АКТУАЛЬНІ ПИТАННЯ ТЕХНІЧНИХ НАУК

UDC 004.056.55
DOI https://doi.org/10.32782/2663-5941/2025.2.2/44

Dyka A. I.
National University "Odesa Law Academy"

INTELLIGENT METHODS FOR STEGANALYSIS OF CODE-CONTROLLED COVERT MESSAGES IN WEB APPLICATIONS

In the modern digital environment, web applications occupy a leading position in communication and in ensuring the exchange of information between users. The rapid growth of multimedia content requires not only increasing its accessibility and ease of use, but also ensuring information security. One of the current problems in this context is the detection of covert data transmission channels that can be used for unauthorized exchange or concealment of confidential information. Steganographic methods that allow integrating covert messages into images, audio, or video files without changing their external characteristics pose a particular danger. To counter such threats, the development of effective steganalysis methods is becoming a key task in the field of information security of web applications. The paper focuses on the steganographic method with code control, which is characterized by high resistance to detection and poses significant complexity for modern steganalysis algorithms. Its feature is the use of balanced Walsh functions with minimal impact on the container pixels, which practically does not change the statistical properties of the image. This makes traditional approaches based on simple features or frequency characteristics ineffective. The paper proposes two approaches to building effective algorithms for detecting covert messages: (1) a method based on the analysis of the envelope of histograms of the Walsh-Hadamard transformants distribution using a multilayer fully connected neural network, and (2) a method based on an ensemble of convolutional neural networks that operates without pre-processing of histograms and can independently highlight informative features. The combination of mathematical apparatus and artificial intelligence tools provides increased classification accuracy and adaptability to various data sets. The results of experimental research have shown the high effectiveness of both methods: for the first approach, the accuracy of detecting covert messages was over 93.3%, and for the second, the accuracy reached a value of 94.3%. At the same time, the values of the Precision, Recall, and F1-score metrics are within the range of 0.95…0.98. In addition, the processing time of one image was less than 0.005 seconds, while the feature extraction time for an image of size 3872x2592 is approximately 0.45 seconds, which makes the developed methods suitable for use in systems close to a real-time mode of operation. Comparative analysis with classical statistical approaches (RS-analysis,χ^2-method, pair analysis) and the StegExpose utility confirmed the significant advantage of the proposed methods in cases of using steganography with code control, where other algorithms turned out to be ineffective or completely incapable of detecting covert data. Thus, the developed methods combine high accuracy and speed, ensuring versatility and practical suitability for integration into web applications with increased requirements for information security. They open up prospects for further improvement of steganalysis systems focused on real-time operation and form the basis for creating new protective mechanisms in the digital environment.

Key words: *steganography; steganalysis; code control; Walsh-Hadamard transform; neural networks; convolutional neural networks (CNN); artificial intelligence; web applications; covert messages; information security.*

Introduction and statement of the problem. In the modern digital environment, web applications play a central role in ensuring communication and information exchange between users. With the increase in the volume of multimedia content uploaded to platforms, the need to not only ensure its accessibility and reliability but also to guarantee security is growing. One of the key problems is the detection of possible covert information transmission channels that can be used by attackers for unauthorized data exchange or the concealment of confidential information. Control of multimedia information in this context becomes an integral part of web application security, as it allows preventing potential threats associated with embedding steganographic messages in images, audio, and video files. The development

of effective methods for analyzing and detecting covert data ensures the reliability of web platforms, increases user trust, and protects their information from possible abuse.

Steganography techniques are actively developing in web applications, and although they can provide new opportunities for data protection and confidentiality, at the same time, this creates the prerequisites for their use for malicious purposes. In particular, modern approaches open additional channels for unauthorized transmission of covert information, which complicates the detection of data leaks and increases cybersecurity threats.

In the research of Azizan [1] and co-authors, the FHWA web application is presented, which allows hiding any type of files in images using steganography. While this expands the functionality for users, eliminating the limitations of existing systems that support only certain file formats, from a cybersecurity point of view, such versatility creates new risks. FHWA supports hiding text documents, images, audio, video, and other types of files, which makes it not only a convenient tool for legal tasks but also a potential tool for cybercriminals. The additional integration of the encryption mechanism with the password enhances the security of the covert information, but at the same time makes it more difficult to detect and intercept. According to the test results, 98% of users successfully used FHWA to hide and recover files, which indicates its practical effectiveness, but also demonstrates the potential for malicious use in covert data channels.

In the research of Ediriweera, Dilhara, and Disanayake [2], a hybrid approach to data hiding in web applications is presented, combining steganography and visual cryptography. This method allows for hiding information in images in such a way that it can be restored only with the help of a special key, which increases the level of security against unauthorized access. The use of visual cryptography provides an additional level of protection, allowing the restoration of original data only by combining certain parts of the images. This approach is promising for use in systems where a high degree of confidentiality and data protection is required.

Digital steganalysis methods, which allow for the detection of covert information in multimedia objects, are also actively developing. In particular, modern approaches are often based on deep learning and complex neural networks, which provide high detection accuracy even when using complex steganographic methods. For example, the review by Kheddar et al. [3] systematizes various deep learning-based steganography methods, categorizes them by data type, describes the difficulties and prospects for development, emphasizing the high computational requirements of current models. Similarly, the work of De La Croix et al. [4] offers a comprehensive review of deep learning-based image steganography methods and highlights the main areas that require significant resources to process large data sets.

Eid et al. [5] and Farooq and Selwal [6] in their reviews emphasize current digital image analysis methodologies and open problems, among which is the need for high-performance computing resources, which limits the possibility of applying these methods in real-time on web platforms. At the same time, Weng et al. [7] propose lighter models for deep learning-based image steganography that demonstrate high efficiency with significantly lower resource costs, making them more suitable for web applications.

Research performed by Kuznetsov et al. [8], Hammad et al. [9], and Ray et al. [10] also demonstrates a wide range of methods, from the use of texture features to edge detection using deep learning models, which allow for increasing the accuracy of detection. However, all these approaches are often cumbersome and resource-intensive, which limits their use in web applications with dynamic processing of multimedia information.

Thus, although digital steganography is developing rapidly and demonstrating high results, the search for lightweight, resource-efficient solutions that combine detection accuracy and real-time performance is still relevant for integration into web applications.

However, the method with code control [11] is particularly stable, and today, there are practically no effective methods for its steganalysis. In [12], a theoretical basis for developing such steganalysis methods was created, but the particular method itself has not yet been developed. This paper proposes the development of two effective methods for steganalysis of covert messages with code control based on the application of the Walsh-Hadamard transform domain and modern artificial intelligence methods.

The *purpose* of this paper is to develop methods for steganalysis of code control embedded covert messages for web applications based on the application of the Walsh-Hadamard transform domain and artificial intelligence methods.

General principles of steganalysis of the steganographic method with code control. The steganographic method with code control is characterized by high resistance to detection and poses a significant challenge to existing steganalysis approaches. In this method, balanced Walsh functions

with a small amplitude of influence of ±1 on each pixel of the container are used to embed additional information, while the embedding of additional information occurs in an additive manner, which also adds to the method's resistance to modern steganalysis methods. Indeed, the specified embedding method practically does not affect the statistical properties of the image, which ensures minimal distortion of the medium and high resistance of the method to modern steganalysis attacks. Because of this, standard analysis methods based on simple features or frequency characteristics demonstrate low efficiency. However, in [12], a significant step was taken towards steganalysis of this method. The author proposed a rigorous mathematical apparatus that allowed to detection of characteristic changes in the histogram of the distribution of transformants of the Walsh-Hadamard transform, namely, the characteristic of bimodality. The paper [12] contains rigorous mathematical evidence that determines the occurrence of exactly this kind of distortion in steganographic messages obtained using the steganographic method with code control. The basis of these results is the two-dimensional Walsh-Hadamard transform, which for a block X is defined as

$$W_X = H'_N X H'^T_N, \quad (1)$$

where $H'_N = \frac{1}{\sqrt{N}} H_N$, X is a matrix of size $N \times N$, and the Hadamard matrix H_N of order N is given by the Sylvester construction

$$H_{2^k} = \begin{bmatrix} H_{2^{k-1}} & H_{2^{k-1}} \\ H_{2^{k-1}} & -H_{2^{k-1}} \end{bmatrix}, \; H_1 = 1. \quad (2)$$

According to the results presented in [12], if we have the given matrices W_{X_j} of the size $N \times N$ of the transformants of the Walsh-Hadamard transform of the n image blocks $X_j, j = 1,2,...,n$, each of which has the form

$$W_{X_j} = \begin{bmatrix} w_{X_j,11} & w_{X_j,12} & \cdots & w_{X_j,1N} \\ w_{X_j,21} & w_{X_j,22} & \cdots & w_{X_j,2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{X_j,N1} & w_{X_j,N2} & \cdots & w_{X_j,NN} \end{bmatrix}, \quad (3)$$

and if we make histograms for sequences of transformants $u_{kl} = \begin{bmatrix} w_{X_1,kl} & w_{X_2,kl} & \ldots & w_{X_n,kl} \end{bmatrix}$, in which additional information was embedded using the steganographic method with code control with codewords based on Walsh functions, then they will have a bimodal character with maxima at points $\pm N^2$.

Despite the theoretical results obtained, in practice, histograms of the distribution of transformants of the Walsh-Hadamard transform show significant diversity and pronounced dependence on a specific image. Such variability is due to both the peculiarities of the spatial-frequency structure of the input data and differences in the statistical characteristics of images of different classes. As a result, even in the presence of previously known informative features, effective steganalysis is significantly complicated.

In Table 1, we give examples of histograms of the distribution of transformant (5,1) sequences for several images.

Based on the data presented in Table 1, it can be seen that the specific type of histogram of the distribution of the Walsh-Hadamard transformants depends largely on the selected image and its internal characteristics. Such high variability means that classical approaches to steganalysis, based only on previously known features, often turn out to be insufficiently effective.

In this regard, the most rational seems to be the use of modern artificial intelligence methods that are able to automatically highlight stable features against the background of significant variations in the data. In this paper, we propose two approaches:

– a method based on pre-processing of input data using an envelope of the histogram, which allows you to highlight key peak-like features and reduce the influence of small fluctuations;

– a method based on an ensemble of convolutional neural networks (CNN), which directly operates with raw histogram data and can automatically detect patterns characteristic of steganographic signals, even under conditions of significant histogram variability.

This combination of classical and modern approaches provides a more reliable and versatile steganalysis system, focused on use in web applications.

Steganalysis method based on histogram envelope calculation. The histogram envelope method is based on extracting the global structure of a signal or distribution, ignoring local fluctuations. In our case, the envelope is applied to the histograms of the transformants distribution, which allows us to detect characteristic peaks associated with unimodality or bimodality of the data. Next, we apply the following definition of the envelope.

Definition 1. For a discrete signal, the envelope can be calculated as the upper envelope, which is given by

$$e[n] = \sqrt{x[n]^2 + H\{x[n]\}^2}, \quad (4)$$

where $H\{x[n]\}$ is the Hilbert transform of the signal $x[n]$, which provides a phase shift to $\pi/2$ and allows us to construct an analytical signal. This approach guarantees that the envelope "follows" the amplitude changes of the signal.

Next, the peaks of the signal envelope are calculated, as a result of which vectors containing the number of peaks and their positions are formed. The obtained data are normalized and fed to the input of the neural network. A multilayer fully connected

neural network (Multilayer Perceptron, MLP) with the specification given below is used as a classifier:

– Input layer: 3 neurons corresponding to the sizes of the feature vectors (number of peaks and positions of the first two peaks).

– Hidden layers: four layers with the number of neurons 400, 250, 100, and 50, respectively; activation function – ReLU (Rectified Linear Unit).

– Output layer: 2 neurons corresponding to unimodal and bimodal distribution classes; activation function – softmax to obtain class probabilities.

– Loss function: cross-entropy, which allows for the correct training of the classification of two classes.

– Optimizer: Adam algorithm with an initial learning rate of 0.001.

After training, the neural network predicts the probability of a new image belonging to each class. The class with the maximum probability is determined as the classification result, which allows us to conclude about the presence or absence of a steganographic signal. Such architecture ensures the ability of the model to automatically detect key patterns in histograms, even with high variability of input data.

Let us describe the proposed steganalysis method using a step-by-step presentation of three main operations: preparing input data for the method, training the neural network, and deciding whether additional information is present in a given image.

Method M(1)1. Input data preparation

1.1. The input image F is divided into 8×8-blocks X_i in a standard way.

1.2. For each block, a two-dimensional Walsh-Hadamard transform (1) is calculated.

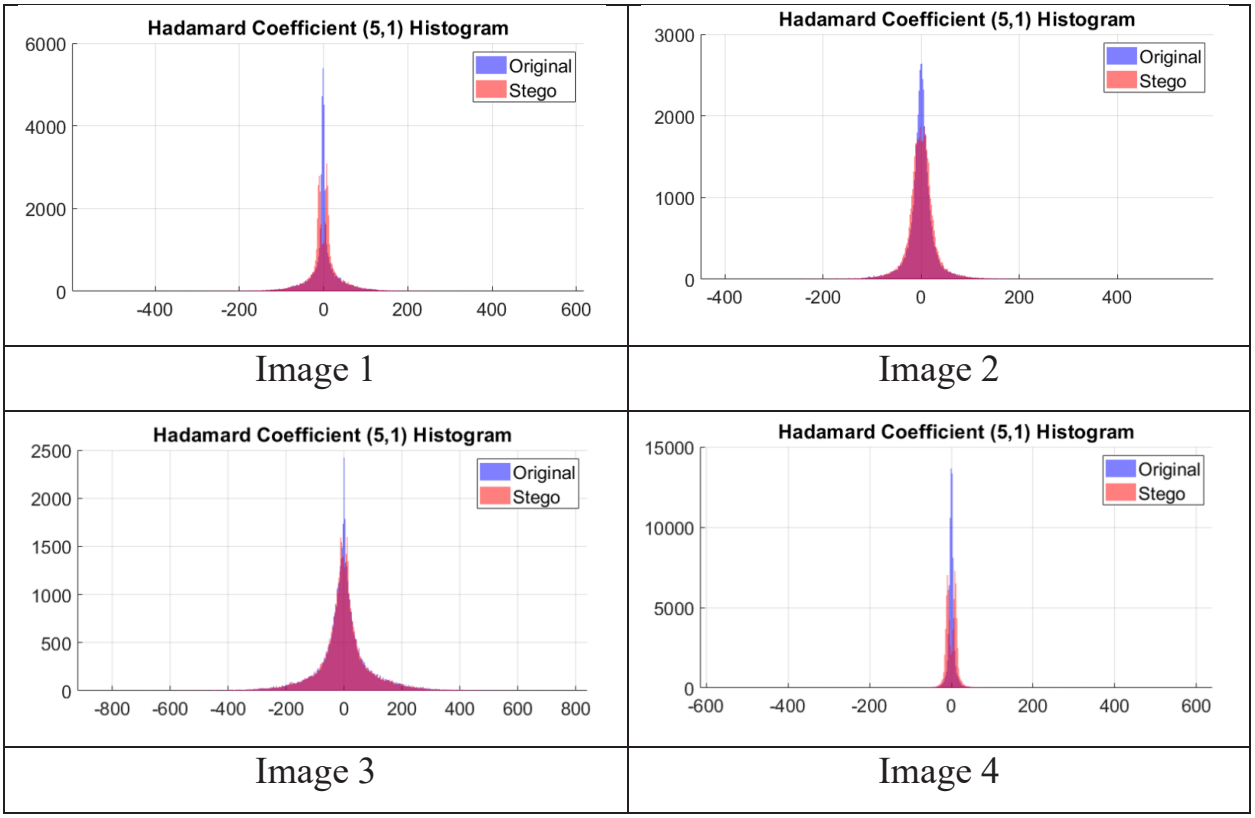
1.3. For the image, for each of the transformants of the Walsh-Hadamard transform (k,l) or for the transformant, for which there is a suspicion of the presence of additional information embedded using the steganographic method with code control, a sequence $u_{kl}=\left[w_{X_1,kl}\quad w_{X_2,kl}\quad \ldots\quad w_{X_n,kl}\right]$ is constructed.

1.4. We construct a histogram of the distribution of the sequence u_{kl}, which will be drawn from 500 bins evenly selected on the interval.

1.5. We construct an envelope (4) for the histogram constructed in step 1.4. We find a vector that will contain the following features: the number of peaks and the positions of the first two peaks of the obtained envelope.

Table 1

Histograms of the distribution of transformant (5,1) sequences of the Walsh-Hadamard transform for original images and steganographic messages



Method M(1)2. Neural network training

2.1. To train a neural network, a sample is formed, including 530 clean images and 530 images with steganographic messages. For each image, the M(1)1 method is used to calculate the envelope of the histogram and determine key features, in particular, the number of peaks and their location. The resulting feature vectors form the basis of the input data for the neural network.

2.2. Each feature vector is designated as belonging to a class: "1" – original (clean) images; "2" – images with steganographic messages.

2.3. The sample is divided into training (80%) and test (20%) parts. This division allows us to evaluate the ability of the network to generalize patterns to new data not used during training.

2.4. We train the neural network using the backpropagation method with the Adam optimizer. The input feature vectors are normalized to increase the stability of the training process. After training is complete, the network is able to predict the probability of a new image belonging to each class, which allows us to conclude whether or not a steganographic message is present.

Method M(1)3. Decision Making

3.1. Based on the image that enters the method input, a feature vector is constructed based on the method M(1)1.

3.2. The resulting feature vector is fed to the input of a trained multilayer fully connected neural network (MLP). The network processes the input data, taking into account the relationships between the features, and generates probabilities of the image belonging to the classes “original (clean) images” or “images with steganographic messages”.

3.3. Based on the output probabilities of the neural network, the class with maximum likelihood is determined. This allows an unambiguous conclusion to be drawn about the presence or absence of additional information in the image. Thus, the proposed method provides automatic and reliable classification of input images based on the statistical features of their histograms.

To visually represent the distribution of features between image classes, the t-Distributed Stochastic Neighbor Embedding (t-SNE) method was used. This method allows for to display of multidimensional feature vectors in two-dimensional space, preserving local data structures and the mutual location of points. Fig. 1 shows the results of the projection of feature vectors of unimodal and bimodal image histograms, which allows us to assess the degree of their resolution. As can be seen, the t-SNE method demonstrates that the features obtained by calculating the peaks of the envelope of the histogram provide a sufficiently informative separation of classes, which confirms the effectiveness of using a neural network for classification.

As can be seen, there is a clear grouping of both classes into separate clusters, which indicates the informativeness of the selected features and their ability to distinguish between the original images and steganographic messages. Despite the presence of zones of partial overlap, which indicates a certain similarity of features for individual cases, most points retain their belonging to their class. The spatial shape

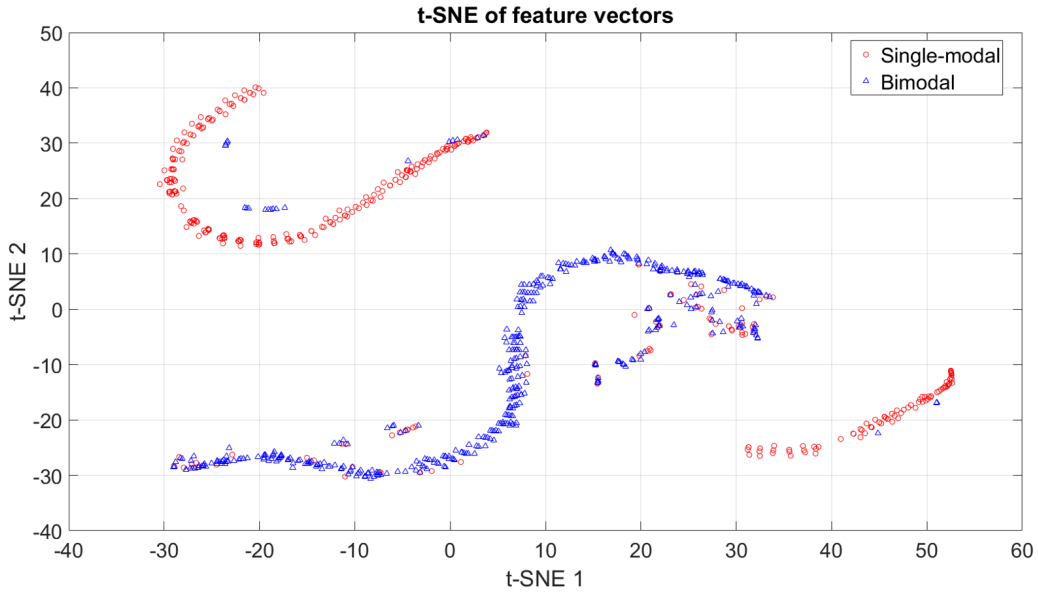


Fig. 1. t-NSE of feature vectors

of the clusters has a pronounced nonlinear structure, which confirms the feasibility of using a multilayer neural network to model complex dependencies between features and increase classification accuracy.

To quantitatively assess the effectiveness of the developed method for detecting steganographic messages, experimental testing was performed on a test data set. To display the classification results, a confusion matrix was used, which is shown in Fig. 2.

Fig. 2 allows us to analyze the relationship between the actual and predicted classes. The main diagonal of the matrix displays the number of correctly classified samples, while the off-diagonal displays the number of erroneous assignments to other classes. This approach allows us to comprehensively assess the recall, precision, and overall accuracy of the algorithm, as well as identify typical classification errors

$$\begin{aligned} Precision &= \frac{TP}{TP+FP} = \frac{192}{192+10} \simeq 0.95; \\ Recall &= \frac{TP}{TP+FN} = \frac{192}{192+4} \simeq 0.98; \\ F1\text{-}score &= 2\frac{Precision \cdot Recall}{Precision + Recall} = \frac{0.95 \cdot 0.98}{0.95+0.98} \simeq 0.96, \end{aligned} \quad (5)$$

where TP (True Positive) is the number of true positive solutions, FP (False Positive) is the number of false positives, and FN (False Negative) is the number of false negatives. This approach allows for a comprehensive assessment of recall, precision, and the harmonic mean of these indicators (F1-score), as well as identifying typical classification errors.

The results obtained demonstrate the high effectiveness of the proposed method for detecting steganographic messages. The Precision value equal to 0.95 indicates that only 5% of the samples classified by the algorithm as "steganographic messages" turned out to be false. At the same time, the Recall indicator equal to 0.98 confirms the ability of the method to correctly detect the vast majority of embedded messages, with a minimum number of missed cases. The calculated harmonic mean of these indicators (F1-score ≈ 0.96) generally characterizes the balance of the algorithm between accuracy and sensitivity, confirming its high practical applicability. The high value of the metrics in combination with the analysis of the confusion matrix indicates that the proposed approach provides reliable detection of covert data and can be recommended for use in information protection systems, where minimizing erroneous decisions is critically important. At the same time, the overall detection accuracy is over 93.3%.

Additionally, an evaluation of the algorithm's performance was conducted. On a computer with an AMD Ryzen 5 7500F processor (6 cores), 32 GB of RAM, and an NVIDIA GeForce RTX 3050 graphics adapter, the average processing time of one image of size 3872x2592 for feature extraction using the proposed method is approximately 0.45 seconds. The classification time of one image was 0.0044 seconds. This means that if there are features available, the system is able to classify more than 200 images per second, which makes it possible to use the developed method in steganographic message detection systems in a mode close to real-time.

Steganalysis method based on an ensemble of convolutional neural networks. Despite the high classification accuracy obtained using the method based on the analysis of the envelope of distribution histograms, this approach has a number of limitations.

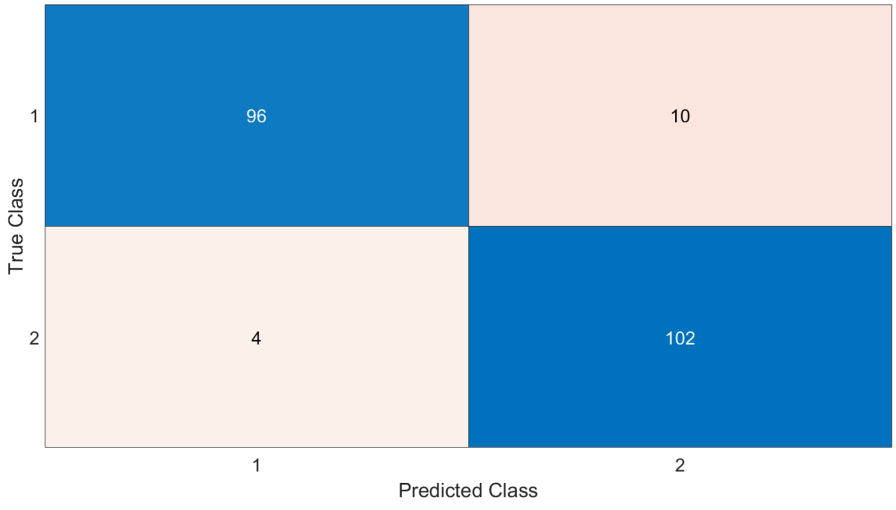


Fig. 2. Confusion matrix for classifying a sample of steganographic messages and original images based on the envelope method

First, it largely depends on the previous stage of signal processing and feature extraction. Any changes in the choice of parameters, normalization methods, or envelope estimation methods can lead to differences in classification quality, which reduces the universality of the model. Second, this approach involves the intervention of the researcher in the process of feature formation and, therefore, is not fully automated. This makes it difficult to scale the method to other types of data or tasks where the characteristic features are a priori unknown.

The development of a method based on the direct use of neural networks without prior envelope formation will eliminate the above limitations. Modern deep learning models are able to independently learn an effective representation of data and isolate significant patterns that may be unobvious or inaccessible to classical feature analysis methods. This creates the prerequisites for increasing the overall accuracy of steganographic message detection and reducing the number of false classifications.

In addition, the approach with greater use of neural networks and less preprocessing provides better generalization ability, which is especially important when operating with large and heterogeneous data sets. It also opens up the possibility of integrating unified architectures (for example, CNN ensembles or transformers), which have proven their effectiveness in computer vision and signal analysis tasks. From a practical point of view, this will allow creating a method that is less dependent on preprocessing, more adaptive to new conditions, and capable of operating in real-time.

Thus, the development of an alternative method based on the direct use of neural networks is a logical and promising direction of research, which will allow both to improve the quality of covert data detection and to ensure the versatility and scalability of steganalysis.

In this research, an ensemble of convolutional neural networks was used to detect steganographic messages, which allows for increasing the stability of classification results and reducing the impact of random errors characteristic of individual models. The ensemble approach was implemented by building three different CNN architectures with a variation in the size of convolutional kernels and the number of filters.

Each neural network has the following structure:

– The input layer receives normalized histograms with the number of $bins = 500$.

– The first convolution operation is performed by filters with sizes $[5\times1],[7\times1],[9\times1]$ for the three models, respectively, which allows analyzing patterns of different scales.

– Normalization and ReLU activation provide stabilization of learning and consideration of nonlinear dependencies.

– Subsampling layers (max pooling) reduce the dimensionality and highlight the most significant features.

– The second convolution operation with a doubled number of filters (128, 256, 384, respectively) allows for a deeper representation of the multi-level data structure.

– The Global Average Pooling layer performs compact aggregation of spatial information.

– Fully connected layers (128 neurons + Dropout 0.5) perform classification at a higher level of abstraction, which increases the generalization ability.

– The output layer implements Softmax classification into two classes ("unimodal" and "bimodal" distributions).

For training, the Adam optimization algorithm was used with an initial learning rate 10^{-4}, the number of epochs 50 and the minipack size is equal to 32. An important component is Dropout regularization, which reduces the risk of overtraining.

The proposed steganalysis method can be presented as a sequence of three key stages: input data generation, neural network training, and determination of the presence of covert information in the researched image.

Method M(2)1. Input data preparation

Perform steps 1.1 – 1.4 of Method M(1)1.

Method M(2)2. Neural network training

2.1. To train the neural network, a sample is formed, which includes 530 clean images and 530 images with steganographic messages. For each image, using the method M(2)1, histograms of the distribution of the transformants of the Walsh-Hadamard transform are calculated, which are represented as normalized vectors containing information about the 500 bins over which the transform values are distributed.

2.2. Each feature vector is designated as belonging to the class: “1” – original (clean) images; “2” – images with steganographic messages.

2.3. The sample is divided into training (80%) and test (20%) parts.

2.4. We train the neural networks using the backpropagation method with the Adam optimizer.

Method M(2)3. Decision making

3.1. Based on the input image, the method constructs a vector containing information about the distribution of values of the given transformants of the Walsh-Hadamard transform across bins. To improve the data representation, the histogram is fed as a tensor of dimension $[bins\times1\times1]$, suitable for processing by a convolutional neural network.

3.2. The input tensor is fed to an ensemble of three convolutional neural networks (CNN). Each network independently processes the data, highlighting local and global features of the histogram, and generates probabilities of the image belonging to the classes "unimodal" (pure) or "bimodal" (with steganographic message).

3.3. The final decision is made by averaging the probabilities of all three ensemble models. The class with the maximum average probability is selected as the final result. This approach allows for the reduction of the influence of errors of individual models and provides a more reliable determination of the presence or absence of covert information in the image.

Thus, the proposed method provides automatic, stable, and highly accurate classification of input images based on statistical features of their histograms using an ensemble of neural networks.

Similar to the description of the method based on envelope analysis, to visualize the classification efficiency, Fig. 3 presents a confusion matrix, which reflects the classification accuracy of the ensemble of neural networks on the test set.

Similar to (5), based on the data presented in Fig. 3, we calculate the main accuracy indicators of the method based on an ensemble of convolutional neural networks

$$\begin{aligned} Precision &= \frac{TP}{TP+FP} = \frac{200}{200+4} \simeq 0.98; \\ Recall &= \frac{TP}{TP+FN} = \frac{200}{200+8} \simeq 0.96; \\ F1-score &= 2\frac{Precision \cdot Recall}{Precision + Recall} = 2 \cdot \frac{0.98 \cdot 0.96}{0.98+0.96} \simeq 0.97. \end{aligned} \quad (6)$$

The results obtained indicate a very high accuracy of the proposed method: the Precision value reached 0.98, the Recall – 0.96, and the F1-score is approximately 0.97. It is worth noting that these indicators even exceed the results of the method based on envelopes, which confirms the high efficiency of the chosen approach even without additional data processing. Such high indicators show the reliability and stability of the proposed method for the task. At the same time, the overall detection accuracy is over 94.3%.

It should be noted that both the time for classifying images by a neural network and the time for extracting features for the method based on an ensemble of convolutional neural networks are the same as for the method based on the envelope calculation.

Comparison of the proposed methods with existing analogues. To ensure a correct assessment of the effectiveness of the proposed approaches to steganography, a comparison was made with a number of well-known methods that are widely used in the practice of researching covert messages. These include classical statistical approaches – RS-analysis, χ^2-method, and pair analysis, which provide a basic check of the probability of the presence of covert data in digital containers. Additionally, the modern method of steganalysis of code control based on perturbations of Walsh-Hadamard transformants, which is characterized by high sensitivity to structural changes, was considered. For a comprehensive comparison, the integrated utility StegExpose was also used, which combines several detection algorithms and is a practical tool for real analysis scenarios. Such a multi-vector comparison allows a comprehensive assessment of the advantages and limitations of the developed methods in relation to established analogues.

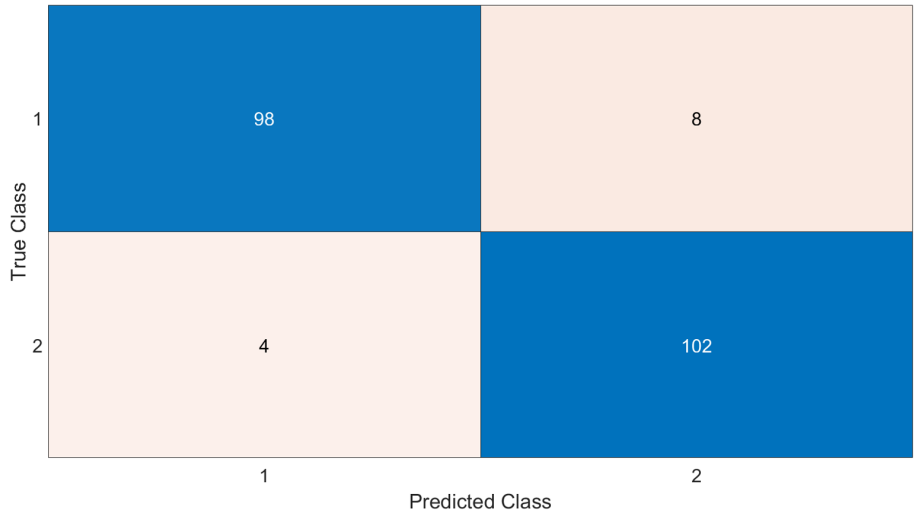


Fig. 3. Confusion matrix for classifying a sample of steganographic messages and original images based on the method based on an ensemble of convolutional neural networks

In Table. 2, we present the results of the experiments.

Table 2

Results of the comparative analysis

Method	Accuracy
Steganalysis method based on envelope calculation	93.3%
Steganalysis method based on an ensemble of convolutional neural networks	94.3%
Method based on the analysis of transformants of the Walsh-Hadamard transform [13]	87.7%
RS-analysis [14]	When performing RS-analysis on a sample of 530 images with full (100%) pixel coverage, it was found: -at a threshold of 5% (value >0.05), signs of embedded information were recorded in 50 cases, which is 9.43% of the total number of images; – at a threshold of 10% (value >0.10) confident signs were detected in 18 cases, which is 3.4%.
Pair analysis [14]	Does not detect embedding using code control
χ^2 analysis [15, 16]	When performing an χ^2-analysis on a sample of 530 images with full (100%) pixel coverage, it was found that signs of embedded information were recorded in 1 case, which is 0.18% of the total number of images;
StegExpose Utility [17]	Does not detect embedding using code control

Conclusions. The paper considers the problem of detecting covert messages formed using the steganographic method with code control, which is characterized by high resistance to existing steganalysis approaches. Two effective methods based on a combination of spectral transformations and modern artificial intelligence technologies are proposed:

– the first method is based on the analysis of the envelope of histograms of the Walsh-Hadamard transformants and subsequent classification using a multilayer neural network;

– the second method implements a direct analysis of the transform distributions using an ensemble of convolutional neural networks, which allows automatically isolating characteristic patterns without prior data processing.

The experiments confirmed the high efficiency of both approaches. For the method based on the envelope of histograms, the accuracy of detecting covert messages was 93.3%, while for the method based on the ensemble of convolutional neural networks, 94.3%. At the same time, the Precision, Recall, and F1-score indicators are at the level of 0.95…0.98, which indicates a balance of methods between sensitivity and specificity. At the same time, the processing time of one image does not exceed 0.005 seconds, which provides the possibility of using the developed methods in web applications in a mode close to real-time.

Comparison with classical steganalysis methods (RS-analysis, χ^2-analysis, pair analysis, StegExpose utility) showed that they are practically unable to detect covert information embedded by the code-controlled method. This confirms the significant advantage of the developed methods, which demonstrate high accuracy even in cases where other approaches are ineffective.

Thus, the developed methods are promising for practical integration into information security systems for web applications. They combine the mathematical rigor of spectral transformations with the adaptability of neural networks, which makes them a universal tool for detecting covert data transmission channels. Further research may be aimed at optimizing computational complexity, adapting to other multimedia data formats, and creating comprehensive protection systems capable of operating under real cyberthreats.

Bibliography:

1. Azizan N. et al. File hiding web application (fhwa) using image steganography. European Proceedings of Multidisciplinary Sciences. 2022. P. 599-609. doi: 10.15405/epms.2022.10.56

2. Ediriweera S., Dilhara B. A. S., Disanayake C. Web-Based Data Hiding: A Hybrid Approach Using Steganography and Visual Cryptography. International Research Conference on Smart Computing and Systems Engineering (SCSE). IEEE, 2023. Vol. 6. P. 1-7. doi: 10.1109/scse59836.2023.10214994

3. Kheddar H. et al. Deep learning for steganalysis of diverse data types: A review of methods, taxonomy, challenges and future directions. Neurocomputing. 2024. Vol. 581. P. 127528. doi: 10.1016/j.neucom.2024.127528

4. De La Croix N. J., Ahmad T., Han F. Comprehensive survey on image steganalysis using deep learning. Array. 2024. Vol. 22. P. 100353. doi: 10.1016/j.array.2024.100353

5. Eid W. M. et al. Digital image steganalysis: current methodologies and future challenges. IEEE Access. 2022. Vol. 10. P. 92321-92336. doi: 10.1109/access.2022.3202905
6. Farooq N., Selwal A. Image steganalysis using deep learning: a systematic review and open research challenges. Journal of Ambient Intelligence and Humanized Computing. 2023. Vol. 14, No. 6. P. 7761-7793. doi: 10.1007/s12652-023-04591-z
7. Weng S. et al. Lightweight and effective deep image steganalysis network. IEEE Signal Processing Letters. 2022. Vol. 29. P. 1888-1892. doi: 10.1109/lsp.2022.3201727
8. Kuznetsov A. et al. Image steganalysis using deep learning models. Multimedia Tools and Applications. 2024. Vol. 83, No. 16. P. 48607-48630. doi: 10.1007/s11042-023-17591-0
9. Hammad B. T., Ahmed I. T., Jamil N. A steganalysis classification algorithm based on distinctive texture features. Symmetry. 2022. Vol. 14, No. 2. P. 236. doi: 10.3390/sym14020236
10. Ray B. et al. Image steganography using deep learning based edge detection. Multimedia Tools and Applications. 2021. Vol. 80. No. 24. P. 33475-33503.
11. Kobozeva A.A., Sokolov A.V. Robust Steganographic Method with Code-Controlled Information Embedding. Problemele energeticii regionale. 2021. No. 4 (52). P. 115-130. doi: 10.52254/1857-0070.2021.4-52.11
12. Dyka A.I. Detection of covert channels in web applications based on unimodality violation in the Walsh–Hadamard spectrum. Informatics and Mathematical Methods in Simulation. 2025. Vol.15, No. 1, P. 5-14. doi: 10.15276/imms.v15.no1.5
13. Lanovska O.O., Sokolov A.V. Steganalysis of a method with code-controlled information embedding in the Walsh-Hadamard transform domain. Informatics and mathematical methods in simulation. 2024. V.1. No 4. P. 1-12
14. Ker A. D. Quantitative evaluation of pairs and RS steganalysis. Security, Steganography, and Watermarking of Multimedia Contents VI. SPIE, 2004. Vol. 5306. P. 83-97. doi: 10.1117/12.526720
15. Westfeld A. Pfitzmann A. Attacks on Steganographic Systems. International Workshop on Information Hiding. 1999. Vol. 1768. P. 61-76. doi: 10.1007/10719724_5
16. Pan I. H., Liu K. C., Liu C. L. Chi-square detection for PVD steganography. International Symposium on Computer, Consumer and Control. IEEE, 2020. P. 30-33. doi: 10.1109/is3c50286.2020.00015
17. StegExpose. [Electronic resource]. 2015. URL: https://github.com/b3dk7/StegExpose.

Дика А. І. ІНТЕЛЕКТУАЛЬНІ МЕТОДИ СТЕГАНОАНАЛІЗУ ПРИХОВАНИХ ПОВІДОМЛЕНЬ ІЗ КОДОВИМ УПРАВЛІННЯМ У ВЕБЗАСТОСУНКАХ

У сучасному цифровому середовищі вебзастосунки посідають провідне місце у комунікації та забезпеченні обміну інформацією між користувачами. Стрімке зростання обсягів мультимедійного контенту зумовлює потребу не лише у підвищенні доступності та зручності його використання, а й у гарантії інформаційної безпеки. Однією з актуальних проблем у цьому контексті є виявлення прихованих каналів передачі даних, що можуть бути використані для несанкціонованого обміну або приховування конфіденційної інформації. Особливу небезпеку становлять стеганографічні методи, які дають змогу інтегрувати приховані повідомлення у зображення, аудіо– чи відеофайли, не змінюючи їх зовнішніх характеристик. Для протидії таким загрозам розробка ефективних методів стеганоаналізу стає ключовим завданням у сфері інформаційної безпеки вебзастосунків. У роботі акцентовано увагу на стеганографічному методі з кодовим управлінням, який вирізняється високою стійкістю до виявлення та становить значну складність для сучасних алгоритмів стеганоаналізу. Його особливістю є використання збалансованих функцій Уолша з мінімальним впливом на пікселі контейнера, що практично не змінює статистичних властивостей зображення. Це робить традиційні підходи, засновані на простих ознаках або частотних характеристиках, малоефективними. У статті запропоновано два підходи до побудови ефективних алгоритмів виявлення прихованих повідомлень: (1) метод на основі аналізу огинаючих гістограм розподілу трансформант перетворення Уолша-Адамара із застосуванням багатошарової повнозв'язної нейронної мережі, та (2) метод на базі ансамблю згорткових нейронних мереж, що працює без попередньої обробки гістограм і здатний самостійно виділяти інформативні ознаки. Поєднання математичного апарату та інструментів штучного інтелекту забезпечує підвищену точність класифікації й адаптивність до різноманітних наборів даних. Результати експериментальних досліджень засвідчили високу ефективність обох методів: для першого підходу точність виявлення прихованих повідомлень склала понад 93.3%, а для другого – 94.3%. При цьому значення метрик Precision, Recall та F1-score перебувають у межах 0.95…0.98. Крім того, час обробки одного зображення становив менше 0,005 секунди, тоді як час виділення ознак для зображення розміру 3872x2592 становить приблизно 0.45 секунди, що робить розроблені алгоритми придатними для застосування у системах, наближених до режиму реального часу. Порівняльний аналіз із класичними статистичними підходами (RS-аналіз, χ^2-метод, аналіз пар) та утилітою StegExpose підтвердив істотну перевагу запропонованих методів у випадках використання стеганографії з кодовим управлінням, де інші алгоритми виявилися малоефективними або зовсім не здатними зафіксувати приховані дані. Таким чином, розроблені методи поєднують високу точність і швидкодію, забезпечуючи універсальність та практичну придатність для інтеграції у вебзастосунки з підвищеними вимогами до інформаційної безпеки. Вони відкривають перспективи подальшого вдосконалення систем стеганоаналізу, орієнтованих на роботу у режимі реального часу, та формують базу для створення нових захисних механізмів у цифровому середовищі.

Ключові слова: *стеганографія; стеганоаналіз; кодове управління; перетворення Уолша-Адамара; нейронні мережі; згорткові нейронні мережі (CNN); штучний інтелект; вебзастосунки; приховані повідомлення; інформаційна безпека.*