

Handwritten Decimal Image Compression using Deep Stacked Autoencoder

Swati Pachare¹, Shubhada Thakare²

¹P.G. Scholar, Dept. of Electronics Engineering, GCOE, Amravati

²Assistant Professor, Dept. of Electronics Engineering, GCOE, Amravati

Abstract - Compression of image is a technique that is used to identify internal data redundancy and to subsequently come up with a compact representation. Compression of image has been a necessary and effective topic of research in the image processing domain. Earlier used algorithms for image compression, like JPEG and JPEG2000, depend on the encoder/decoder (codec) block diagram. They use the fixed transform matrixes, i.e wavelet transform and Discrete Cosine Transform (DCT), together with quantization and entropy coder to compress the image together with quantization and entropy coder to compress the image. A compression of image is important in the applications image processing like storing of data, classification of image, recognition of image etc. However, the image compression with auto-encoder has been found for a small number of the improvements. Therefore, the proposed work is to study and demonstrate the image compression algorithm using the deep neural network (DNN). Deep learning has much potential to enhance the performance in various computer vision tasks.

Key Words: Image compression, image processing, auto-encoder, deep neural network

1. INTRODUCTION

Image compression is the art and science art diminishing the data required in the representation of an image. It is most helpful and commercially successful techniques in the area of the digital image processing. Compression enables image transmission at very low bandwidth and decreases space needed for the storage of the data. Image compression has turn out to be necessary owing to increased demand for information transfer and storage.

J. Jiang presents a review on image compression with neural network shows image compression existing technology like, MPEG, H.26X and JPEG standards are being developed with assisting with neural network to provide improvement over traditional algorithms [1].

Chao Dong et.al shows lossy compression methods like JPEG introduces ringing effects, blocking artifacts. In wavelet transform thersholding and DCT transform in shape-adaptive are used to remove ringing and blocking artifacts but gives the blurred output. To eliminate the undesired artifacts Artifacts Reduction-Convolutional Neural Network (AR-CNN) is used shows impressive results it suppressed the blocking artifacts while retains the edge patterns and sharp

details [2]. G. E. Hinton, R. R. Salakhutdinov proposed dimensionality reduction method of data using the Restricted Boltzmann Machine (RBM) which outperforms over the Principal Component Analysis (PCA) method [3]. Adna Sento proposed the method of image compression method using auto-encoder algorithm uses extended Kalman filter algorithm as learning algorithm which updates the weights in the network [4].

Hongda Shen [5] proposed the lossless compression method of curated erythrocyte images using stacked auto-encoder and their variants. The dimensionality reduction of the images preserves all discriminative features of the images. This compression gives good compression performance as compared to JP2K-LM, JPEG-LS and CALIC. J. Almotiri et.al [6] presents the comparison between auto-encoder and Principal Component Analysis (PCA) in the compression and classification of the data. The auto-encoder gives 98.1% accuracy while PCA gives 97.2% accuracy. Robert Torfason et al. proposed two distinct computer vision tasks from compressed image representations which are image classification and semantic segmentation are consider. When combining classification and image compression training, observe an increase in MS-SSIM and SSIM and at the same time, a better segmentation and classification precision [7]. A.B. Said et al. [8] presents the combined compression and classification of EMG and EEG signals by means of deep learning approach. Deep architecture extracts the features and also reconstructs the data using greedy layer wise training. This experiment is conduct on DEAP dataset. It consists of EEG, EMG and multiple physiological signals recorded from 32 participants. For both EEG and EMG data they have 23,040 samples of dimensionality 896. While comparison with DWT and compress sensing shows that this approach performs better with high compression ratio. J. Papitha, G. Merlin Nancy, D. Nedumaran proposed eight compression algorithms were compare on 200 and more MR images for the evaluation of quality and also the performance of it [9]. Gaurav Kumar and Pradeep Kumar Bhatia [10] expound the brief comparison between wavelet, Discrete Cosine Transform and Neural network methods used for image compression. The comparison is based on performance parameters such as PSNR, retained energy and output image size. From this comparison wavelet based image compression shows the optimum results as compared with ANN and DCT. W. K. Yeo et al., proposed the method to compress the MRI images using the feed forward neural network and for training back-propagation algorithm is used. The lossless algorithms like JPEG, JPEG 2000

outperforms well in terms of image quality and compression ratio as compared to feed forward neural network (FNN) but the FNN in the field of medical [11].

2. PROPOSED SYSTEM

This section includes the required compression and decompression of the handwritten decimal numerals images from the dataset. For this compression and reconstruction of the images stacked auto-encoder is used. The auto-encoder encodes the image into its compressed representation and from this compressed representation reconstruction of the images is carried out.

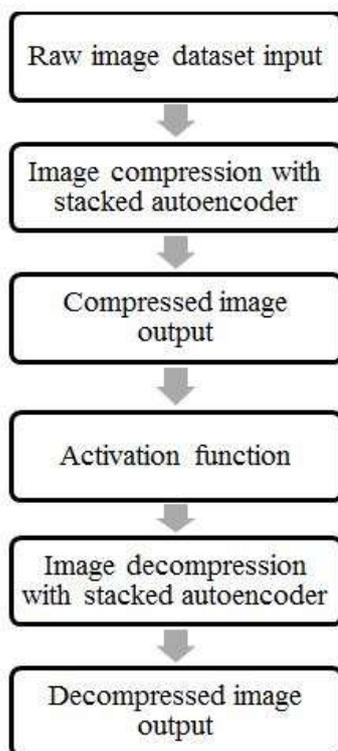


Fig- 1: Block Diagram

1. Modified National Institute of Standard and Technology (MNIST) is the database used consisting of collection of handwritten digit images. It consists of 60,000 training images and 10,000 testing images [12].

2. Image compression with stacked auto-encoder:

Auto-encoders [13] can be stacked to form a deep network by feeding the internal representation (output code) of the Auto-encoder at the layer below as input to the considered layer. The unsupervised pre-training of the architecture is done one layer at a time. Suppose we consider k layers, once the first k layers are trained, it is possible to train the $(k + 1)^{th}$ layer using the internal representation of the k^{th} layer. The advantage of this architecture is that by using more hidden layers than a single auto-encoder, a high-dimensional input data can be reduced to a much smaller code representing the important features.

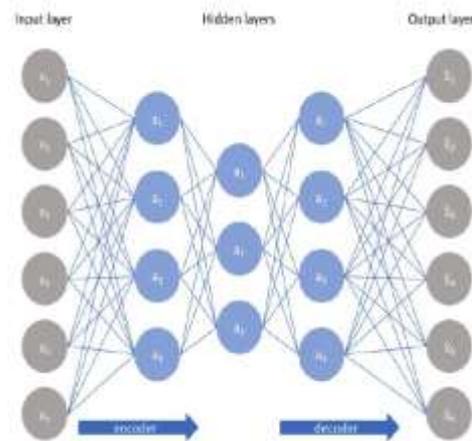


Fig-2: Stacked Autoencoder

Training of Auto-encoder

- Step 1: Start
- Step 2: Create and configure the auto-encoder network
- Step 3: Initialize the weights and biases
- Step 4: Train the auto-encoder network with Scaled Conjugate Gradient (SCG) method
- Step 5: Output of hidden layer of 1st auto-encoder network is given as input to the 2nd auto-encoder
- Step 6: Repeat step 4
- Step 7: Again the output of hidden layer of 2nd auto-encoder network is given as input to the 3rd auto-encoder.
- Step 8: Repeat step 4
- Step 9: And after training of all three auto-encoder the output of hidden layer 3rd Auto-encoder is given as input to the softmax layer which is the activation function.
- Step 10: All this auto-encoder are stacked together to form deep stacked auto-encoder.
- Step 11: End

Scaled Conjugate Gradient (SCG) [14] training algorithm is used. SCG is the supervised learning algorithm. SCG is fully automated, includes no critical user-dependent parameters, and avoids a time consuming line search, which CGB and BFGS uses in each of its iteration in order to determine an appropriate step size. SCG belongs to class of conjugate gradient methods which shows super-linear convergence on most problems. The speed is depends upon convergence criteria i.e bigger demand for reduction of error, bigger the speed-up.

3. EXPERIMENTATION

1. Data collection

The dataset chosen for the analysis of the deep networks is MNIST which contains a set of handwritten digit samples from 0 to 9. Each digit is in the form of a grey-scale image, of size 20×20 pixels, with values in range $[0;1]$

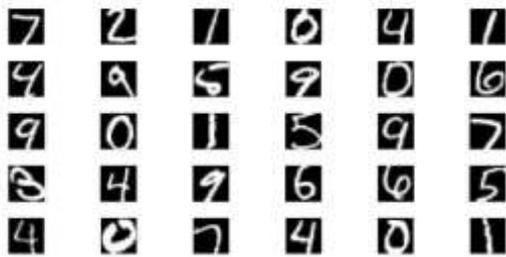


Fig- 3: MNIST Sample Images

2. Training of auto-encoder

The MNIST handwritten digits dataset have been used for this experiment. This architecture is consists of three hidden layers where first layer has 100 units, second layer has 50 units and third layer has 25 units. This structure can be denoted as [400->100->50->25]. The training of these three auto-encoders is performed one by one. Then after training all three auto-encoders are stacked together to form a deep architecture.

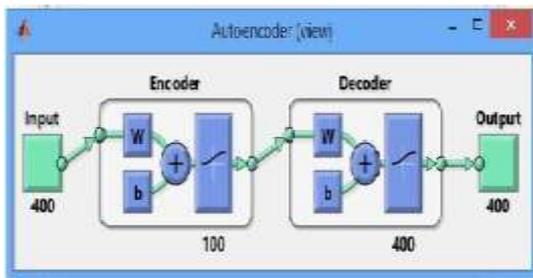


Fig- 4 : Auto-encoder Layer 1

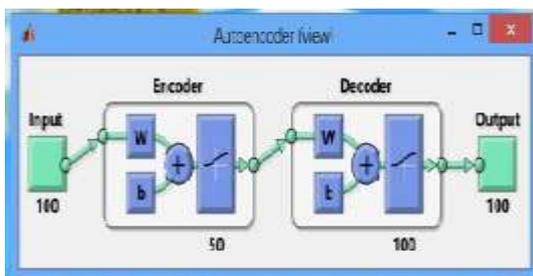


Fig- 5 : Auto-encoder Layer 2

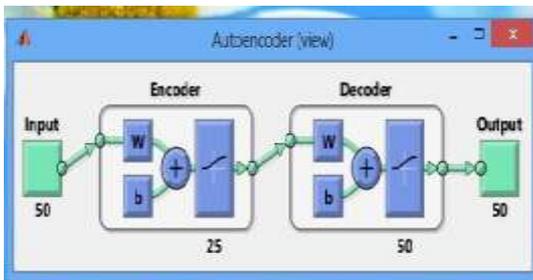


Fig- 6 : Auto-encoder Layer 3

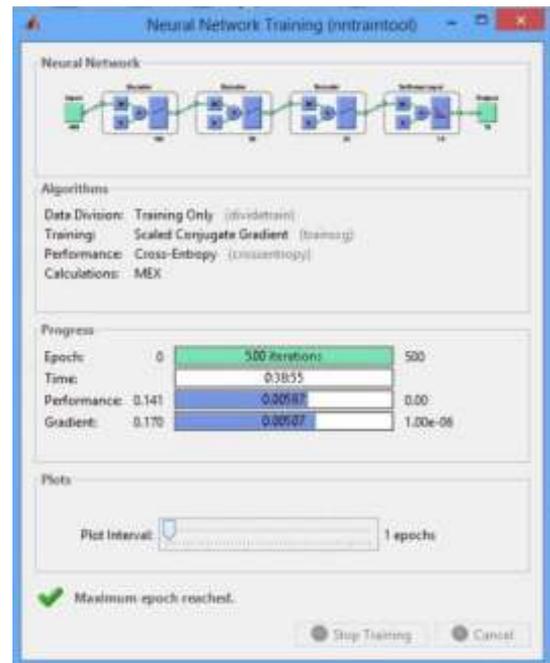


Fig-7: Stacked Auto-encoder

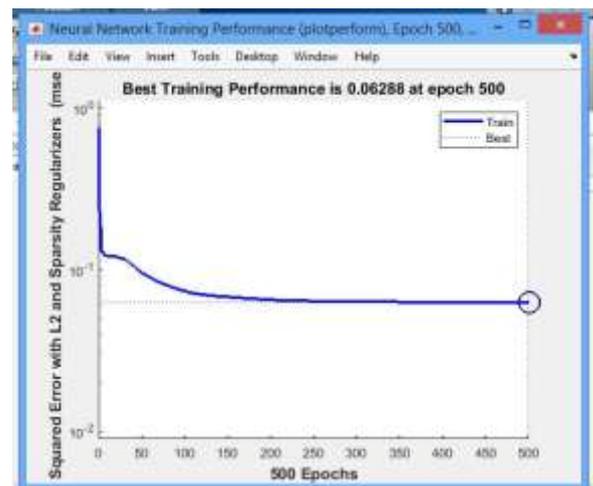


Fig- 8: MSE obtained at layer 1

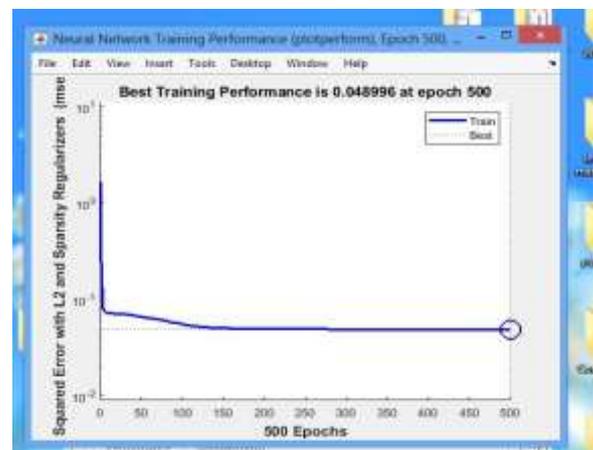


Fig- 9: MSE obtained at layer 2

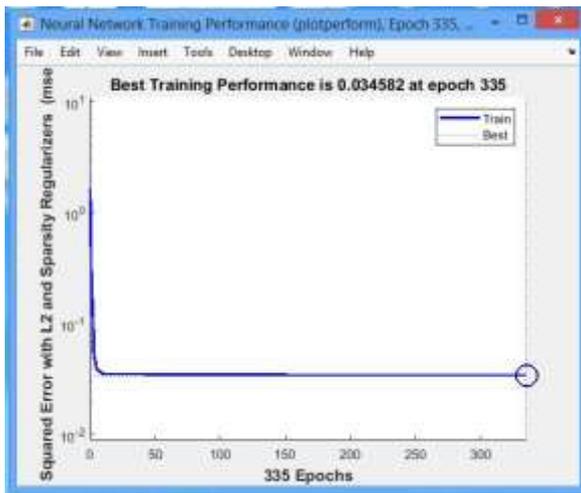


Fig- 10: MSE obtained at layer 3

3. Testing of Auto-encoder

While testing, 10,000 images from the MNIST dataset are taken for the testing. All trained neural networks are saved as a .mat file that can be used for testing. These .mat file can be directly loaded when testing is carried out. So one by one image or array of images are loaded from the dataset and compression at each layer is shown on the output window and after compression decompressed image is obtained.

4. EXPERIMENTAL RESULTS

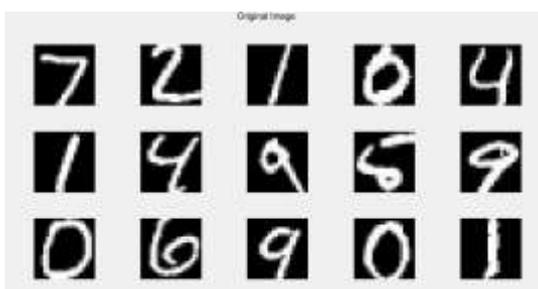


Fig-11: Input Image

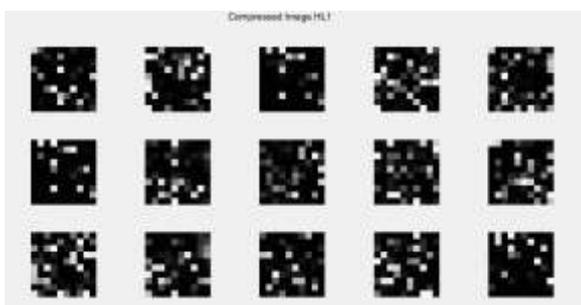


Fig-12: Compressed image at hidden layer 1

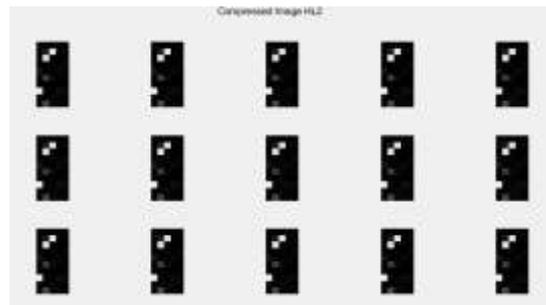


Fig-13: Compressed image at hidden layer 2

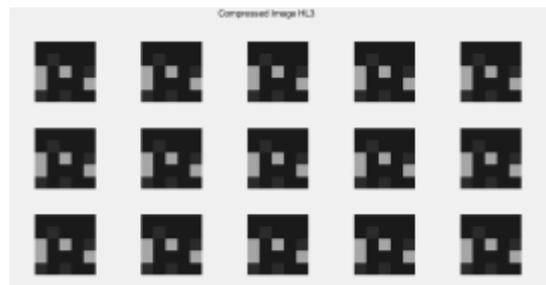


Fig- 14: Compressed image at hidden layer 3

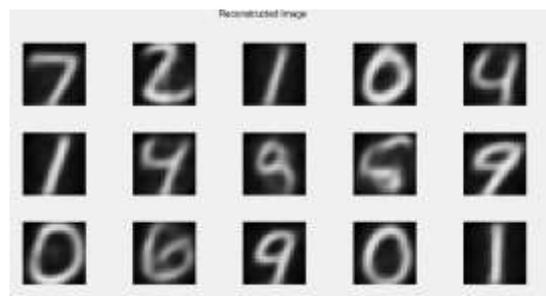


Fig-15: Decompressed image output

5. CONCLUSION

In this paper the review of various compression of image techniques are studied and compared. The main purpose of the image compression to obtain less storage and low bandwidth which is satisfied by the deep learning as compared to machine learning and conventional image compression technique by comparing the results of above mentioned papers. In the deep learning, it learns features and classifies automatically which is the advantage deep learning. Due to large dataset the training required more time and this is the main disadvantage of deep learning. But once training is done testing can be done in less time. The dimensionality reduction of the input image can be done very precisely from the very large dataset of images. And compression at each layer gives the more compressed representation which is required and also as we move on increasing the layers can get more compressed image. The decompression of image can also be done very easily because there is no need to apply separate decompression method.

REFERENCES

- [1] J. Jiang, "Image compression with neural networks- A survey", *Signal Processing: Image Communication* 14(1999) 737-760
- [2] C. Dong, Y. Deng, C. C. Loy and X. Tang, "Compression Artifacts Reduction by a Deep Convolutional Network," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 2015, pp. 576-584.
- [3] G.E. Hinton and R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science*, vol. 313, no. 5786, pp. 504-507, 2006
- [4] Adna Sento "Image Compression with Auto-encoder Algorithm using Deep Neural Network (DNN)" 2016 IEEE Management and Innovation Technology International Conference (MITiCON-2016), pp. 99-103
- [5] Hongda Shen, W. David Pan, Yuhang Dong, and Mohammad Alim, "Lossless Compression of Curated Erythrocyte Images Using Deep Autoencoders for Malaria Infection Diagnosis", 2016 IEEE
- [6] Jasem Almotiri, Khaled Elleithy, Abdelrahman Elleithy, "Comparison of Autoencoder and Principal Component Analysis Followed by Neural Network for E-Learning Using Handwritten Recognition", 2017 IEEE
- [7] Robert Torfason, Fabian Mentzer, Eirikur Agustsson, Michael Tschannen, Radu Timofte, Luc Van Gool "Towards image understanding from deep compression without decoding" Published as a conference paper at ICLR 2018
- [8] Ahmed Ben Said, Amr Mohamed, Tarek Elfouly, Khaled Harras, Z. Jane Wang, "Multimodal deep learning approach for joint EEG-EMG data compression and classification", 2017 IEEE
- [9] J. Papitha, G. Merlin Nancy and D. Nedumaran, "Compression Techniques on MR Image-A Comparative Study", 2013 IEEE
- [10] Gaurav Kumar, Pradeep Kumar Bhatia, "Empirical analysis of Image Compression using Wavelets, Discrete Cosine Transform and Neural Network", 2016 International Conference on Computing for Sustainable Global Development (INDIACom), pp. 3862-3866
- [11] W. K. Yeo, David F. W. Yap, T.H. Oh, D.P. Andito, S. L. Kok, Y. H. Ho, M. K. Suaidi, "Grayscale medical image compression using feedforwrd neural network", 2011 International Conference on Computer Applications and Industrial Electronics (ICCAIE 2011), pp.,633-638
- [12] <http://yann.lecun.com/exdb/mnist>
- [13] Josh Patterson and Adam Gibson, "Deep Learning A Practitioner's Approach", Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA 95472
- [14] Martin Fodslette Meiller, "A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning", *Neural Networks*, Vol. 6, pp. 525-533, 1993
- [15] Salman Khan, Hossein Rahmani, Syed Afaq Ali Shah, Mohammed Bennamoun, "A Guide to Convolutional Neural Networks for Computer Vision", A Publication in the Morgan and Claypool Publishers series