, or restore sharpened images that are accompanied with ringing effects. Inspired by the success of deep convolutional networks (DCN) on super-resolution, This work formulate a compact and efficient network for seamless attenuation of different compression artifacts. The Artifacts Reduction Convolutional Neural Network is the innovative method to reduce the artifacts appears during image compression. The parameters used to evaluate the results are PSNR, PSNR-B, SSIM. This ARCNN achieves 93% of structural similarity which is best than any other model.

Keywords - Image compression, Joint picture expert group, Peak signal to noise ratio.

I. INTRODUCTION

Image compression is the application of data compression on digital images. In effect, the objective is to reduce redundancy of the image data in order to be able to store or transmit data in an efficient form. Uncompressed multimedia (graphics, audio and video) data requires considerable storage capacity and transmission bandwidth. Despite rapid progress in mass-storage density, processor speeds and digital communication system performance, demand for data storage capacity , data transmission bandwidth continues to outstrip the capabilities of available technologies. The recent growth of data intensive multimedia-based web applications have not only sustained the need for more efficient ways to encode signals and images but have made compression of such signals central to storage and communication technology.

II. PRINCIPLES BEHIND COMPRESSION

A common characteristic of most images is that the neighboring pixels are correlated and therefore contain redundant information. The foremost task is to find less correlated representation of the image. Two fundamental components of compression are redundancy and irrelevancy reduction. Redundancy reduction aims at removing duplication from the signal source (image/video). Irrelevancy reduction omits parts of the signal that will not be noticed by the signal receiver, namely the Human Visual System. In general, three types of redundancy can be identified:

A. Coding Redundancy

A code is a system of symbols used to represent a body of information or set of events. Each piece of information or events is assigned a sequence of code symbols, called a code word. The number of symbols in each code word is its length. The 8-bit codes that are used to represent the intensities in the most 2-D intensity arrays contain more bits than are needed to represent the intensities.

B. Spatial Redundancy and Temporal Redundancy

The pixels of most 2-D intensity arrays are correlated spatially, information is unnecessarily replicated in the representations of the correlated pixels. In video sequence, temporally correlated pixels also duplicate information.

C. Irrelevant Information

Most of the 2-D intensity arrays contain information that is ignored by the human visual system and extraneous to the intended use of the image. It is redundant in the sense that it is not used. Image compression research aims at reducing the number of bits needed to represent an image by removing the spatial and spectral redundancies as much as possible.

III. LOSSY COMPRESSION ARTIFACTS

A.Blocking Artifacts

JPEG image is compressed by 8×8 non-overlapping blocks. Blocking artifacts are the discontinuities along the block boundaries of 8×8 blocks

B.Ringing Artifacts along the sharp edges

To efficiently compress an image, quantization of high frequency components is done to remove some high frequency signals from the images. However, when the edges are sharp, there are ringing artifacts like a wave near the sharp edges when quantization is too strong.

C.Blurring

Loss of high frequency components also introduces blurring. These artifacts affect other routines such as super-resolution and edge detection.

IV. DEEP CONVOLUTIONAL NEURAL NETWORK

Deep learning architectures such as deep neural networks, deep belief networks and recurrent neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design and board game programs, where they have produced results comparable to and in some cases superior to human.

A.Working principle of Deep Learning

Computer programs that use deep learning go through much the same process. Each algorithm in the hierarchy applies a nonlinear transformation on its input and uses what it learns to create a statistical model as output. Iterations continue until the output has reached an acceptable level of accuracy. The number of processing layers through which data must pass is what inspired the label deep. n traditional machine learning, the learning process is supervised and the programmer has to be very, very specific when telling the computer what types of things it should be looking for when deciding if an image contains a dog or does not contain a dog. This is a laborious process called feature extraction and the computer's success rate depends entirely upon the programmer's ability to accurately define a feature set for "dog." The advantage of deep learning is that the program builds the feature set by itself without supervision. Unsupervised learning is not only faster, but it is usually more accurate.

Initially, the computer program might be provided with training data, a set of images for which a human has labeled each image "dog" or "not dog" with meta tags. The program uses the information it receives from the training data to create a feature set for dog and build a predictive model. In this case, the model the computer first creates might predict that anything in an image that has four legs and a tail should be labeled "dog." Of course, the program is not aware of the labels "four legs" or "tail;" it will simply look for patterns of pixels in the digital data. With each iteration, the predictive model the computer creates becomes more complex and more accurate. Because this process mimics a system of human neurons, deep learning is sometimes referred to as deep neural learning or deep neural networking. Unlike the toddler, who will take weeks or even months to understand the concept of "dog," a computer program that uses deep learning algorithms can be shown a training set and sort through millions of images, accurately identifying which images have dogs in them within a few minutes.

To achieve an acceptable level of accuracy, deep learning programs require access to immense amounts of training data and processing power, neither of which were easily available to programmers until the era of big data and cloud computing. Because deep learning programming is able to create complex statistical models directly from its own iterative output, it is able to create accurate predictive models from large quantities of unlabeled, unstructured data. This is important as the Internet of Things (IoT) continues to become more pervasive, because most of the data humans and machines create is unstructured and is not labeled.

Use cases today for deep learning include all types of big data analytics applications, especially those focused on Natural Language Processing (NLP), language translation, medical diagnosis, stock market trading signals, network security and image identification.

V. INTRODUCTION OF CONVOLUTIONAL NETWORKS

In this work using network in which adjacent network layers are fully connected to one another. That is, every neuron in the network is connected to every neuron in adjacent layers.

In particular, for each pixel in the input image, encoded the pixel's intensity as the value for a corresponding neuron in the input layer. For the 28×28 pixel images it has been using, this means our network has 784 (=28×28) input neurons. Then trained the network's weights and biases.

Our earlier networks work pretty well: It is obtained classification accuracy better than 98 percent, using training and test data from the MNIST handwritten digit data set. But upon reflection, it's strange to use networks with fully-connected layers to classify images. The reason is that such a network architecture does not take into account the spatial structure of the images. For instance, it treats input pixels which are far apart and close together on exactly the same footing. Such concepts of spatial structure must instead be inferred from the training data.

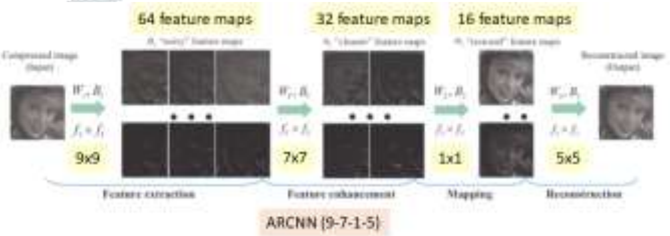
VI. ARTIFACTS REDUCTION CNN (ARCNN)

ARCNN is used to reduce the following image artifacts:

Blocking Artifacts: JPEG image is compressed by 8×8 non-overlapping blocks. Blocking artifacts are the discontinuities along the block boundaries of 8×8 blocks.

Ringing Artifacts along the sharp edges: To efficiently compress an image, quantization of high frequency components is done to remove some high frequency signals from the images. However, when the edges are sharp, there are ringing artifacts like a wave near the sharp edges when quantization is too strong.

Blurring: Loss of high frequency components also introduces blurring. These artifacts affect other routines such as super-resolution and edge detection.



AR-CNN consists of four layers, namely the feature extraction, feature enhancement, mapping and reconstruction layer.

VII.PARAMETERS

A.Structural Similarity Index

The SSIM formula is based on three comparison measurements between the samples of x and y: luminance (l), contrast (c) and structure (s). The individual comparison functions are

B.Peak signal-to-noise ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity.

$$PSNR = 10log_{10}\frac{MAX^2}{MSE}$$

C.PSNR-B

Peak signal-to-noise ratio added with Blocking effect factor results in PSNR-B.*Blocking Effect Factor* is a high frequency noise which is less visible in highly detailed areas but very visible in the smooth regions.

$$PSNR - B(x,y) = 10log_{10}\frac{MAX^2}{MSE - B(x,y)}$$

VIII.EXPERIMENT

A.LIVE1 DATASET

B.CLASSICAL DATASET IMAGE

C. AVERAGE RESULTS OF LIVE1 DATASET

Parameters	Quality	Jpeg	ARCNN
PSNR	20 30 40	28.063635 29.457774 30.486149	29.847724 31.457271 32.461254
SSIM	20 30 40	0.866455 0.902583 0.921592	0.896317 0.928708 0.942917
PSNR-B	20 30 40	25.156030 26.543923 27.656231	29.219181 30.791400 31.904289

D.AVERAGE RESULTS OF CLASSICAL DATASET

Parameters	Quality	Jpeg	ARCNN
PSNR	20 30 40	30.493456 31.831258 32.753215	31.588737 32.934269 33.712266
SSIM	20 30 40	0.850902 0.878414 0.893191	0.866819 0.892044 0.902908
PSNR-B	20 30 40	27.953935 29.354499 30.379653	31.161171 32.579682 33.386919

IX.CONCLUSION

Applying deep model on low-level vision requires more layers. More the layers in the convolutional network , more accurate results will be produced.SSIM is effective way to find the difference between original image and the obtained restored image. In this work up to 94% structural similarity were produced, which shows the originality of the restored image.

IX.REFERENCES

[1] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *TPAMI*, 33(5):898–916, 2011.

[2] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. A. Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In *BMVC*, 2012.

[3] P. Doll´ar and C. L. Zitnick. Structured forests for fast edge detection. In *ICCV*, pages 1841–1848. IEEE, 2013.

[4] C. Dong, C. C. Loy, K. He, and X. Tang. Image super-resolution using deep convolutional networks. *arXiv:1501.00092*, 2014.

[5] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In *ECCV*, pages 184–199. 2014.

[6] A. Foi, V. Katkovnik, and K. Egiazarian. Point-wise shape adaptive DCT for high-quality de-noising and de-blocking of gray-scale and color images. *TIP*, 16(5):1395–1411, 2007.

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *CVPR*, pages 580–587. IEEE, 2014

[8] M. A. Gluck and C. E. Myers. Hippocampal mediation of stimulus representation: A computational theory. *Hippocampus*, 3(4):491–516, 1993.

[9] R. C. Gonzalez and R. E. Woods. Digital image processing, 2002.

[10] K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassing human-level performance on image net classification. *arXiv:1502.01852*, 2015.

[11] V. Jain and S. Seung. Natural image de-noising with convolutional networks. In *NIPS*, pages 769–776, 2009.

[12] J. Jancsary, S. Nowozin, and C. Rother. Loss-specific training of non-parametric image restoration models: A new state of the art. In *ECCV*, pages 112–125. 2012.