Convolution Neural Networks for Blind Image Steganalysis: A Comprehensive Study

Hanaa Mohsin Ahmed [1]

University of Technology-Iraq
Department of Computer Science
salman.hanaaa@yahoo.com

Halah Hasan Mahmoud [2]

University of Baghdad
Computer Center
halah@uob.edu.iq

Abstract:
Recently, Convolution Neural Network is widely applied in Image Classification, Object Detection, Scene labeling, Speech, Natural Language Processing and other fields. In this comprehensive study a variety of scenarios and efforts are surveyed since 2014 at yet, in order to provide a guide to further improve future researchers what CNN-based blind image steganalysis are presented its architecture, performance and limitations. Long-standing and important problem in image steganalysis difficulties mainly lie in how to give high accuracy and low payload in stego or cover images for improving performance of the network.

Keywords: deep learning, convolution neural networks, blind image steganalysis, payload

Hanaa .M/ Halah .H

1. Introduction

For solving difficult real-world problems quickly, deep learning techniques which is most important sub-field of machine learning has been used for classification and regression problems. High performance of computer-hardware and huge dataset lead to training deeper models. Today, deep learning implemented in many area of computer science like: Classification and Recognition for Image and Video, Natural Language Processing, Management Relationship of Customer, Recommendation Systems, and Bioinformatics (Heaton, 2015).

The study of hiding secret information into innocent looking cover media like the digital images, videos, audios, and texts in such a way that existence of message is not visible is known as Steganography (Baldominos, 2018). Covered and writing are meaning associated to "Steganos" and "Graphy" respectively. The cover file with secret message hidden in it is referred as "stego" file. Among the various media types the digital images are mostly used for communication as they contain redundant bits and are most widely used on internet. Steganography apart from having positive applications such as smart identity cards which holds individuals confidential information, circulating secret data of a company, medical imaging systems etc. has also been found to be utilized by anti-social elements for terrorist activities (He *et.al*, 2016).

Steganalysis is the counterpart of steganography, which wants to identify if there is data hidden inside a digital medium. According to this there is a changing in properties of type statistical for cover image while implementing steganography algorithms. Steganalysis can be used to detect these changes which are required high computation. The traditional is the first one of steganalysis which primary concern with hand-extracted features and machine learning classifier while other is based on deep learning technology appears as a result to major increase in data huge and computing performance.

According to steganography algorithm, steganalysis has been broadly studied under two main types; firstly "specific" and secondly "blind" steganalysis. Aim of specific steganalysis is to detect information hidden in a carrier by having prior knowledge of underlying steganography algorithm. This makes it very accurate but at the same time it is not capable of detecting new steganography algorithms. To overcome this shortcoming blind steganalysis has been introduced which aims to detect any known or unknown steganography algorithms (Dong.*et.al*, 2016).

2. Convolution Neural Network Architecture

This is a popular deep learning model which is inspired by the actual visual perception model of the living creatures. In 1959, Hubel & Wiesel (Fukushima *et.al*, 1982) discovered that cells in animal visual cortex hold the responsibility for the detection of light in receptive areas. Based on that discovery, K. Fukushima presented recognition in 1980 that might be considered the base of the Convolutional Neural Network (CNN) Lecun, *et.al*, 1998). Several approaches were introduced for overcoming the issues that are encountered in training deep Convolutional Neural Network. Since the year of 2012, CNNs proved to be of high efficiency in complicated Pioneered by LeCun (Anselmo *et.al*, 2017) and image classifications like gender detection (Hanaa *et.al*, 2018), the most significant advantages Convolutional Neural Networks according to traditional fully-connected NNs is:

1) Post training, it is possible using the result of a network's layer as property vector combined with an external classifier, instead of simply depending on net classification layer

2) Convolutional layers are comprised with small size kernels allowing a sufficient extractionn of high-level properties which are input into fully-connected layers.

3) Reduced amount of parameters to be learned.

4) Mis-classification error drives the weights altering of each of the convolutional and fully-connected layers.

5) Training of a Convolutional Neural Network is accomplished via back-propagation and stochastic gradient descent.

Basic layers of Convolutional Neural Networks are recorded as follows (Taigman *et,al*, 2014):

1) **Input layer:** in which data is input to the network. This data may be raw image pixels or their transformations, whichever more efficiently emphasizes some certain image elements.

2) **Convolution layers:** include a set of filters that have preset size utilized for performing convolution on the data, producing a feature map. These filters may highlight some patterns which are helpful for image characterizing, like edges, textures, and so on.

- **Number of filters**: More filters give to a convolutional layer, the more features it can detect.

- **Filter Size**: or "Kernel" or "Feature detector," size of filter matrix.

- **Stride:** slide filter matrix over the input matrix according to specific number of pixels. If equal to 1, then moving filters one

Hanaa .M/ Halah .H

pixel at a time. Having a larger value of it will produce smaller feature maps.

- **Padding**:fundamentally including columns (on a level plane) or (vertically) columns of zeros to the borders of image to regulator yield measure of the layer in which it is included.

- **Activating or Transferring Functions:** establish bounds for the result of neurons. Neural Networks may utilize several distinct activation functions. Choosing an activation function for your neural network is a significant consideration because it may impact the way you must format input data. Table 1 shows different types of activation functions and corresponding uses (purposes) of it (He *et.al*, 2015).

3) Pooling layers: Selectively, this operator may be found following a convolutional layer. The reason of pooling is the reduction of dimensions of the input via achieving down-sampling, taking the place of part of a property map in a specific index with a statistic summary of the near locations. The most popular of the examples is the max-pooling, in which the input is reduced via taking a sub-tensor of the property map and replacing it by its max or average value (Gong et.al, 2015).

4) Fully connected layers: utilized for the purposes of understanding patterns that are produced by the preceding layers. Neurons in this layer are of full connections to all activations in the +preceding one. They are referred to as the inner product layers as well. After training, transfer learning methods are capable of extract properties in these layers for train other classifiers (Taigman *et.al*, 2014). Figure 1 shows layers in CNNs.
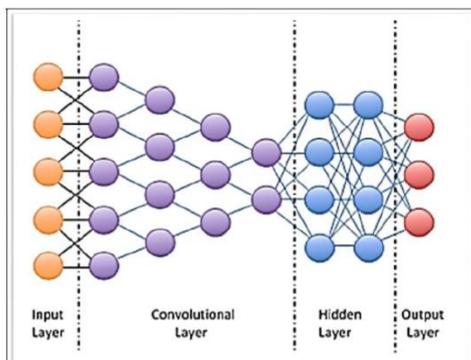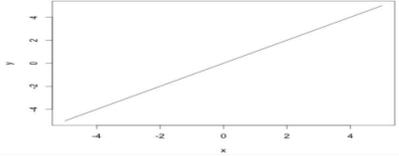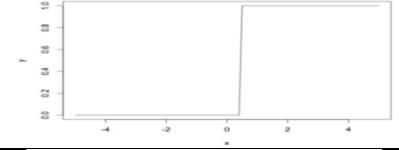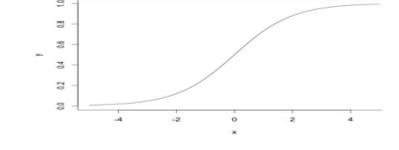

Figure (1) Layers of convolution neural networks (Taigman *et.al*, 2014).

Hanaa .M/ Halah .H

Table (1): Mathematical Equations of Activation Functions (He *et.al*, 2015)

Function Type	Math Equation	Characteristics	Graph
Linear	$\emptyset(x) = x$	At all, value of neuron doesn't update.	
Step Or threshold	$\emptyset(x) = \begin{cases} 1, if\ x \geq 0.5 \\ 0, otherwise \end{cases}$	It is simple. For fixed threshold, yields 1 (true) for values that are above it.	
Sigmoid or Logistic	$\emptyset(x) = \dfrac{1}{1 + e^{-x}}$	Results simply number value positive. With feed-forward NNs it is general. It makes sure values stay within a relatively small range.	
Hyperbolic Tangent	$\emptyset(x) = \tanh(x)$	For output results in range [-1, 1] it is most suitable.	
Rectified Linear Unit (ReLU)	$\emptyset(x) = \max(0, x)$	It is a linear, non saturating function. ReLU does not saturate to -1, 0, or 1. Moves towards and eventually attains a value can be defined as saturating.	
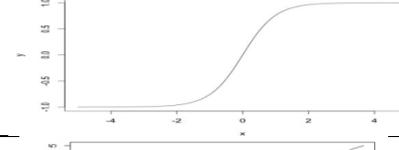
Hanaa .M/ Halah .H
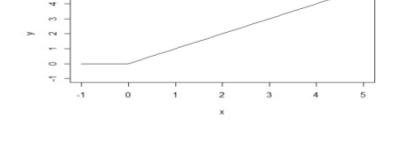
3. Related Works

Recently, there are numerous of studies adopted deep learning especially convolution neural network. Our motivation is to review different optimization of CNN architectures and some details for solving blind image steganalysis.   The primarily goal of improving design of network is to obtained more accuracy and performance and reduce complexity.

In the field of steganalysis, (Tan *et al*., 2014) Propose network composite of stack of convolutional auto-encoders. HUGO stego algorithm with payload equal to 0.4 bit per pixel (bpp), and Bossbase dataset is used in its experiments. This method has been failed to compete Spatial-domain Rich Model (SRM).

Two steps of traditional approach have been replacement by model of (Qian *et al.* 2015) presented. dropout technique is applied to reduce the problem of over-fitting. The algorithm unifies feature extraction and classification under a single network architecture, and the parameters in the two steps are automatically optimized by the back-propagation method at the same time. HUGO, WOW, and S-UNIWARD algorithms are used for testing which yields detection accuracy percentage of 3% to 4% less than to Ensemble Classifier and SRM features.

In 2016, framework of Qian's has examined by (Pibre *et al.*,2016) For improvement performance of detection in the scenario of reusing the same embedding key for different images by reducing error more than 16% with S-UNIWARD at 0.4 bpp payload. To obtain diversity in this net 64 and 16 filters in first and second layer respectively. For embedding, Steganography algorithms have been implemented using C++ which can be downloaded from DDE Lab Binghamton web site.

Rather than tasks of conventional computer vision a somewhat model relatively different presented by (Ye *et al.*, 2017). the model is able to well replicate and optimize residual computation, feature extraction, and binary classification in a unified framework and extracts hierarchical feature representations directly from raw images. Ten layers constitutes net and ends with a fully-connected layer followed with a 2-way softmax, which produces the distribution over 2 class labels. To capture the structure of embedding signals effectively, truncated linear unit is introduced to the first layer, while rectified linear unit activation function is applied to the other layers. Approach achieves superior performance across SRM and maxSRMd2 for a variety of payloads.

Also, (Salomon *et al.*, 2017) designed a CNN-based steganalyzer for stego images with a unique embedding key. The proposed structure embeds less convolutions but with much larger filters in the final convolutional layer, which can deal with larger images and lower payloads. The architecture took 512×512 image as input image, and the input image was firstly filtered by a single kernel of size 3×3, then followed by a layer of 64 filters as large as possible with zero padding. The experiments showed that the proposal outperformed the existing CNN-based steganalyzer and defeated many state-of-the-art image steganography schemes in case of the "same embedding key".

On the other hand, multi-columns framework instead of single one get high performance and better precisely regarding to state-of-art- model is presented by (Ke *et al.*, 2018) Two different payload "0.1 bpp" and "0.4" bpp for embedding message has been chosen with HUGO, and S-UNIWARD adaptive algorithms. Detection accuracy for stego and cover is increased by "3%" compared to Pibre et al. and Couchot et al works.

Table 2 illustrates Characteristics Obstacles and solutions of related works for convolution neural network tools to design blind image steganalyzes.

A. Steganographic algorithms

Depending on the domain of embedding message and available information, steganographic algorithms have been categorized as three types: **Spatial**: embedding messages by modifying pixel values. **JPEG**: embed process into quantized DCT coefficients and **Side-informed**: utilize the knowledge for cover image and then non quantized DCT coefficients and rounding errors. Spatial domain include following types.

- Least Significant Bit: Non-adaptive and applied together with embedding "ternary matrix".
- Edge-Adaptive: A change of pairs of pixels has been kept large in absolute value.
- HUGO: purpose of it is for minimizing distortion in embedding within "high-dimensional" of features added of 4 pixels neighboring. It is content-adaptive method.
- WOW: distortion for embedding is evaluated by using "wavelet filter". It is a highly content-adaptive and dissimilar of HUGO.
- S-UNIWARD: the case of "UNIWARD" spatial domain based on wavelet, likely to WOW.

B. Datasets

On the other hand, to build network of blind image steganalysis datasets must be used through three phases of creation: *Firstly*, *Training dataset*, primarily model has been fit on training dataset to fit parameters like connection weights. It consists of input (scalar) and output (label) vectors. The result have yield after running model with suitable dataset and check with target for similar input vector in this dataset. According to this result and specific algorithm selected parameters are adjusted. *Secondly*, *Validation dataset* is used to predict responses. The main evaluation obtained

here through iterations of hyper-parameters is unbiased model. Last of all dataset named *Test* or sometimes *holdout* delivers "unbiased evaluation" for final model fit on training set.

Primarily Images Datasets consisting of images or videos are used for solving different problems like classification, object detection and facial recognition. Bossbase, ImageNet, INRIA (Hanaa al.et, 2018) are most popular dataset which has been choose for building models regarding blind image steganalysis. See Table (3) and Table (4).

Hanaa .M/ Halah .H

Table (2) Characteristics Obstacle and solution of works designing a blind image stegabalysis using Deep learning tool

Work	Characteristics	Obstacle	Solution
Tan *et al*, 2014	• Filters have been set with ranodomize value. • Phase for extraction of features are systematically and significantly better. • Diversity due to special structure of the filters.. • Error rate of 48% • Ajust convolution kernel through backpropagation calculation. • Implementation using DeepLearn Toolbox.	• stuck at localplateaus or diverege lead to poor solution in case of randomize initalization.	• Multiplied kernel of first convolution by linear filter called shift-invariant-impulse response. • The non linear activations that is used "Sigmoid".
Qian *et al*, 2015	• Image processing, convolution and fully connected layers consistued net. • Operations that have been applied in convolution layer are convolution, non linearity and pooling. • "Softmax" activation for class label to get more distribution. • Implementation using C++ GPU convolution library.	• Image content may be impacted due to added stego noise to cover. • Prediction error value is not better	• Same filter for image processing layer. • Increase performance can be done by sharing parameters. • Better prediction error with "Gaussian" non linearity comparing to "Signmoidal".
Pibre *et al*, 2016	• Good chosen for CNN structure. • In learn process, one global optimization. • Optimized have been done with kernels of filter. • Compare CNN and Fully-Connected Networr to Rich Model with an Essemle Classifier.	• "Gaussian" unimprove result. • Pooling have been canceled due information loss. • Key in security error weaks accepting steganography.	• "ReLU" produce positive value. • "Robustness" mismatch of cover-source issue due to transformation by first two layers and change propability of key. • Diversity due to large number of filters.
Ye *et al*, 2017	• Ten layers and 2 labels for class resulted from fully-connected layer with "Softmax" for two ways. • For each convolution layer "one type of non linear activation" applied. • Between first to three layers, operation o pooling have been canceled. • Implementation using MatLab.	• Equal or more than two fully-connected layers (too trianed parameters). • Overfitting.	• Implement Pipeline of Steganalysis. • Suffusient for small set. • Batch normalization and local constract normalization used in network.
Salomon et al, 2017	• "Hyperbolic tangent" is chosen in convolution which consists of two layers. • Size of kernel for first layer is 3×3. • Zero-padding with 64 filters. • Features map equal to 256. • Cancelled pooling for two convolution layers.	• Kernel optimum to used first filter has no certainty. • Not reach Converge.	• For hidden message, detection of sequential images due to result two "Softmax" neurons for one fully connected layer.
Ke et al, 2018	• architecture consist of 3 parallel (columns) CNNs. • Unalike filter sizes (3×3, 5×5, 12×12 ,15×15), complex stego information. • A single output layer with two softmax neurons. • DDE Lab offer by Binghamton web site with C++ for implementation. • "Dropped" pooling operation of all layers.	• Unsuitable for stegananlysis because average operation is computed in MDNNs. • Less no.of features maps. 59	• Increase performance due to (ReLU) for each convolution layer. • Reduced number of filters with large size of it is better choice for minimze complexity of computation. • "Estimate of distortion" solved by normalize images input with same size.

Hanaa .M/ Halah .H

Table (3) steganography and dataset for related studies

work	Stego algorithm	payload	Dataset	Trianing	Testing
Tan *et al*	HUGO	0.4 bpp	Bossbase 1.01, 10,000 (512*512) grey-scale cover images	5000	5000
Qian *et al*	HUGO, WOW, S-UNIWAED	0.3, 0.4, 0.5 bpp	Bossbase 1.01, 10,000 (512*512) grey-scale cover images. ImageNet 15,000,000	100,000	
Pibre *et al*	S-UNIWARD	0.4 bpp	Cropped Bossbase 1.01, 10,000 (256*256*8bit) grey-scale. LIRMM-Base, 1008 grey-level image	40,000	5000
Ye *et al*	S-UNIWARD, HILL, WOW	0.2 bpp	Bossbase 1.01, 10,000 512*512 grey-scale cover images. BOWS2 downsampled , cropped and greyscale 512*512*8-bit	5000 1000	5000
Salomo n et a	HUGO, WOW, and J-UNIWARD	0.1, 0.4 bpp	Bossbase 1.01 10,000 (512*512) grey-scale cover images. Raise database	4000	2000
Ke et al	HUGO, S-UNIWARD	0.4, 0.1 bpp	Bossbase 1.01, 10,000 512*512 grey-scale cover images. RAISE database.	40,000	20,000

Table (4) Feature extraction and classification steps for related studies

Paper	Input layer	Feature Extraction			Classifcation	
		Convolution layer Kernal	Poolying layer	feature maps	Fully connected	Output
Tan *et al, 2014*	1st stage: 512×512	40 (5×5)	Max-pooling, 4×4 , down sampling	40(128×128)		
	2nd stage: 40(128×128)	10 (5×5)	Max-pooling, 4×4 , down sampling	400 (32×32)		
	3rd stage: 400 (32×32)		Max-pooling, 4×4 , down sampling	4000(8×8)	3000 neurons	2 neurons
Qian *et al, 2015*	1st :256×256	5×5		16 (124×124)		
	2nd: 16 (124×124)	16 (5×5)	Gaussian non-linearity, Average-pooling, window size:3×3, step: 2	16(61×61)		
	3rd: 16 (61×61)	16 (3×3)	Gaussian non-linearity, Average-pooling, window size:3×3, step: 2	16(29×29)		
	4th: 16(29×29)	16 (3×3)	Gaussian non-linearity, Average-pooling, window size:3×3, step: 2	16(13×13)		
	5th: 16(13×13)	16 (3×3)	Gaussian non-linearity, Average-pooling, window size:3×3, step: 2	16(4×4)		
	6th: 16(4×4)	16 (5×5)	Gaussian non-linearity, Average-pooling, window size:3×3, step: 2	256	FC1:128 neurons FC2:128 neurons	Non-saturating, non-linearity, ReLU 2 neurons ,

Hanaa .M/ Halah .H

| Paper | Input layer | Feature Extraction | | | Classifcation | |
		Convolution layer Kernal	Poolying layer	feature maps	Fully connected	Output
						Two-way softmax
Pibre *et al, 2016*	1st :256 ×256	64 (7×7), ReLU, stride:2	No pooling layer	64 (127×127)		
	2nd :64 (127×127)	16 (5×5), ReLU	No pooling layer	16 (127×127)	FC1:1000nerurons	ReLU
					FC1:1000nerurons	ReLU , Softmax
Ye *et al, 2017*	1st :256 ×256	30(5×5×1),TLU, stride:1		30(252×252)		
	2nd:30(252×252)	30(3×3×30),ReLU, stride:1		30(250×250)		
	3rd:30(250×250)	30(3×3×30), ReLU, stride:1		30(248×248)		
	4th: 30(248×248)	30(3×3×30), ReLU, stride:1	Max-pooling 2-2, Stride:2	30(123×123)		
	5th: 30(123×123)	32(5×5×30), ReLU, stride:1	Max-pooling 3-3, Stride:2	32(59×59)		2 neurons , Two-way softmax
	6th: 32(59×59)	32(5×5×32)), ReLU, stride:1	Max-pooling 3-3, Stride:2	32(27×27)		
	7th: 32(27×27)	32(5×5×32)), ReLU, stride:1	Max-pooling 3-3, Stride:2	32(11×11)		
	8th: 32(11×11)	16(3×3×32), ReLU, stride:1		9(16×16)		
	9th: 9(16×16)	16(3×3×16), ReLU, stride:1		3(16×16)		2 neurons , Two-way softmax
	10th: 3(16×16)	(3×3×1)		510×510		
Salomon et a, 2017	512×512	64(509×509)	hyperbolic tangent	64		
	510×510		Zero-padding, Stride:1	256		2 neurons , Two-way softmax
Ke et al, 2018	Col1:img	16(3×3), ReLU	Dropped	30(7×7)		
		30(7×7), ReLU	Dropped			

Hanaa .M/ Halah .H

Paper	Input layer	Feature Extraction			Classifcation	
		Convolution layer **Kernal**	Poolying layer	feature maps	Fully connected	Output
	Col2:img	20(7×7), ReLU	Dropped	40(5×5)		
		40(5×5), ReLU	Dropped			
	Col3:img	24(5×5), ReLU	Dropped	40(3×3)		
		40(3×3), ReLU	Dropped			
	30(7×7),40(5×5), 40(3×3)	1×1				2 neurons , Two-way softmax

BN: Batch Normalization, TLU: Truncate Linear Unit

4. Conclusion

The study mainly focuses on the comparison of number of steganalyzers built since 2014 year designed to predict stego images for universal image steganalysis techniques using convolution neural network tool. It includes characteristics, whole model and each layer and activation function to obtain better performance. Different popular image datasets are used through experiments to validate model with more than payload according to specific steganography algorithm selected for experimental.

5. References

[1] J. Heaton, "Deep Learning and Neural Networks", Heaton Research, Inc., 2015.

[2] A.o Baldominos, Y. Saez, P. Isasi, "Evolutionary Convolutional Neural Networks: an Application to Handwriting Recognition" , Elsevier , Neurocomputing, vol. 283, 29 March pp. 38-52, 2018

[3] K. He, X. Zhang, S. Ren and, J. Sun, "Deep residual learning for image recognition", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 770-778, 2016.

[4] C. Dong, C. C. Loy, K. He, and X. Tang,. "Image super-resolution using deep convolutional networks", TPAMI , vol.38 no.2, pp: 295-307, 2016.

[5] K. Fukushima, S. Miyake, "A self-organizing neural network model for a mechanism of visual pattern recognition", Competition and Cooperation in Neural Nets, Proceedings of the U.S.-Japan Joint Seminar held at Kyoto, Japan February 15–19, pp: 267–285, 1982

[6] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, issue. 11, 1998, pp. 2278–2324, 1998.

[7] F. Anselmo, G. Gilson, "Convolutional Neural Network approaches to granite tiles classification" ,   Expert Systems With Applications 84 .May, 2017.

[8] M. A Hanaa, H. Mahmoud Halah, "Effect of Successive Convolution Layers to Detect Gender", Iraqi Journal of Science, vol. 59, no.3C, pp: 1717-1732, 2018

[9] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, "Deep face: Closing the gap to human-level performance in face verification", in: Proceedings of the conference on Computer Vision and Pattern Recognition, IEEE, Columbus, USA, pp:1701-1708, 2014.

[10] K. He, J. Sun, " Convolutional neural networks at constrained time cost", IEEE Conference on Computer Vision and Pattern Recognition, pp:5353-5360, 2015.

[11] Y. Gong, L. Liu, M. Yang, L. Bourdev, "Compressing deep convolutional networks using vector quantization", International Conference on Learning Representations ICLR May 7 - 9, in San Diego, CA, 2015

[12] T. Tan , L. Bin "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. 2014,

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

[14] L. Pibre, P. Jérôme, D. Ienco, M. Chaumont, "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. 2016.

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

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

[17] Q. Ke , L. Dongming , "Image Steganalysis via Multi-Column Convolutional Neural Network", Guangzhou University , Feb 2018.

[18] Hanaa M. A, Halah H. Mahmoud, "Review on Image Steganalysis using INRIA Dataset", Journal of Al-Nahrain University, vol. 21 no.4 December, 2018, pp:94-108.

Hanaa .M/ Halah .H

الشبكات العصبية التلافيفية لتحليل الاخفاء الاعمى في الصورة : دراسة شاملة

هاله حسن محمود ²   هناء محسن احمد ¹
مركز الحاسبة   قسم علوم الحاسوب
جامعة بغداد   الجامعة التكنولوجية

المستخلص:

يتم اليوم تطبيق الشبكات العصبية التلافيفية على نطاق واسع في تصنيف الصور، اكتشاف الكائنات، وضع العلامات على المشهد، الكلام ، معالجة اللغات الطبيعية وغيرها من المجالات. في هذه الدراسة الشاملة ، تم استطلاع مجموعة متنوعة من السيناريوهات والجهود منذ عام ٢٠١٤ حتى الآن ، من أجل توفير دليل لمواصلة تحسين الباحثين المستقبليين حول ما يقدمه التحليل البصري للصورة العمياء القائم على الشبكات العصبية التلافيفية ، والهندسة المعمارية والأداء والقيود. تكمن المشكلة القديمة والمهمة في صعوبات تحليل الصورة بشكل أساسي في كيفية إعطاء دقة عالية وحمولة منخفضة في صورة أو تغطية صور لتحسين أداء الشبكة.