International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-4, February 2020

An Image Compression Based Technique to Watermark a Neural Network

R. S. Kavitha, U. Eranna, M. N. Giriprasad

Abstract: ***While neural networks have made considerable progress in the area of digital representation, training of neural models requires an enormous data and time. It is well known that the use of trained models as initial weights often leads in less training error than un-pre-trained neural networks. We propose in this paper a digital watermarking system for neural networks. We formulate a new challenge: the integration of watermarks into neural networks through discrete cosine transform (DCT) based approach. For discrete wavelet transform (DWT)-based digital image watermarking algorithms, additional performance enhancements could be obtained by combining DWT with DCT. Throughout the neural networks, we also describe specifications, embedded conditions, and attack forms of watermarking. The technique presented here does not affect the network performance in which a watermark is positioned as the watermark is embedded while the host network is being trained. Finally, we perform detailed image data experiments to demonstrate the potential of neural networks watermarking as the basis for this research attempt.***

Keywords: Digital Watermarking, Neural Network, Vision and Image Processing.

I. INTRODUCTION

The concept of digital watermarking is to establish ownership of digital content, such as photographs, audio and videos [1]. Digital watermarking can be done on text, image, audio, video and graphics in spatial or frequency domain. The watermark can be of noise type (pseudo noise, Gaussian random and chaotic sequences) or image type (binary image, stamp, logo and label). Based on the deployment conditions various watermarking techniques can be used. For public use, visible watermarks are preferred while for private applications and to arrest unauthorized copying invisible watermarking can be used. Fragile watermarks are used in tamper-proof applications whereas robust watermarking is used in applications where the watermark should remain intact even after modification or tampered with.

According to the detection stage, visual watermarking is more robust which needs the original media and the embedded watermark for detection while blind watermarking does not require any of these.

Revised Manuscript Received on February 28, 2020.
***** Correspondence Author
R. S. Kavitha*, Department of ECE, GATES Institute of Technology, Gooty, AP, India. E-mail: drkavithakavana@gmail.com
U. Eranna, Department of ECE, BITM, Ballari, Karnataka. India. E-mail: jayaveer_88@yahoo.com
M. N. Giriprasad, Department of ECE, JNTUA, Anantapur, AP. India. E-mail: mahendragiri1960@gmail.com

This is the most demanding type of watermarking. Watermarking is generated and embedded at the transmitter, while detection and extraction will happen at receiver. Watermarking is also done even in prepossessing stage.

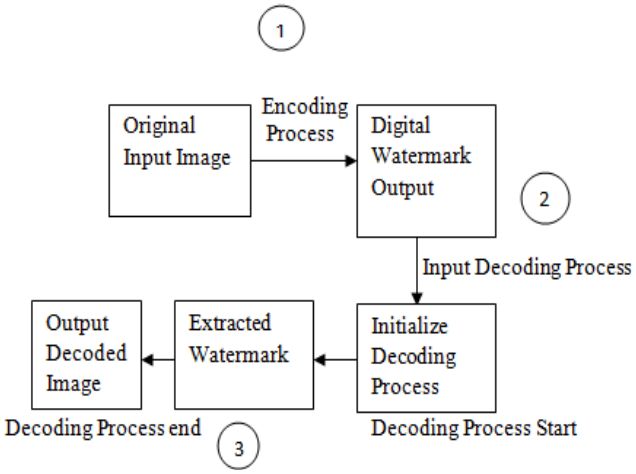


Figure 1. A Generalized Block Diagram

Deep neural networks in the area of multimedia analysis have made huge progress. It aims to model high-level data abstractions by using deep architectures composed of multiple dynamical transforms. In addition, there have been published several deep learning frameworks [2]. They allow engineers and researchers build deep learning-based systems or do less effort-based research. Accessing trained models is very important for the rapid advancement of deep neural network systems research and development. For the owner(s) who train the models, the trained models could be important assets. Quality and quantity of data setting directly affect the performance of large network activities. Deep neural network success was achieved not only through algorithms, but also by massive amounts of data and computational power [3]. The trained models can be viewed as intellectual property, and it is a worthy challenge to provide copyright protection for trained models. We emphasis on how the copyrights of trained models can be protected computationally and propose for neural networks a digital watermarking technology [4].

We propose a conceptual framework for integrating a watermark into models of deep neural networks to safeguard copyrights and identify violation of trained models of intellectual property. The Figure 1 illustrates the generalized block diagram of this conceptual framework with all the three steps are presented in detail in the next section of the methodology adopted. The theory of applying two transforms that is DWT and DCT is based on the fact that combined transforms can compensate for each other's disadvantages, leading to efficient watermarking.

II. METHODOLOGY ADOPTED

This section describes the detailed methodology adopted in applying the digital watermark to the neural network while elucidating the specifications, embedded conditions, and attack forms of watermarking [5]. The cover image (color image) size is given by mc x nc, the mc being the no of rows and nc being number of columns.

The watermark image is of size mw x nw. Next the mid band matrix is selected appropriately. The Figure 2 illustrates the flow chart for watermark embedding process with the encoding methodology. During the Encoding Process the midband coefficients selection matrix is given as:

midband= [mid11, mid12.....midij,..,mid88]

for 1<=i<=8 and

1<=j<=8 and

midij = 0 or 1.

Here the blocksize is selected as 8.

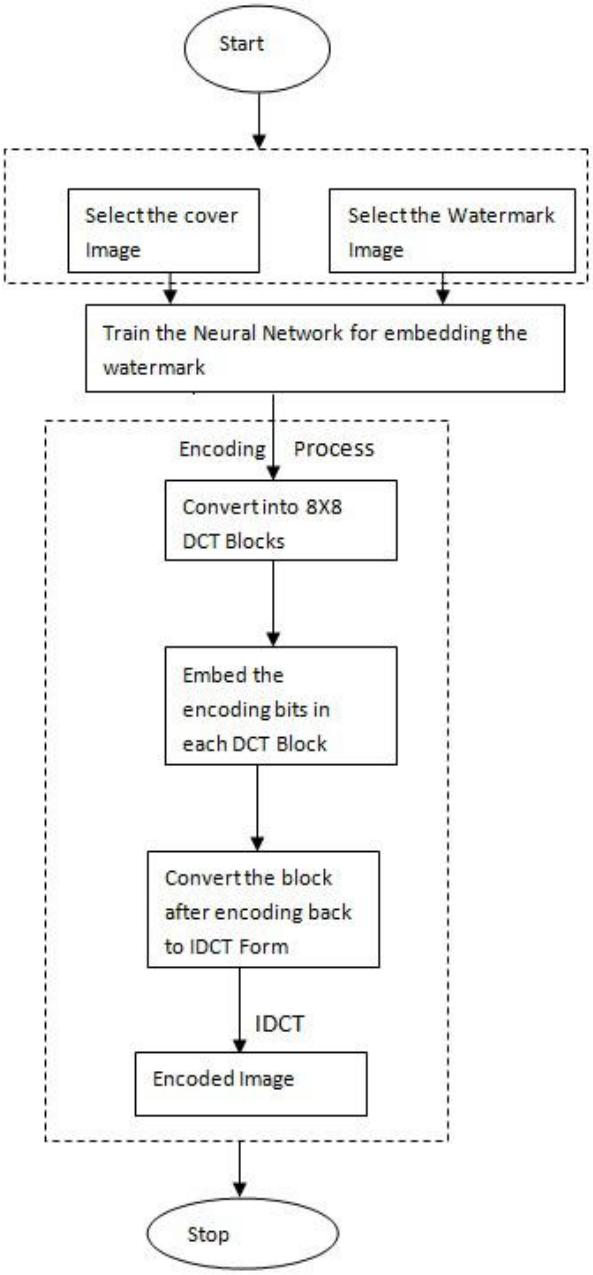


Figure 2. Flow Diagram for embeddeding the watermark

Next we define the Sum of Midband as ΣΣmidband(i,j) and the message to be inserted is message = (mc x nc)/blocksize2. Here the mc represents the number of rows and nc represents the number of columns in the encoding binary message matrix. This message is generated by using random integer function of MATLAB. The variables p=mc/8, q= nc/8 and R = p*q where R is total no of blocks of size 8x8. This will generate the random numbers equal to number of mid-band coefficients. The cover image and watermark are then supplied to the input layer of neural network. The mid band matrix is useful in selecting the discrete cosine transform (DCT) block for encoding the message, with the condition is if the midband (i,j)=1 then only encode that respective DCT block. As the neural network is trained the cover image is divided into blocks of 8x8 one by one. With the neural network training process each block is then transformed into its equivalent DCT coefficient block [6].

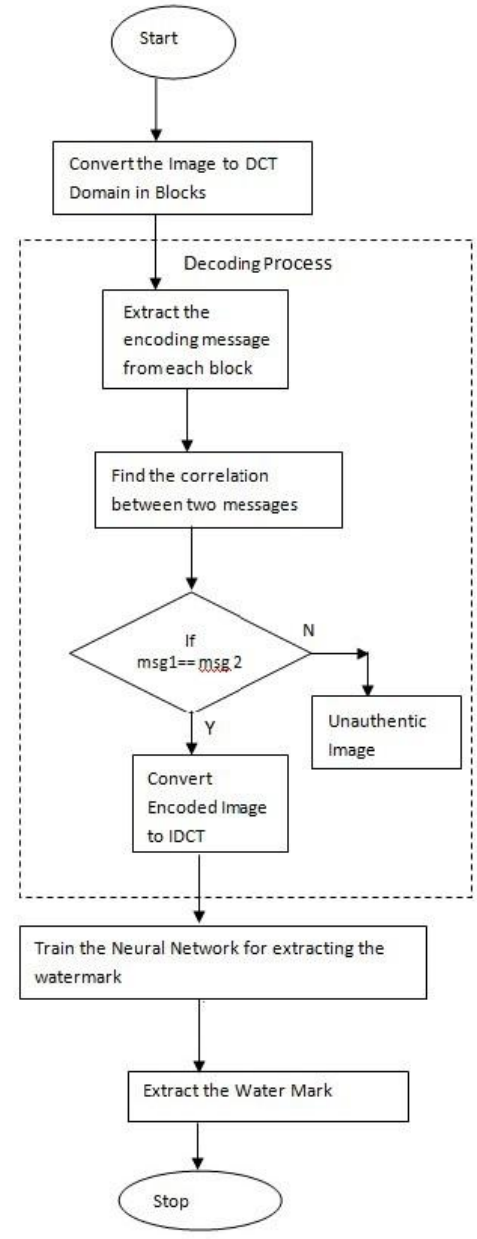


Figure 3. Flow Diagram for extracting the watermark

Then the DCT block is further converted into inverse DCT block with subsequent iterations. If total number of columns are exceeded then with reinitialization process the next row is considered.

This process is explained mathematically as:

if(x+8)<nc then;

x=x + (blocksize),

x=1 and y = y + blocksize for (x+8)>nc.

Next the total number of blocks in the host image is incremented for encoding the next block as (k =k+1 until k=p*q).

The encoded image is converted to column vector in the following form: Cover Image= [X1, X2, X3,. . . . . .., Xn1] Where n1=mc x nc is the total number of pixels in the cover image. Also, the watermark W= [Y1, Y2, Y3,. . . . . ,Yn2] where (n2=mw x nw) is converted to the column vector.

The encoded process ends at this point and decoding process starts with the DCT block of watermarked image is obtained blockwise in the first step.

The initial index of the dct block is set to 1. The embedded sequence is obtained as sequence of positions (pos)= dct block(i,j), for 1<i<blocksize, 1<j<blocksize, midband (i,j)=1 where, for all new pair (i,j) the position is pos=pos+1.

The correlation of this obtained sequence is initialized with a zero sequence. Now as x is incremented and if x exceeds the total number of columns , then x is reinitialized, and next row is taken.

This process is explained mathematically as follows:

x=x+ blocksize for (x+8) < nw and

x=1 and y=y + blocksize for (x+8) > nw.

The message(msg) from the correlation of the two sequences is found as follows: (msg)=0 for correlate(msg)>0.5 and 1 otherwise. For every message (msg): 1<=msg<=mm x nm and further a Boolean flag variable is checked and if flag = 1 then the decoding process is successful, if ag = 0 then image is not authentic, and it is not supplied to counter propagation network [7] for extracting the watermark and the decoding process ends at this point.

Next we discuss the detailed neural network training process.

III. NEURAL NETWORK TRAINING PROCESS

In this section the neural network training process is described with initial step as importing the image data to hold the original image and the corresponding image with the watermark [8]. Here the we standardize and reshape the images to fit into the network. The Number of neurons in the hidden layer is considered to be 10 and the learning rate for input layer α=0.4 and the learning rate for the output layer β=0.3 [9].

Next we split the data into the training and validation sets and create the response flag variables for the image data. The images with the watermark have a response 1(Flag=1), while the images without a watermark have a response 0(Flag=0).

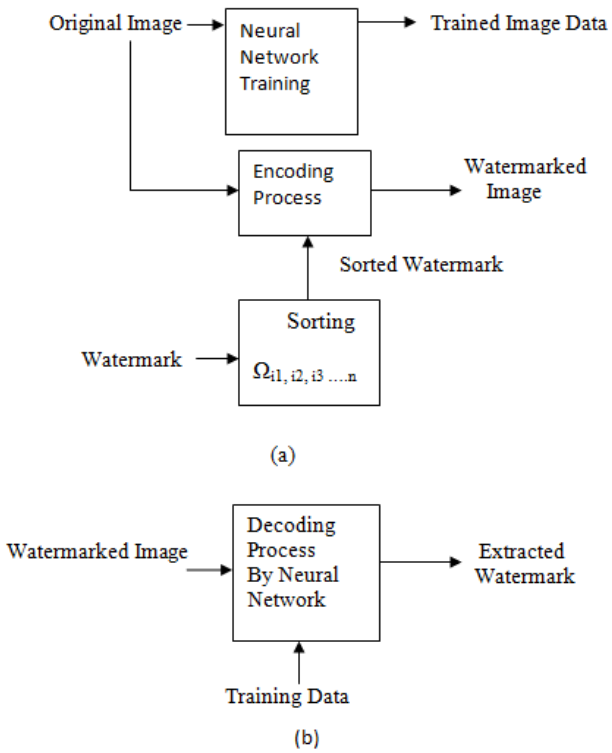


Figure 4. Neural Network Training Steps for (a) Watermark Embedding (b) Watermark Extracting

The watermark embedding steps is as illuminated in the Figure 4(a) and the water extraction sequence is as depicted in Figure 4(b).

The Neural Net was reused for recognized unmodified data in the extraction segment. The watermark bit data is ' 1 ' if a data is recognized by the network, and ' 0 ' if a data is not recognized by the network [10].

IV. RESULTS

In this section we train the neural network and apply the watermark and the corresponding results are illustrated. Displaying training progress at 5000 iterations with a total error of 0.0017274 and training starts with initializing the image normalization.

As illustrated in Figure 5(a) displays the neural network training progress with graph of % accuracy versus iteration. The Figure 5(b) displays the Loss versus iteration characteristics. The utilized hardware resource is the single CPU with a constant learning rate schedule with 0.01 learning rate. After reaching the final iteration the training completes with a validation accuracy of 99.68%. The training cycle is divided into 4:4 epochs with number of iterations as 232:232 with 58 iterations per epoch and the maximum number of iterations is 232. The validation cycle frequency is observed to be 30 iterations. The Figure 6 illustrates the experiment on watermark positioning and embedding on the standard experimental image considered as 'Lena'. The image

'Panda' is selected to be the image to be embedded and is considered as watermark image. The experimental result has proved that there can be a better performance on several different standard images and more robustness against various attacks.

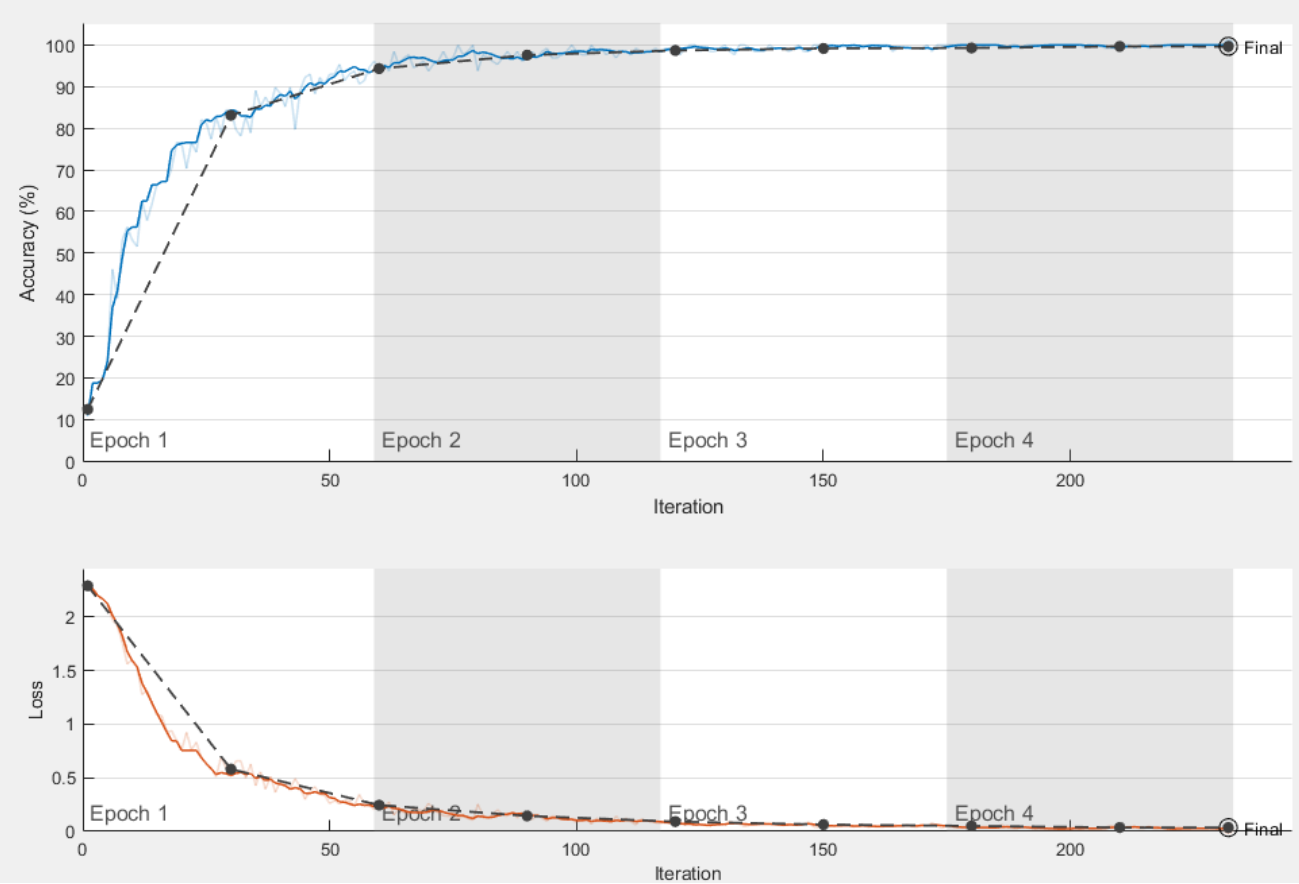


Figure 5. Neural Network Training process (a) % Accuracy versus iterations plot (b) Loss versus iterations plot

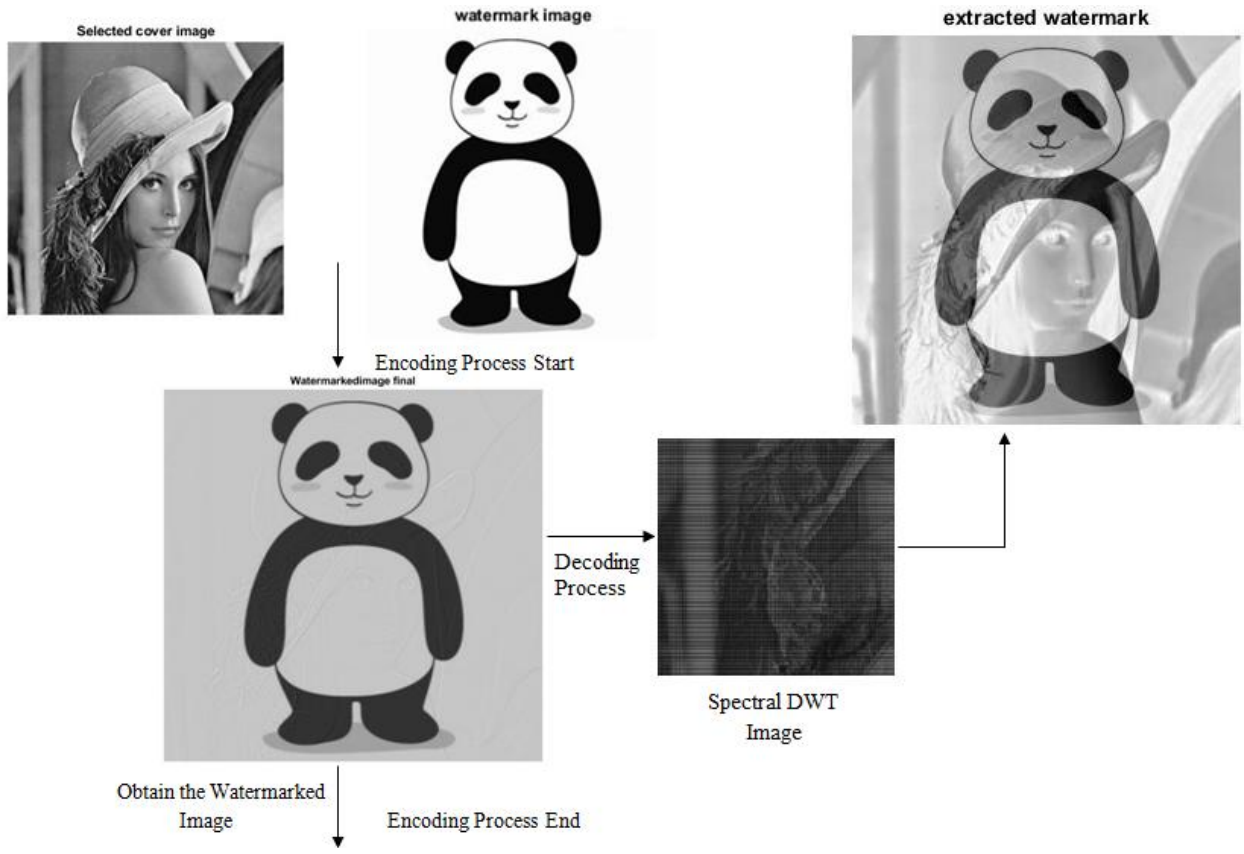


Figure 6. Experiment on Watermark positioning and embedding on image data

Retrieval Number: D9064019420/2020©BEIESP
DOI: 10.35940/ijitee.D9064.029420
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

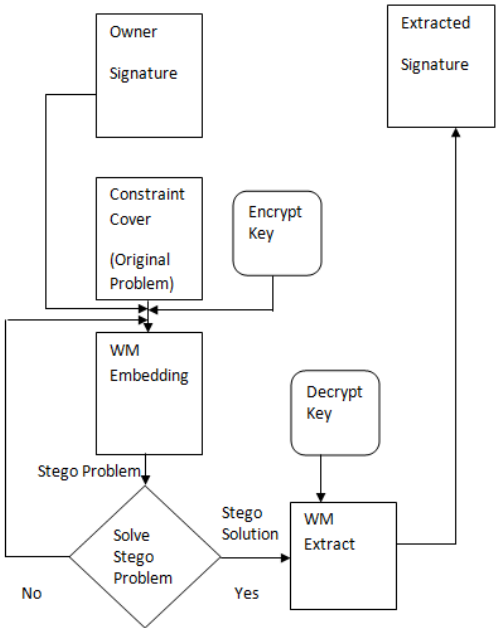
V. CONCLUSION

Before we presented a new combined block-based DCT and Neural Network watermarking scheme. DCT block-based transform is used to boost imperceptibility, and Neural Network recognition power is used for extraction of performance efficiency. Results of the experiment show our proposed scheme's better performance and robustness.

While frequency domain-based digital watermarking strategies tend to be good in perceptibility but poor against geometric attacks, the current approaches may be more likely to combat geometric attacks. Utilizing Neural Network is one of the popular approaches digital watermarking. In future the research would continue in the direction towards digital watermarking based on intelligence.

APPENDIX

The Figure below illustrates the typical constraint based watermarking system with a conceptual framework with all the mathematical modelling steps describing the stego key with the set of various constraints. The theory of applying two transforms is based on the fact that combined transforms can compensate for each other's disadvantages, leading to efficient watermarking.



ACKNOWLEDGMENT

The authors would like to express their gratitude to JNTU, Hyderabad for providing the necessary infrastructure to carry-out this work.

REFERENCES

1. Huang, J et.al.,\Embedding Image Watermarks in DC Components", IEEE Transactions on Circuits and System for Video Technology, vol. 10, no. 6, pp.974-979, 2000.
2. Y. H. Zhang, \A Blind Digital Watermarking Algorithm Based on HVS and RBF Neural Network," Proceeding of the 3rd WSEAS International Conference on Computer Engineering and Applications, vol. 09, no. 2, pp. 202-205, February 2002.
3. Han S, Liu X, Mao H, Pu J, Pedram A, Horowitz MA, Dally WJ "efficient inference engine on compressed deep neural network.", In: Proceedings of ISCA 2016.
4. Abadi M et al "Tensor Low: Large-scale machine learning on heterogeneous distributed systems." arXiv:1603.04467, 2016.
5. T. K. Tewari and V. Saxena, \An Improved and Robust DCT based Digital Image Watermarking Scheme," International Journal of Computer Application, vol. 3, pp. 272-277, June 2010.
6. D. T. Meva and A. D. Kothari , \Adoption of Neural Network Approach in Steganography and Digital Watermarking for Convert Communication and Copyright Protection," International Journal of Information Technology and Knowledge Management, Vol.4, no. 2, pp. 527-529, July 2011.
7. N. Bansal and P. Pathak, \A Review of Applications of Neural Network in Digital Watermarking," Proceedings published in International Journal of Computer Applications (IJCA), Vol.2, no. 1, pp. 127-132, December,2011.
8. S. Oueslati, A. Cheris and B. Solaimane, \Adaptive Image Watermarking Scheme Based on Neural Network International Journal of Engineering Science and Technology (IJEST), vol. 3, no. 1, pp. 748-756, Jan. 2011.
9. J. Kung, D. Kim, and S. Mukhopadhyay, "On the impact of energy accuracy tradeoff in a digital cellular neural network for image processing," IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 34, pp. 1070-1081, July 2015.
10. Cox I, Miller M, Bloom J, Fridrich J, Kalker T, "Digital watermarking and steganography". Morgan Kaufmann Publishers, 2008. Inc., 2nd edn.

AUTHORS PROFILE

R. S. Kavitha, is a research scholar at JNTUA Ananthapuramu. She is currently associated with GATES Institute of Technology, Gooty, Andhra Pradesh. She received her M. Tech from BITM, Ballari, affiliated to VTU, Belgaum and B Tech from RYMEC, Ballari. Affiliated to Gulbarga University, Gulbarga. She is Life Member of MIE. She has published different novel works and her research interest span over few areas like Digital Image Processing, Digital Signal Processing, and Deep Learning.

Dr. U. Eranna, is leading the Department of ECE at BITM, Ballari as Head of the Department. He received his PhD from Sri Krishnadevaraya University, Ananthapuramu. He received his M.E. from MS university Baroda and B. Tech from JNTUA, Ananthapuramu. He is life member of MISTE. He has rendered his service as Principal BITM, Ballari. He has published different novel works in IEEE conferences and Scopus indexed journal. His research interests span over the fields of Communication Engineering and Image Processing.

Dr. M. N. Giriprasad, has obtained his B.Tech degree from JNT University College of Engineering, Anantapur in 1982, M.Tech degree from SV University College of Engineering, Tirupati in 1994, Ph.D. from JNT University, Hyderabad in 2003. He has 104 research publications in both National and International Journals. Out of these publications, around 10 are SCI and around 55 are SCUPUS. A few areas of his research are Bio Medical Instrumentation, Image and Signal Processing, Medical Image Processing, Digital System Design and Embedded Systems etc., He is life member in Professional societies, viz., ISTE, IEI and NAFEN.