UDC 004.8:519.2

R.V. Davydenko, O.M. Suprun, PhD
(Vocational College of Engineering and Management
of the National Aviation University, Ukraine)

Artificial neural network for blind JPEG steganography method

A blind method of detecting embedded steganographic information in digital images is considered. The feature space of such an image is given on the basis of point estimates of the distribution of detailed wavelet coefficients. An artificial neural network to analyze the image database and divide it into two classes - stego and containers had been proposed and described in the article.

Hiding messages in images (steganography) is used for both legal and illegal purposes. Detecting hidden messages in images stored on websites and computers (steganalysis) is very important area of cyber forensics. Hiding information in image and sound files, as well as other types of data on computers, is becoming more common these days. Automating the detection of hidden messages is a mandatory requirement, as the large number of images stored on computers or websites makes it impossible to examine each image individually.

Models that describe the processes of embedding and retrieving data can be represented as a communication channel. In this case, the cover image acts as a communication channel through which data is transmitted, and the embedded message plays the role of the data stream being sent. Thus, the cover image can be considered as noise; simple initial steps are devoted to model the image as a realization of Gaussian noise. Under these signal detection assumptions, theoretical information concepts can be used to design the embedding algorithms, retrieval algorithms, and message detection algorithms. An artificial neural network is used to classify samples of data on features extracted from the cover-image and data of this stego-image [1].

The algorythm of the blind method of steganography was first described by Memon and Farid and was presented as follows. First, a three-level image wavelet transformation was performed, then for the detailed wavelet coefficients (horizontal, vertical and diagonal subbands), the values of average, variance, excess and asymmetry coefficient were found. Also the vector of error prediction of wavelet coefficients based on the values of adjacent coefficients and also point statistical characteristics were calculated. Thus, a 72-bit characteristic vector of the image was obtained. Then a blind detector based on a 72-bit characteristic vector was trained on the basis of containers and stacks. After that, the classification was performed, using the Fisher linear discriminator. With the development of steganographic algorythms, this method has lost its relevance, but the methodology is used in all blind algorythms.

The choice of features is one of the most important stages in the construction of a blind method of detecting steganographic information. The pixel space of the image is converted into a feature space and the definition of the built-in message is already in the feature space.

Signs for classification are tha following: the first four statistical moments of the coefficient value of the wavelet transform subband; the first four statistical

moments of error in estimating the linear prediction of the values of the coefficient value of the wavelet transform subband; selection of image quality indicators; selection of feature values based on the distance between the values of the parameters of calibrated images; parameters estimated by selecting the generalized Gaussian distribution to the values of the wavelet subband; a feature based on the Xi-square attack proposed by Westfeld [2].

During the performance of any blind algorithm of the built-in information detection the following stages take place:

1) construction of a multidimensional space of image features;

2) analysis of differences between the original images and stego in the space of features;

3) classification of the database of source images and stego into two groups;

4) assigning the analyzed image to the image-container or to the stego according with the results of 2 and 3.

Stegoanalysis methods based on neural networks are still insufficiently studied. They use statistical methods of stegoanalysis combined with the ability of neural networks to learn and classify. Existing methods are based on algorithms for teaching with a teacher. Frequency forms of image representation obtained by wavelet transforms are used as input data vectors. In [3] it was proposed to use RBF neural networks for stegoanalysis. This method refers to the "blind" methods of detecting embedded information.

The best results were obtained using a multilayer neural network of direct spread. These types of networks are formed by multilayer perceptrons, in which each computing element uses a threshold or sigmoidal activation function. Theoretically, a multilayer perceptron can form arbitrarily complex decision boundaries and implement arbitrary Boolean functions.

The development of an efficient backpropagation learning algorithm for determining weights in a multilayer perceptron has made these networks most popular with researchers and users of neural networks.

The best classification parameters were obtained for a two-layer neural network with 150 neurons in the first layer (fig. 1).

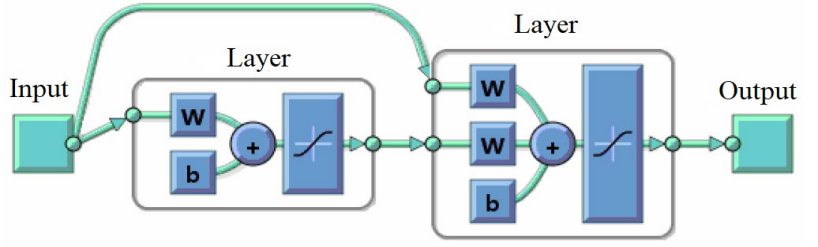

Fig. 1. Neural network model.

The structure of convolutional neural networks (CNN) is widely used in solving the problems of detecting spatial dependencies in digital images, and the peculiarities of the formation of CNN can reduce the number of parameters and improve the quality of feature detection. In the general case, CNN consists of the

following basic blocks: convolutional layers; sub-sample layers; fully connected layers.

Conclusions

Detecting transmission of covered data, hidden by one of the many existing methods of steganography in different container formats is a rather complex process. A comprehensive approach is needed to solve this problem. This approach must be aimed too analyze all possible methods of encoding, from the simplest ones to the most complex. Thus, today the urgent problem is to improve existing and create new methods of steganoanalysis, as well as the development of software based on them, with which with some probability the presence of hidden information in the container or its absence can be detected. The possibility of using the approches of convolutional neural networks to detect steganographic attachments to digital images is considered. The advantages of the proposed method of detecting attachments include sufficient accuracy and simplicity of implementation, finding attachments without the use of complex statistical algorithms. Of the shortcomings, it is worth noting the need to solve the problem of forming a representative sample of digital images used in the training phase of the neural network.

Conventional or classic image stegnography are usually designed in a heuristic way. Generally, these algorithms decide whether to conceal information into a pixel of the cover image and how to conceal 1-bit information into a pixel. So the key of the classic steganography methods is well hand-crafted algorithms, but all of these algorithms need lots of expertise and this is very difficult for us. The best solution is to mix the secret image with the cover image very well without too much expertise. Deep learning, represented by convolutional neural networks, is a good way to achieve this exactly. What we need to do is to design the structure of the encoder and the decoder as described below.

References

1. Yudin O., Barannik N., Ziubina R., Buchyk S., Frolov O., Suprun O. Efficiency Assessment of the Steganographic Coding Method with Indirect Integration of Critical Information: In.: 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT), pp. 36-40 (2019).

2. A. Westfeld and A. Pfitzmann, “Attacks on Steganographic Systems,” 3rd International Workshop on Information Hiding, Lecture Notes in Computer Science, vol. 1768, pp.61–75, Springer-Verlag, 2000.

3. A.Zh. Abdenov, LS Leonov, The use of neural networks in blind methods for detecting embedded steganographic information in digital images. Polzunovsky Vestnik. - 2010. - P. 221 - 225.