Effect of levels in Dual Tree Complex Wavelet Transform when design Universal image stego-analytic

Hana Mohsen[1], Hala Hassan*[2]

[1]Department of Computer Science, University of Technology, Baghdad, Iraq
[2]Computer Center, University of Baghdad, Baghdad, Iraq

Abstract

   Universal image stego-analytic has become an important issue due to the natural images features curse of dimensionality. Deep neural networks, especially deep convolution networks, have been widely used for the problem of universal image stego-analytic design. This paper describes the effect of selecting suitable value for number of levels during image pre-processing with Dual Tree Complex Wavelet Transform. This value may significantly affect the detection accuracy which is obtained to evaluate the performance of the proposed system. The proposed system is evaluated using three content-adaptive methods, named Highly Undetetable steGO (HUGO), Wavelet Obtained Weights (WOW) and UNIversal WAvelet Relative Distortion (UNIWARD). The obtained precision 0.98387, 0.96659 and 0.98387 for the three content-adaptive methods, applied on BOSS image dataset, respectively. The obtained results show that number of level equals to 5 outperforms other numbers in terms of detection accuracy. Also it minimizes the ime required for both training and testing phases.

Keywords: Deep Learning, Dual-Tree Complex Wavelet Transform, stego-analytic, Convolution Neural Network, Detection.

<div dir="rtl">

تأثير المستويات في تحويل شجرة المويجات المزدوجة عند تصميم محلل الصورة الشاملة

هناء محسن[1]،هالة حسن*[2]

[1]قسم علوم الحاسوب، الجامعه التكنولوجيه، بغداد، العراق

[2]مركز الحاسبه، جامعة بغدادن بغداد، العراق

الخلاصه

يعتبر تصميم محللات الصورة العالمية المحتوية على معلومات ضمنية من القضايا المهمة نسبة الى الابعاد الكبيرة لخصائص الصور الطبيعية. ان الشبكات العصبية العميقة ، وخاصة شبكات الالتواء العميقة من الطرق التي لديها اهتمام بمشكلة تصميم محلل الصورة العالمية. تصف هذه الدراسة تأثير تحديد قيمة مناسبة لعدد من المستويات عند إجراء المعالجة المسبقة على الصورة باستخدام مجمع المويجات المتحول المزدوج الشجرة. ان لهذه القيمة تأثير في دقة الكشف عن الصور التي تم الحصول عليها لتقييم النظام. توضح النتائج 0.98387 ،

</div>

*Email: halah@uob.edu.iq

0.96659 و 0.98387 للطرق القابلة للتكيف مع المحتويات هوكو، واو ويونيوورد على الترتيب. أن عدد

المستويات عندما تكون تساوي خمسة تعطي أداء أفضل للنظام مقارنة بعدد المستويات الأخرى بالاضافة الى تقليل

الوقت للازم للتدريب واختبار الصور.

1. Introduction

The use of digital media has been has been rapidly increased by internet users to communicate, share and disseminate information. Two main crucial issues to the simplicity of malicious processing: first is security and second is trustworthiness of it. Image can be considered as an ideal media for steganography.It can hide a large amount of secret messages [1]. Therefore, several methods for data hiding have been developed in the last few years [2].

The process of hiding secret information inside innocent looking cover media named as steganography, in which the secret message existence may not be noticed. Four different media types may be considered in steganography which are: text, image, audio, and video. Steganography is derived from two main Greek words: "Steganos" followed by "Graphy", which mean covered and writing, respectively. Cover file with secret message hidden constitutes the file of "stage" [3]. A Number of requirements must be satisfied when designing embedding algorithms, these are: Un-detectability, Imperceptibility, Security, Capacity and Robustness [4].

On the other side, steganalysis, which is the process of discovering the secret embedded considered as a difficult process. In case of unauthorized transferring between users, any evidence may be taking in place which is sufficient to make damage in cover file. Different architectures have been developed of stego detection named as detector or stego-analytic [5]. It can be defined as a method which takes a media file as input, and outputs either 'steganogram' or 'innocent'". There are two types of steo-analytic, the first one is to detect embedding information and the second is to extract or destroy these information in addition to detection [6].

In steganography, various media types are used to conceal the secret message. According to alterations in the transport media, there are different types of methodologies that show the presence of the embedding process like visual, structural, statistical, and learning. Steganalysis can be classified into two types; Specific/Target and General/Universal. The main consideration when designing specific steganalysis model is the ability to attack only one known embedding technique. It is used in case with sufficient prior knowledge of the steganography method. It gives a more accurate results but it is not capable to detect new steganography algorithms. Type of blind can be considered as universal, which can be categorized into supervised and unsupervised types, is characterized by a low accuracy in comparison with the target type [7].

2. Related Work

Recently, various architectures for universal image steganalysis have been developed by researchers. Bright prospects have been achieved toward deep learning, especially Convolution Neural Network (CNN). A number of attempts have been presented for designing models of universal image steganalysis. These models and their drawbacks are outlined in this section. Three main Content-adaptive spatial domain steganography algorithms are selected on experiments with different datasets: Highly Undetetable steGO (HUGO) [8] Wavelet Obtained Weights (WOW) [9] and UNIversal WAvelet Relative Distortion (UNIWARD) [10]. Content-adaptive steganography constrains its embedding changes to those parts of covers that are difficult to model, such as textured or noisy regions. When combined with advanced coding techniques, adaptive steganographic methods can embed rather large payloads with low statistical detectability at least when measured using feature-based steganalyzers trained on a given cover source [11].

First attempt was suggested to design image stego-analytc by **Tan and Bin [**12**],** their method improved Spatial-domain Rich Model (SRM) [18] (which captures a large number of different types of dependencies among neighboring pixels to give the model the ability to detect a wide spectrum of embedding algorithms) with auto-encoder in architecture of CNN. SRM included pre-training procedure with unsupervised model. It included a number which extracts features that are relevant comparing to only one stage. Results yields through experimental, proved that their method the ability for attacking the

HUGO algorithm in case of embedding rate selected to 0.4 bpp with Break Our Steganographic System": The Ins and Outs of Organizing BOSS (Bossbase) [13] dataset image contain embedding information. For low payload, efficiency of steganalysis was improved. Its performance is still inferior to what is accomplished by the model (SRM). Graphic Processing Units (GPUs) was new hardware revolution to give better results. "Deep Learn" Toolbox was used for implementation.

Later, another model of CNN achieved performance close to SRM was developed by **Qian, Wang and Tan** [14]. Non-linear activation function named Gaussian was used within layers of the net. Their model has achieved a much lower detection error comparing to Subtractive Pixel Adjacency Matrix (SPAM) [15] (method for detection of steganographic methods that embed in the spatial domain by adding a low-amplitude independent stego signal). Detection error was ranging in 2%-5% higher depending on the payload. Personal Computer (PC) with specification Intel Xeon E5-2650 2.0GHz for CPU and Tesla K40 12G for GPU was used to run the net. Convolution library within C++ language has been written for the models with both Bossbase and ImageNet Datasets for training and testing images. Two main drawbacks have been yielded through experiments in two cases, situations: 1) difference between cover and stego was small and 2) low payload during process of embedding.

Qian Wang and Tan framework have been examined by **Pibre, Jerome, Inco and Chaumont** [16] to give better detection. A new shape of CNN was developed which has the capability to minimize error in detection by more than 16% comparing with other spatial UNIWARD (S-UNIWARD)[17] (The practitioner's goal is to design the distortion to obtain a scheme with a high empirical statistical detectability) at 0.4 bpp. 64 and 16 filters are set in the first and second layers, respectively, to give diversity for the net. C++ language has been used to implement the model from Digital Data Embedding (DDE) Lab Binghamton web site.

In study of **Ye, Ni and Yi [18],** the network used another type of activation function called Truncated Linear Unit (TLU) which combined the selection channel inside network. The model consists of a number of layers equal to ten layers ended by two consequent fully-connected layers. Basic high-pass filters were used to initialize weights in first layer for computation of residual maps in SRM, which easily reaches better local minima as a regularize. They concluded the work with comparing to SRM and its selection-channel-aware variant max SRMd2, has yielded superior performance for all steganography techniques with different payloads. "Caffe" toolbox was chosen with two main dataset BOSSbase and BOWS2 datasets for implementation.

Salomon, Couturier, Guyeux and Bahi [19] have presented criterion with their method to make decision in two situations 1) the CNN or may be 2) spatial rich models with ensemble classifier (SRM+EC)[20] method which describes how to obtain rich features. Three methods UNIWARD, Minimizing the Power of the Most Powerful Detector (MiPOD) [21] (the chosen model is used to obtain new fundamental insight regarding the performance limits of empirical steganalysis detectors built as classifiers), and (High-pass, Low-pass, and Low-pass) (HILL) **[22]** (cost function is designed by using a high-pass filter to locate the less predictable parts in an image, and then using two low-pass filters to make the low cost values more clustered) and Tensorflow platform were chosen. Images of BOSSBase dataset was selected for experiments. Error rates reached 16%, 16% and 17% and 39%, 38%, and 41% with payload 0.4 and 0.1 bpp, respectively, for three algorithms S-UNIWARD, MiPOD, and HILL, respectively. These rates were better than those obtained by SRM+EC.

Another model **c**onsists of fewer convolutions and much larger filters in the final convolutional layer developed by **Salomon, Couturier, Guyeux and Bahi** [23]. It has capability to deal with larger images in size and payloads with lower value. 512×512 is size of image feed into the system. Firstly, it was filtered by a single kernel, it is equal to 3×3, and also ends by a layer of 64 filters with zero padding. Experiments approved that their method has outperformed the previous CNN-based stego-analytic and also defeated many state-of-the-arts for the case of the "same embedding key".

Final efforts was suggested by **Ke and Dongming** [24] which consists of multi-columns framework instead of single one in order to obtain high performance with better precisely comparing to state-of-art-model. HUGO, and S-UNIWARD algorithm with only two payloads "0.1 bpp" and "0.4" bpp" were used

for implementations. Their model improved detection accuracy which was increased by "3%" compared to Pibre et al method.

3. Architecture of the network

Convolution Neural Network (CNN) is one of the most important models of deep learning. It consists of a set of layers. Each layer formed by a number of neurons which each current and successive neurons are connected together through vector of weights that has one value. In each layer there is one type of activation function computed by neurons. The network end with one or more fully connecting layer. For each layer, one type of abstraction obtained which may be low or high level features. In first layer, low features yields like, edges, or noise, etc. While the deeper layers give the high features which was considered as the most relevant features. There are four types of layers in CNN: input, convolution, pooling and fully connected [25].

This paper discusses the effect of dual-tree complex wavelet transform pre-processing and chooses number of levels on network presented by [26]. The network presents different architecture of CNN with another type of transform, named dual tree complex wavelet during pre-processing phase before input images into system. The main task of this transform is for exploiting the difference between cover and stego images through shift variance property. The net consists of five successive convolutions layers. Each one followed by normalization and pooling layers and ends with fully connected layer. Table-1 shows first image of BOSS base dataset implement DT_CWT with four different levels, respectively to explain edge detection. The framework of network consists of three phases: Image Pre-processing, Feature Extraction and finally the classification of image as shown in Figure (1). In 1998, Kingsbury [27] introduced the Dual-Tree Complex wavelet Transform which can be considered as one of the effective approaches for implementing an analytic wavelet transforms. It handles the four issues Oscillations, Shift Variance, Aliasing, and Lack of Directionality and enhancement the Discrete Wavelet Transform (DWT) [28]. Different types of transformations are used for pre-processing data to extract features similar to work in [29].

Dual Tree Complex Wavelet (DT-CWT) uses one type of transformation to generate complex coefficients by using their real and imaginary parts to get as feature set. It is used for emphasizing the edge elements of images because it has the property of less shift variance and more directional selectivity. As result, two main problems are solved. First is preserving properties of perfect reconstruction, and second is computational efficiency with good well-balanced frequency responses. The following steps must be performed:-

1. Remove the coefficients for low frequency.
2. Reconstruct images with high pass wavelet transform.
3. For better learning, normalize image elements to have the values in the interval [0, 1].

Table 1-effect of levels in DT-CWT [25]

| Image	No.of Levels	Pre-processing image with DT-CWT
	6	
	5	
	4	
	3	
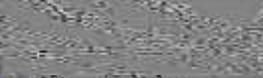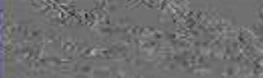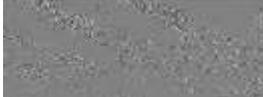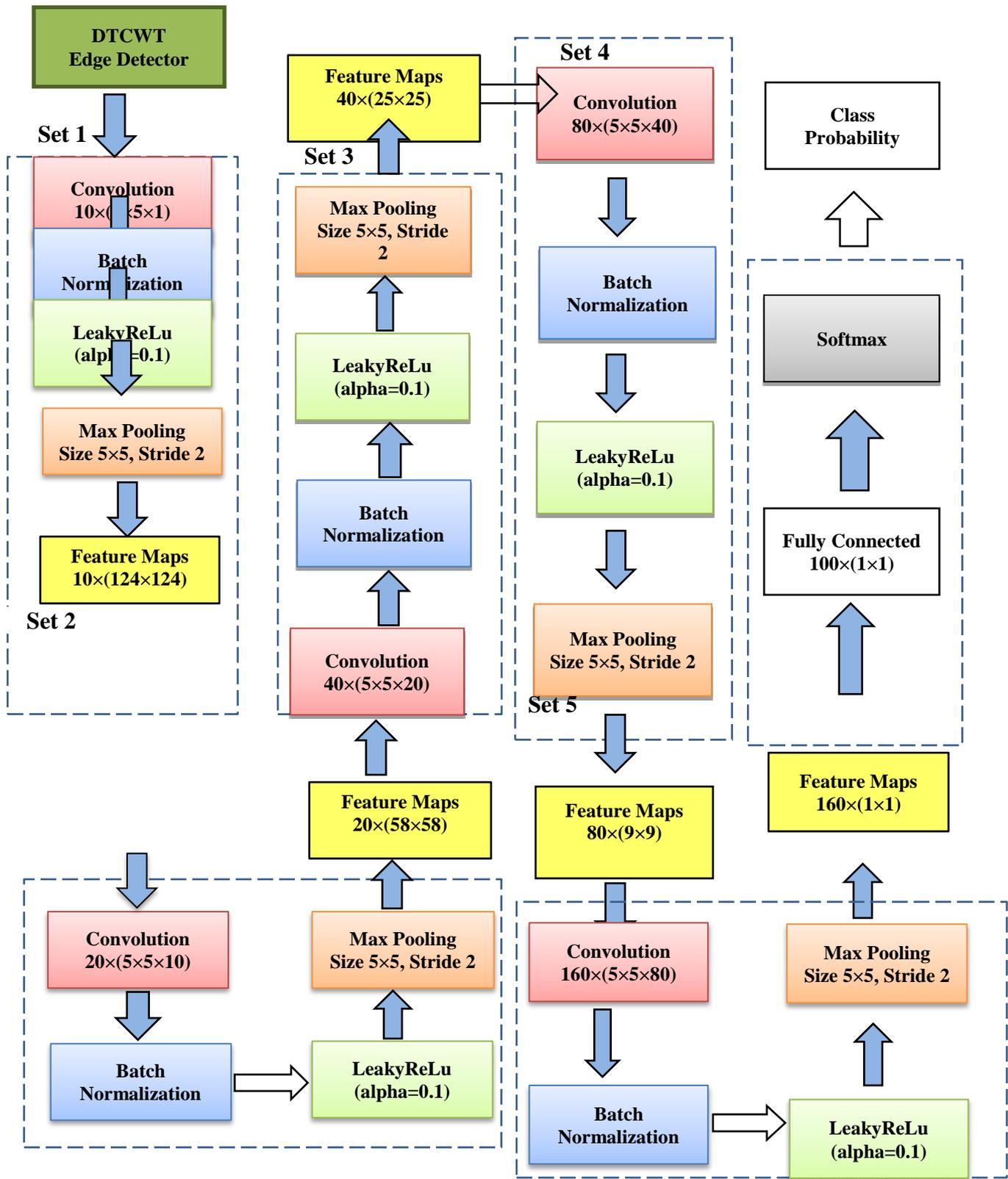
Figure 1-Convolution Neural Network of system

One of the important types of training method is named stochastic mini-batch gradient descent with setting momentum. It provides main function that computes batch. Automatically, this function can restart after each training epoch by reaching the check pointing. Also it has the capability to train data on CPU or if needs on one or more GPUs to increase speed. Below are some terms that are needed in training.

- Batch size – It defines the number of training data points for every mini batch. The significant choice will affect the estimation of the gradient for the full training dataset. On the other hand, there must have enough noise to avoid the bad local minima which in turn does not give good generalization.Batch size value set to 100 in the system.

- Number of batches – It refers to the total number of mini batches within whole training dataset. The operation of dividing the count of the total training data points by the batch size will yield the number of batches.

- Epochs – It can be represented as one full pass of training over the entire dataset. Full pass consists of forward pass in addition to one back-propagation. As a result, each epoch consists of n number of (forward pass + backpropagation) where n denotes the number of batches.

Number of epochs is set to 100 in the system.

Graphic Processing Units (GPUs), can be defined as a new revolutionized related to deep-learning world. The characteristics of GPU are outlined below:

- Better resolution in game application because it displays more screens per second.

- Higher computation especially in matrix-to-matrix multiplication. Process data is performed in parallel since there are several thousand cores of GPU utilized to complete it. As a result, it speeds up training phase.

4. Datasets

Bossbase image dataset is used for experimentswhich provides two databases of images, the BOSSBase and the BOSSRank.

- BOSSBase is composed of 8,156 never-compressed cover images coming from 7 different cameras. This database is provided as the source of cover images used for the development of steganalyzers. All images were created from full-resolution color images in RAW format. The images were first resized so that the smaller side was 512 pixels long, then they were cropped to 512×512 pixels, and finally converted to grayscale. The whole process was published in a script along with the original images in RAW format and their EXIF headers. . It includes 8 datasets, two for cover with 2 extension jpg and pgm and other 6 dataset for content-adaptive steganography HUGO, UNIWARD, and WOW algorithms with two payload 0.1 and 0.4 bpp. Use the following link to download these images, (http://info.iut-bm.univ-fcomte.fr/staff/couturie/Code/DL/images/) followed by the name of dataset.

The BOSSRank database is composed of 1, 000 grayscale images with 512×512 dimensions obtained by the same processing script. 482 of them were randomly chosen to carry the secret payload of approximately 0.4 bpp while keeping the rest without any payload. Participants did not know that 847 images were obtained by Leica M9 in RAW format and 153 images came from Panasonic Lumix DMC-FZ50 captured directly in JPEG6 format.Table 2- show first image named "1" for cover and stego in all 8 sub-folders.

Table 2-first Cover and Stego images with two different payload [12]

1.Jpg Cover	HUGO 0.1 bpp	WOW 0.1 bpp	UNIWARD 0.1 bpp
1.pgm Cover	HUGO 0.4 bpp	WOW 0.4 bpp	UNIWARD 0.4 bpp

5. Evaluation and Testing

Accuracy, sensitivity, and specificity are three important scalars used to evaluate the performance of the classifier. Difficult problems like sensitivity to data may appear if no balance. It can be easily derived a number of measures from confusion matrix for evaluating [30]. In this paper there are four

measures used for evaluation like precision, Recall, F-score, and finally Receiver operating characteristics (ROC) which yields clear explanation for generating of the ROC curve.

The main task of Confusion matrix is to evaluate the performance of any model. The purpose of any classifier is to distinguish between two instances and state its corresponding class which may be either true or false. True Positive (TP), True Negative (TN), False Positive (FP), and finally the False Negative (FN) are four main possible classifications. Equation (1) describes how to calculate this measure [31].

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \qquad \ldots\ldots\ldots (1)$$

Precision, recall and f1-score are three measurements which can be derived from confusion matrix which give a good understanding of the results. Equations (2 to 4) show how to calculate these metrics [29].

$$Precision = \frac{TP}{TP + FP} \qquad \ldots\ldots\ldots (2)$$

$$Recall = \frac{TP}{TP + FN} \qquad \ldots\ldots\ldots (3)$$

$$F - Score = 2\frac{Precision * Recall}{Precision + Recall} \qquad \ldots\ldots\ldots (4)$$

Finally, the primary graph for assessment named receiver operating characteristics (ROC) curve can be outlined as a 2-dimensional. The True Positive Rate and False Positive Rate represent the x-axis and the y –axis respectively [32].

Table-3show results obtained after training pairs of 1000 images (cover and stego) and testing image

Table 3-Precision, Recall and F-score Results with DT-CWT [25]

	Payload	Precision	Recall	F-Score
HUGO	0.1 bpp	0.98387	0.9833	0.9836
HUGO	0.4 bpp	0.98387	0.9833	0.9836
WOW	0.1 bpp	0.96875	0.96667	0.96771
WOW	0.4 bpp	0.96875	0.96667	0.96771
UNIWARD	0.1 bpp	0.98387	0.98333	0.9836
UNIWARD	0.4 bpp	0.98387	0.98333	0.9836

Table -4 shows the effect of no. of level chosen for DT_CWT to obtain more detection accuracy

Table 4-Effects of no. of levels for DT-CWT

Algorithm	No.of Levels	Precision	Recall	F-Score
HUGO 0.1 bpp	3	0.92857	0.91667	0.92258
	4	0.96875	0.96667	0.96771
	5	0.98387	0.98333	0.9836
HUGO 0.4 bpp	3	0.92857	0.91667	0.92258
	4	0.96875	0.96667	0.96771
	5	0.98387	0.98333	0.9836
WOW 0.1 bpp	3	0.90769	0.8986	0.90654
	4	0.92897	0.94765	0.94898
	5	0.96659	0.96975	0.96965
WOW 0.4 bpp	3	0.90769	0.8986	0.90654
	4	0.92897	0.94765	0.94898
	5	0.96659	0.96975	0.96965
UNIWAED 0.1 bpp	3	0.92857	0.91667	0.92258
	4	0.96875	0.96667	0.96771
	5	0.98387	0.98333	0.9836
UNIWARD 0.4 bpp	3	0.92857	0.91667	0.92258
	4	0.96875	0.96667	0.96771
	5	0.98387	0.98333	0.9836

From the above results, it can be noticed that when state number of level equal to 5, it gives more detection accuracy comparing to number of levels equal to 3. This step for pre-processing has impact to minimize time for training and testing images in comparison with the case of not using dual tree complex.

6. Implementation

MATLAB 2017b is selected to implement the network. Figure-2 shows the main front interface.
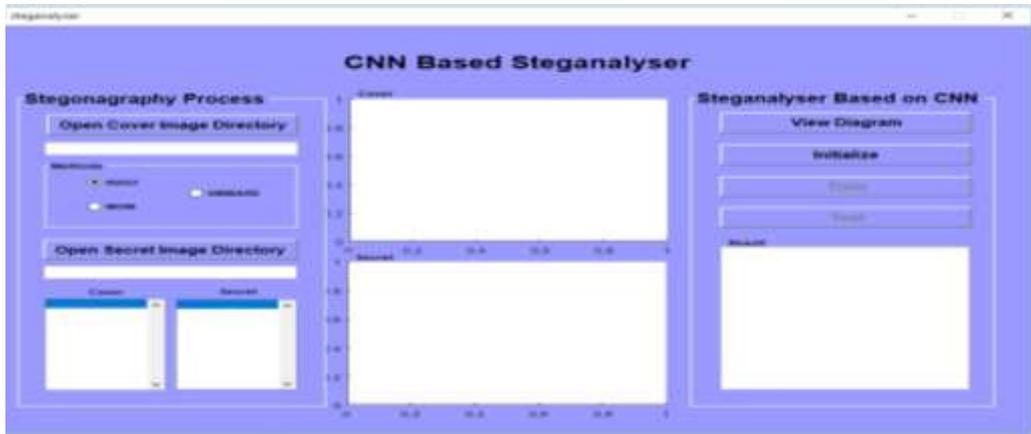
Figure 2-Front Interface

For selecting cover image directory, Click on "Open Cover Image Directory" button for select cover directory and on "Open Secret Image Directory" to select stego directory. State stego method on the suitable option buttons. See Figure-3.
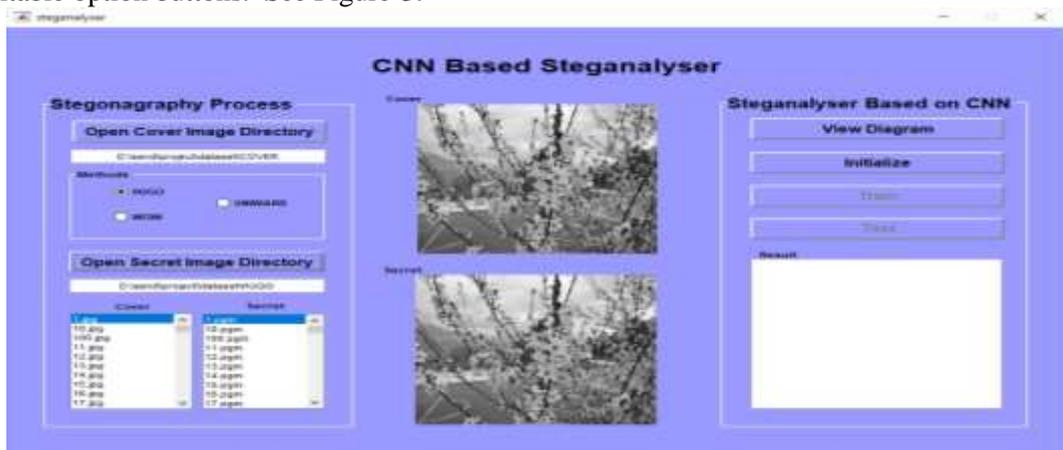
Figure 3-Select cover and stego images

After images for cover and stego are displayed, click on "**Initialize**" button to start building the CNN structure by adding layers and split data into training and testing as shown in Figure -4.
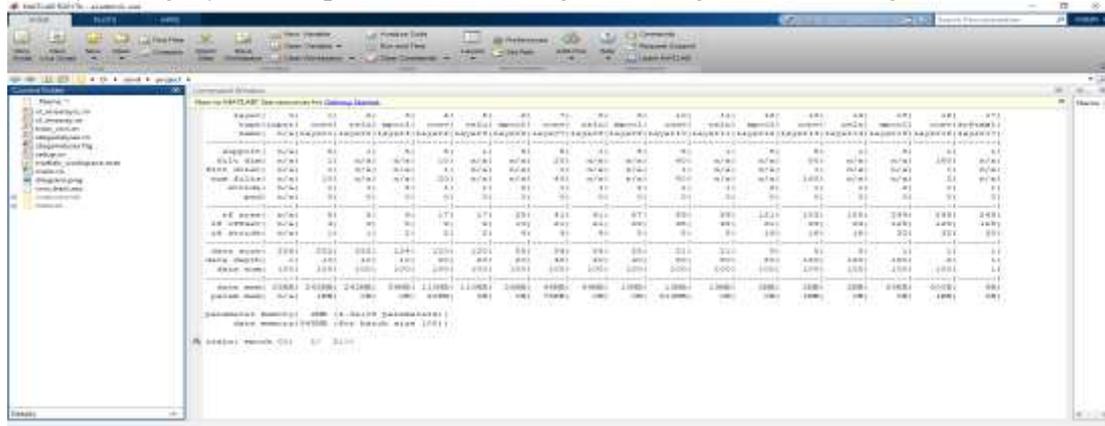
Figure 4-Build CNN

The two final steps are "Train" button used for training analyzer and showed the progress with graph and the testing step to show the results.

6. Conclusion

Image steganalysis is one important type for classification of Image. Pre-processing phase on image is an important step to give most relevant features before images feed into the system. The curse of dimensionality of features is main problem when designing universal image stego-analytic method. This paper has discussed results of assessment classification measures presented in architecture of Convolution Neural Network with Dual Tree Complex Wavelet Transform Preprocessor for Universal Image Steganalysis and has given a new test with different level of DT-CWT to show which number of levels gives better result. This phase yields little time for training and testing images comparing to system without pre-processing and also better detection accuracy comparing to previous architectures presented earlier.

For future work, we recommend to implement either one of two solutions; First: "Hybrid Model" by selecting one of Artificial Intelligent algorithms or any supervised machine learning tools like support vector machine in preprocessing phase to select good feature. This will be used to decrease dimensionality of data and thus will minimize no. of sets (CNN) for training images. Accordingly, computation and complexity of system degree means low hardware requirement. Second: "Parallel Kernel-Multi Feature", in design of CNN, after image input into system and in the convolution layer, many kernels will be built and making convolution at parallel. This will yield different feature map output from convolution, which will be selected to pass to the next layer in the net.

References

1. Parnas, D.L. **1972**. "On the Criteria To Be Used in Decomposing Systems into Modules", Communications of the ACM , vol. 15, no. 12.
2. Petitcolas, F. A. P. **2002**. "Information Hiding", 5th International Workshop, Noordwijkerhout, The Netherlands, springer, October 7-9.
3. Johnson N. F. and Jajodia S. **1998**. "Exploring Steganography: Seeing the Unseen", IEEE Computer Society, vol. 31, no. 2, pp. 26–34.
4. Alaa A. J. A., Shahrin b. S. and Mazdak Z. **2012**. "An Introduction to Image Steganography Techniques", Conference: International Conference on Advanced Computer Science Applications and Technologies, Kuala Lumpur, Malaysia, vol. ACSAT'12.
5. Fridrich, J. **2010**. "Steganography in Digital Media: Principles, Algorithms, and Applications", Cambridge University Press.
6. Hans, G.S. **2012**. "Machine Learning in Image Steganalysis", Wiley.
7. Nissar, A. and Mirb, A.H. **2010**. "Classification of steganalysis techniques: a study", Digital Signal Processing, vol. 20, no. 6, pp. 1758–1770.
8. Pevn`y, T., Filler, T. and Bas, P. **2010**. "Using high-dimensional image models to perform highly undetectable steganography," in Information hiding, 161–177, Springer.
9. [9] Holub, V. and Fridrich, J. **2012**. "Designing steganographic distortion using directional filters" in [The IEEE International Workshop on Information Forensics and Security (WIFS)], pp: 234–239.
10. Holub, V. and Fridrich, J. **2013**. "Digital image steganography using universal distortion" in , Proceedings of the first ACM workshop on Information hiding and multimedia security, 59–68, ACM.
11. Fridrich J., Kodovsky J., Holub V. and Goljan M. **2011**. Steganalysis of Content-Adaptive Steganography in Spatial Domain, , Information Hiding 13th International Conference, Prague, Czech Republic, May 18-20.
12. Tan. T. and Bin, L. **2014**. "Stacked convolutional auto-encoders for steganalysis of digital images", Signal and Information Processing Association Annual Summit and Conference (APSIPA), Asia-Pacific, conference: 9-12 Dec.

13. Patrick B., Tomas F. and Pevny T. **2011**. "Break Our Steganographic System": The Ins and Outs of Organizing BOSS, Information Hiding 13th International Conference, IH, Prague, Czech Republic, May 18-20.

14. Qian,Y. Dong J., Wang, W. and Tan T. **2015**. "Deep learning for steganalysis via convolutional neural networks," in Proceedings of the IS&T/SPIE Electronic Imaging, p. 94090J, 2015, San Francisco, Calif, USA.

15. Pevný T., Bas, P. and Fridrich J. **2010**. "Steganalysis by subtractive pixel adjacency matrix". IEEE Transactions on Information Forensics and Security, 5(2):215–224, June.

16. Pibre L., Jérôme, P., Ienco, D. and Chaumont, M. **2016**. "Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover source-mismatch, Media Watermarking, Security, and Forensics, IS&T Int. Symp. on Electronic Imaging, SF, California, USA, 14-18 Feb.

17. Holub V., Fridrich J. and Denemark T. **2014**. "Universal distortion design for steganography in an arbitrary domain". EURASIP Journal on Information Security, Special Issue on Revised Selected Papers of the 1st ACM IH and MMS Workshop.

18. Ye J., Ni J. and Yi Y. **2017**. "Deep learning hierarchical representations for image steganalysis," IEEE Transactions on Information Forensics and Security, TIFS, vol. 12, no. 11, pp. 2545–2557, Nov.

19. Jean-Francois C. , Rapha¨el C. and Michel S. **2017**. Improving Blind Steganalysis in Spatial Domain using a Criterion to Choose the Appropriate Steganalyzer between CNN and SRM+EC, International Conference on ICT Systems Security and Privacy Protection.

20. Kai L., Jianhua Y. and Xiangui K. **2017**. "Ensemble of CNN and rich model for steganalysis", International Conference on Systems, Signals and Image Processing (IWSSIP), IEEE, 22-24 May.

21. [21] Sedighi V., Cogranne R. and Fridrich J. **2016**. "Content-adaptive steganography by minimizing statistical detectability". IEEE Transactions on Information Forensics and Security, 11(2), pp: 221– 234.

22. Li B., Wang M. and Huang J. **2014**. "A new cost function for spatial image steganography". In Proceedings IEEE, International Conference on Image Processing, ICIP, Paris, France, October 27–30.

23. Salomon M., Couturier R., C. Guyeux, Couchot J.F. and Bahi J.M. **2017.** "Steganalysis via a convolutional neural network using large convolution filters for embedding process with same stego key", vol. 6, Issue 2, pp: 79-92, July.

24. Ke Q. and Dongming L. **2018**. "Image Steganalysis via Multi-Column Convolutional Neural Network", Guangzhou University, Feb.

25. Dong Y. and Li D. **2013**. "Deep Learning: Methods and Applications", Foundations and Trends R_ in Signal Processing, Vol. 7, Nos. 3–4 , 197–387, the essence of knowledge, Now.

26. Hanaa,M.A. and Halah, H.M. **2019**. "Convolution Neural Network with Dual Tree Complex Wavelet Transform Preprocessor for Universal Image Steganalysis", Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.11(3).

27. Kingsbury N.G. **1999**. "Image processing with complex wavelets," *Philos. Trans. R. Soc. London A, Math. Phys. Sci*., vol. 357, no. 1760, pp. 2543–2560

28. Kingsbury N.G. **2001**. "Complex wavelets for shift invariant analysis and filtering of signals," *Appl. Comput. Harmon. Anal.*, vol. 10, no. 3, pp. 234–253, May.

29. Israa M. H. **2018**. Wavelet Energy and Shape Features for Plants Recognition, Iraqi Journal of Science, Vol. 59, No.2B, pp: 956-962

30. Shaikh S. **2011**. "Measures derived from a 2 x 2 table for an accuracy of a diagnostic test" 2 , 128, Journal of Biometrics & Biostatistics.

31. Alaa T. **2018**. "Classification assessment methods", Applied Computing and Informatics, sciencedirect.

32. Fawcett T. **2006**. "An Introduction to ROC Analysis", Pattern Recognition Letters, 27 , pp: 861-874.