Research article Available online www.ijsrr.org ISSN: 2279–0543

International Journal of Scientific Research and Reviews

A Novel Approach for Image Compression Using the Concept of Dct nd Gan

Akhilesh Sharma[1] and Tripti Sapaha[2*]

[1] Assistant Professor, Chouksey Engineering College,Bilaspur, Chhattisgah, India
[2*] M. Tech Scholar, Chouksey Engineering College,Bilaspur, Chhattisgah, India

ABSTRACT:

As we know currently we are living in the era of Internet world, where we have to transfer video/images in high speed, so there is need of good quality level compression unit which is capable to compress the image with good quality level. In this paper, it is discussed why the concept of compressing data using generative models should be used and a suggestion that it could be a direction which could lead to produce more accurate and visually pleasing reconstructions at much deeper levels of compression. Here an approach based on the concepts of Generative Adversarial Network(GAN) and Discrete Cosine Transforms(DCT) is used to generate a good quality compressed image. Here we use Python as a core language.

KEYWORDS: Image compression, DCT, GAN, DIP, OpenCVComputer Vision.

***Corresponding author**

Miss. Tripti Sapaha

M.Tech Scholar,

Department of Computer Science & Engineering

Chouksey Engineering College,

Bilaspur , Chhattisgarh, India

Email: tri001pti@gmail.com

Contact: 9131060969

INTRODUCTION:

Data Science is a multi-disciplinary field that utilizes logical strategies, procedures and frameworks to extricate learning and bits of knowledge from organized and unstructured data[1,2]. Data science is a similar idea as information mining and huge information: "utilize the most dominant equipment, the most dominant programming frameworks, and the most proficient calculations to tackle issues". Data Science is an "idea to bring together insights, information examination, AI and their related techniques" so as to "comprehend and investigate real wonders" with information. Picture preparing is broadly utilized these days to get bits of knowledge from Image Data. Large abundance of Image Data present everywhere demands for analysis of this data. Using Image Processing techniques, different models can be developed to automate different processes. One such application of Image Processing is in Intelligent Character Recognition which is automated extraction of data from handwritten forms in scanned jpg/png/tif format. There is very important role of data science in image processing. In present universe of innovation and information sciences, regularly greater part of information is being produced. Transmission of such enormous measure of information requires a nature of administration. So here comes the requirement for smooth information pressure. Rich corruption is a nature of-administration term used to portray the thought, as data transfer capacity drops or transmission mistakes happen, client experience falls apart however keeps on being important. Picture pressure is the way toward lessening picture record size in bytes without debasing the nature of picture to an unsatisfactory level. Conventional pressure procedures, for example, Joint Photographic Experts Group(JPEG), are freethinker to the information being compacted and don't corrupt nimbly. Expanding upon the thoughts of [3] and the ongoing guarantee of profound generative models [4] In this paper basically we proposed a novel approach which is able to compress the image by using of GAN & DCT approach. The remaining paper is sub categorized as follows. Necessary background and underlying principle of Image compression Algorithm is given in Section II whereas Section III demonstrates Implementation of proposed approach. Experimental results and its analysis are given in Section IV and last one, in Section Vis conclusion which wraps up the entire paper.

LITERATURE REVIEW:

In this section we study about the previous research in area of image compression but before that here we try to explain what is image compression: An image can be represented as an m by n matrix, where m, the number of rows, is the pixel height of the image, and n, the number of columns, is the pixel width of the image. When a computer creates or stores an image, each and every pixel is assigned a number to represent its relative darkness or brightness [3]. Each value contained inside the matrix decides the brightness of the corresponding displayed pixel. In case of a grayscale image the range of values within the matrix is from 0 (black) to 1 (white), where each number simply represents light or dark. But for colour images, which are much more space-consuming, the computer must split the image into three layers composed of red, green and blue in the image. Each single-colour picture is then calculated much like a grayscale picture based on darkness, and then recombined at the end to reproduce the original image. In other words, in case of a colour image, each colour pixel is broken down to three primary components: red, green and blue(RGB) [4]. Hence there are three values associated with each pixel, each ranging from 0 (colour is absent) to 1 (completely saturated). These 3 values are assigned to red, green and blue respectively. For instance, if a 9 megapixel grayscale image is considered, it can be represented as a 3000 × 3000 pixel s matrix [4]. Since it is a grayscale image, each pixel in the image matrix can be represented by a certain integer whose value falls between 0 and 255 [4]. If a storage space of 1 byte is assigned to each pixel, then the entire image requires approximately 9MB space. For a colour image, this storage space value is larger since it has three components, red, green and blue (RGB). Each component is represented by a matrix, so storing color images takes three times the space (27Mb)[4].

There are many researches done in the area of image compression. Some of the previous approaches are discussed below:

1. DCT Based [5,6,7]
2. Fractal Based [8,9,10]
3. SVD Based [11,12]
4. K-MEAN Based [13,14,15]
5. CNN Based [21,22]
6. GAN Based [23,24,25,26,27,28]
7. RNN Based [29,30,31]

Discrete Cosine Transform (DCT) tries to de-relate the photo data. After de-relationship each change coefficient can be encoded unreservedly without losing weight efficiency. This region portrays the DCT and a part of its basic properties. In reference to figure from introduction, remembering the true objective to achieve awesome weight execution, connection between the shading parts is first lessened by changing over the RGB shading space into a de related shading space. In standard JPEG, a RGB picture is first changed into a luminance-chrominance shading space, for instance, YCbCr. The upside of changing over the photo into luminance-chrominance shading space is that the luminance and chrominance portions are especially de related between each other [5,6], DCT based approach is not upto the mark but in terms of time consumption is a good approach. Fractal image compression (FIC) is one of the important methods of gray scale image coding, fully automated by Jacquin [8]. The basic FIC is based on the observation that images usually exhibit affine redundancy. Here an image is segmented into a number of different sized blocks and the encoding process consists of approximating the small image blocks, called Range blocks (RBs), from the larger blocks, called Domain blocks (DBs), of the image, by searching the best matching affine transformation from a DB pool, much akin to the image compression by vector quantization (VQ) method [9]. In the encoding process, separate transformations for each RB are obtained. For decoding, the set of affine transformations, when iterated upon arbitrary initial image, produces a fixed point (attractor) that approximates back the target image. This scheme, named Partitioned Iterative Function System (PIFS), was proposed by Fisher[10]. The recent developments and improvements in the field of image security have increased awareness on the importance of digital signal processing for image recognition and image compression. These developments have made it essential to reduce the digital information that needs to be stored and transmitted. The reduction in both the storage space capacity of the image and its transmission bandwidth is exploited using image compression. The main advantage of image compression is that the percentage of irrelevance and redundancy is reduced. This also optimizes the storage space and enhances the transmission rate. Image compression enables image reconstruction [11, 12]. The digital information contained by the image determines the degree of compression achieved. Singular Value Decomposition (SVD) is one of the most effective tools for image compression and also for biometric recognition such as face recognition.

Deep neural networks can be effectively applied to the problem of lossy image compression [13, 14, 15, 16]. Those methods extend the basic auto-encoder structure and generate a binary representation for an image by quantizing either the bottleneck layer or the corresponding latent variables. Several options have been explored for encoding images at different bit rates including training multiple models [17], learning quantization-scaling parameters [18], and transmitting a subset of the encoded representation within a recurrent structure [19, 20].Deep convolution neural network (CNN) which makes the neural network resurge in recent years and has achieved great success in both artificial intelligent and signal processing fields, also provides a novel and promising solution for image and video compression. In this paper, we provide a systematic, comprehensive and up-to-date review of neural network based image and video compression techniques. The evolution and development of neural network based compression methodologies are introduced for images and video respectively. More specifically, the cutting-edge video coding techniques by leveraging deep learning and HEVC framework are presented and discussed, which promote the state-of-the-art video coding performance substantially. Moreover, the end-to-end image and video coding frameworks based on neural networks are also reviewed, revealing interesting explorations on next generation image and video coding frameworks/standards. The most significant research works on the image and video coding related topics using neural networks are highlighted, and future trends are also envisioned. [21, 22]. Standard lossy image compression techniques such as JPEG and WebP are not data-specific, i.e. they are not designed specifically to handle individual datasets, in which the images are semantically related to each other. Hence these techniques do not make use of the semantic relations among the images in a specific dataset, and do not achieve the best possible compression rates. This has led to a growth in the research towards deep neural network based compression architectures. These models tend to achieve orders-of-magnitude better compression rates while still maintaining higher accuracy and fidelity in their reconstructions. [23, 24,] Santurkar et al. [25] combine Generative Adversarial Networks (GANs) with Variational Auto-encoders (VAEs) in their compression scheme, where the decoder of the compressor is the generator part of the trained GAN, which is later combined with the encoder of the VAE. By constructing the compressor with an encoder obtained from a VAE and a decoder from a GAN, authors aim to make use of the strengths of both models. Specifically, GANs are known to produce high quality, perceptual outputs that are different from the training data, whereas VAEs output images that have high fidelity to their originals, though their reconstructions are not necessarily as visually appealing to humans due to the pixel-wise loss they use for training. In a compression scheme, we need our reconstructions to be both high-quality perceptual images and to be true to their originals. This is why the literature tries to combine these two models. Another end-to-end Convolutional Neural Network for image compression is proposed by Jiang et al. [26]. Additionally, Theiset al. [27] introduce Compressive Auto-encoders, where they deal with the problem of the non-differentiable nature of the quantization component of compression. A quite recent paper by Agustsson et al.[28] show the power of incorporating GANs into these compression schemes. Auto-encoders have been used to reduce the dimensionality of images [Hinton and Salakhutdinov, [29], convert images to compressed binary codes for retrieval [Krizhevsky and Hinton, [30], and to extract compact visual representations that can be used in other applications [Vincent et al., [31]. These are the previous research that have following research gaps:

- Optimized in Time Complexity
- Reduction in Quality
- Reduction in Accuracy
- Lack in HD image Analysis

In this paper, basically we try to reduce the aforesaid issues by the use of our proposed approach which is based on comparatively advanced concepts.

PROPOSED METHODOLOGY:

In this Section we present our novel approach which is basically combination of approximate DCT & GAN based image compression. As we can see in figure (3.1), the very first step is to insert image and after that we apply the approximation logic and based on that we apply image resizing and convert a big image in to a small image due to this step we are able to save lots of time. Now after that next step is apply approximate DCT image compression approach where we use our own DCT coefficient which follows the 2-bit logic. After the image compression we apply GAN based image compression and thus the first step is training the network, for that we use Tensor flow as a machine & deep learning library and by using Python we did the training of images.

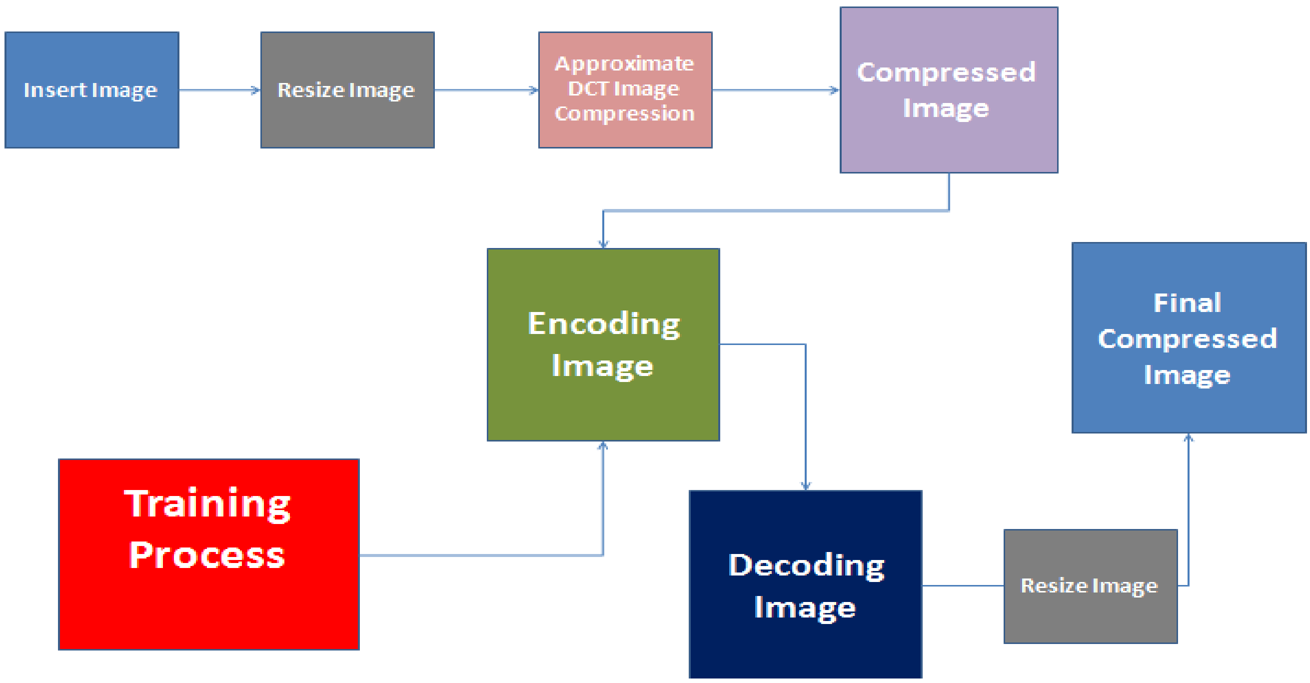

Figure3.1 :Proposed Approach

After the training process we perform the image encoding approach where we use our training data as a reference point and DCT based compressed image as an input image. Now after the encoding process there is encoded file is generated by using of GAN concept. Now the next step is to apply image decoding where we use the encoded file and finally we got the compressed image. Now the last step to reconstruct the image in to original size so for that again we use the image resize. Here we use the following libraries to perform the steps mentioned above:

1. Tensor Flow
2. Open Cv
3. Matlplot
4. Numpy

RESULT & ANALYSIS:

In this section we introduce the relative investigation of all with past existing methodology. In this section we perform the image quality analysis based on parameters below:

1. PSNR
2. SSIM [32]
3. FSIM [33]
4. GMSD 34
5. RFSIM [35]

Here we take **barbara** image as a test image for comparative analysis we use the followings approaches:

1. DCT based Image Compression
2. Fractal Based Image Compression
3. Median Cut Based Image Compression
4. SVD Based Image Compression.

Test Image:

Figure 4.1 : Test Image

- Output From DCT Based Approach:

Figure 4.2 : DCT Compressed Image

- Output From Fractal Based Approach:

Figure 4.3 : Fractal Compressed Image

- Output From Median Cut Based Approach:

Figure 4.4 : Median Cut Compressed Image

- Output From SVD Based Approach:

Figure 4.5 : SVD Compressed Image

- Output From Proposed Based Approach:

Figure 4.6 : Proposed Compressed Image

As we can see in all images our proposed approach have the best quality. Now in terms of scientific result we also did the comparative image quality analysis which we present in table 4.1.

Table 4.1 : " Image Quality Analysis"

S.No.	Approach	PSNR	SSIM	RFSIM	FSIM	GMSD
1	DCT	21	0.89	0.78	0.88	0.72
2	FRACTAL	20.2	0.86	0.75	0.86	0.70
3	MEDIAN CUT	19.3	0.83	0.72	0.83	0.66
4	SVD	18.7	0.83	0.69	0.81	0.62
5	Proposed	24.56	0.93	0.83	0.90	0.80

So as per the analysis we can say that the proposed approach is far better than the existing approaches.

CONCLUSION:

As per current innovation future is completely founded on virtual world. At this moment each diminishes depends on online like shopping, films, pictures, instructions estimated time of arrival. So for these kind of utilization there is need of some other strong system which is able to produce the quality level result. As we are living in the era of 3D and 4G technology , where everyone demand high quality based colour image and videos in their devices like mobile and laptop application. Data science is the most important tool for present system where we can use ML & DL logic and based on that we can create a good quality system. In this paper basically we focus on the data science concept and based on that we use ML & DL logic in terms of GAN with DCT and based on that as result we got a novel approach which is able to generate a quality level result as compare to previous existing approaches.

REFERENCES:

1. A. Radford, L. Metz, and S. Chintala 'Unsupervised representation learning with deep convolutional generative adversarial networks', arXiv preprint arXiv:1511.06434, 2015.
2. Suyog Gupta et al. 'Deep learning with limited numerical precision', In ICML, 2015; 1737–1746.
3. Dan Kalman "A Singularly Valuable Decomposition: the SVD of a matrix", The American University, Washington D.C, February 13, 2002
4. Prasantha H S, Shashidhara H L, Balasubramanya Murthy K N, "Image Compression using SVD", International Conf. on Computational Intellignece and Multimedia Applications, 2007.

5. Gonzalez RG,Woods RE Digital image processing. Publishing House of Electronics Industry 2007.
6. Beijing Gottlieb D, Shu CW (1997) On the gibbs phenomenon and its resolution. SIAM Rev Guo X, Zhang W, Zhong W 2014; 39(4):644–668
7. Doing topology optimization explicitly and geometrically—a new moving morphable components based framework. ASME J Appl Mech 81(8):081009-1– 081009-12
8. Jacquin AE . Image coding based on a fractal theory of iterated contractive im- age transformations. IEEE Trans Image Process 1992;1(1):18–30 .
9. Ramamurthi B , Gersho A . Classified vector quantization of images. IEEE Trans Commun 1986; 34(11):1105–15 .
10. Fisher Y , Jacobs E , Boss R . Fractal image compression using iterated transforms. In: Image and text compression. Springer; 1992; 35–61 .
11. Dan Kalman "A Singularly Valuable Decomposition: the SVD of a matrix", The American University, Washington D.C, February 13, 2002.
12. M. Grossber, I. Gladkova, S. Gottipat , M. Rabinowitz, P. Alabi, T. George1, and A.Pacheco, "A Comparative Study of Lossless Compression Algorithms on Multi-Spectral Imager Data", Data Compression Conference, IEEE, 2009; 447.
13. L. Theis, W. Shi, A. Cunningham, and F. Huszar. Lossy image compression with compressive autoencoders. In Int'l. Conf. on Learning Representations (ICLR2017), 2017; 1(2, 5, 6, 8)
14. G. Toderici, S. M. O'Malley, S. J. Hwang, D. Vincent, D. Minnen, S. Baluja, M. Covell, and R. Sukthankar. Variable rate image compression with recurrent neural networks. ICLR 2016; 1, 2, 3
15. G. Toderici, D. Vincent, N. Johnston, S. J. Hwang, D. Minnen, J. Shor, and M. Covell. Full resolution image compression with recurrent neural networks. CVPR, abs/1608.05148, 2017.
16. K. Gregor, F. Besse, D. Jimenez Rezende, I. Danihelka, and D.Wierstra. Towards conceptual compression. In D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett, editors, Advances in Neural Information Processing Systems, Curran Associates, Inc., 2016; 29(1, 2): 3549–3557
17. G. K. Wallace. The jpeg still picture compression standard. Communications of the ACM, 1991; 1: 30–44

18. Z. Wang, A. Bovik, A. Conrad, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Process- ing, , 2004; 13(4):600–612, 1, 4, 5
19. Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale structural similarity for image quality assessment. In Sig- nals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on,. IEEE, 2003; 2(1, 5, 8): 1398–1402
20. H. Zhao, O. Gallo, I. Frosio, and J. Kautz. Loss functions for image restoration with neural networks. In IEEE Tran- scations on Computational Imaging, March 2017; 1(3).
21. P. List, A. Joch, J. Lainema, G. Bjontegaard, and M. Karczewicz, “Adaptive deblocking filter,” IEEE Trans. on circuits and systems for video technology, 2003;13(7): 614–619,.
22. A. Norkin, G. Bjontegaard, A. Fuldseth, M. Narroschke, M. Ikeda, K. Andersson, M. Zhou, and G. Van der Auwera, “HEVC deblocking filter,” IEEE Trans. on Circuits and Systems for Video Technology, 2012; 22(12): 1746–1754,.
23. Eirikur Agustsson, Michael Tschannen, Fabian Mentzer, Radu Timofte, and Luc Van Gool. Generative adversarial networks for extreme learned image compression. CoRR, abs/1804.02958, 2018. URL http://arxiv.org/abs/1804.02958.
24. Feng Jiang, Wen Tao, Shaohui Liu, Jie Ren, Xun Guo, and Debin Zhao. An end-to-end compression framework based on convolutional neural networks. CoRR, abs/1708.00838, 2017. URL http: //arxiv.org/abs/1708.00838.
25. Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. CoRR, abs/1710.10196, 2017. URL http://arxiv. org/abs/1710.10196.
26. Zachary C. Lipton and Subarna Tripathi. Precise recovery of latent vectors from generative adversarial networks. CoRR, abs/1702.04782, 2017. URL http://arxiv.org/abs/1702.04782.
27. Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. Deep learning face attributes in the wild. In Proceedings of International Conference on Computer Vision (ICCV), December 2015.
28. Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. CoRR, abs/1511.06434, 2015. URL http:// arxiv.org/abs/1511.06434.
29. Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian

Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL http://tensorflow.org/. Software available from tensorflow.org.

30. Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555, 2014.
31. Djork-Arné Clevert, Thomas Unterthiner, and Sepp Hochreiter. Fast and accurate deep network learning by exponential linear units (elus). CoRR, abs/1511.07289, 2015. URL http://arxiv. org/abs/1511.07289.
32. Hore, Alain, and Djemel Ziou. "Image quality metrics: PSNR vs. SSIM." 2010 20th International Conference on Pattern Recognition. IEEE, 2010.
33. Zhang, Lin, et al. "FSIM: A feature similarity index for image quality assessment." IEEE transactions on Image Processing 2011; 20(8): 2378-2386.
34. Zhang, Lin, Lei Zhang, and Xuanqin Mou. "RFSIM: A feature based image quality assessment metric using Riesz transforms." 2010 IEEE International Conference on Image Processing. IEEE, 2010.
35. Xue, Wufeng, et al. "Gradient magnitude similarity deviation: A highly efficient perceptual image quality index." IEEE Transactions on Image Processing 2013; 23(2): 684-695.