IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 24, 2019, accepted November 17, 2019, date of publication November 25, 2019,
date of current version December 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2955452

Coverless Image Steganography: A Survey

JIAOHUA QIN[1], YUANJING LUO[1], XUYU XIANG[1], YUN TAN[1], AND HUAJUN HUANG[2]
[1]College of Computer Science and Information Technology, Central South University of Forestry and Technology, Changsha 410004, China
[2]School of Information Technology and Management, Hunan University of Finance and Economics, Changsha 410001, China

Corresponding author: Huajun Huang (hhj0906@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61772561, in part by the Key Research and Development Plan of Hunan Province under Grant 2018NK2012, in part by the Science Research Projects of Hunan Provincial Education Department under Grant 18A174, in part by the Degree and Postgraduate Education Reform Project of Hunan Province under Grant 209, and in part by the Postgraduate Education and Teaching Reform Project of Central South Forestry University under Grant 2019JG013.

ABSTRACT With the digitalization of information, a lot of multimedia data are under attack, information security has become a key issue of public concern. Image steganography, aiming at using cover images to convey secret information has become one of the most challenge and important subjects in the field of information security recently. Different from the traditional image steganography, coverless image steganography does not need to employ the designated cover image for embedding the secret data but directly transfers secret information through its own properties such as pixel brightness value, color, texture, edge, contour and high-level semantics. Therefore, it radically resist the detection of steganalysis tools and significantly improves the security of the image. Its basic idea is to analyze the attributes of the image and map them to the secret information according to certain rules based on the characteristics of the attributes. This paper includes more than 50 key contributions to provide a comprehensive survey in this field, covers the main aspects of coverless image steganography research: the fundamental frameworks, pre-processing, feature extraction, generation of hash sequence and mapping relationships. The existing methods are evaluated and the prospect of future work is also summarized.

INDEX TERMS Steganography, coverless image steganography, information hiding, information security.

I. INTRODUCTION

Due to the multimedia data may contain private, valuable even confidential information, the popularization of personal computers and the proliferation of multimedia data on the internet provided convenient conditions for the disclosure of personal privacy. In addition, the published information also faces some potential threats such as illegal tampering, copy and distribute. In order to achieve hidden communication and copyright protection, image steganography has become an imminent problem.

The traditional image steganography designates a cover image and embeds the secret information into the carrier data(digital image) with the slight modification [1]. The common steganography methods are divided into two types: spatial domain-based and transform domain-based methods [2]. The spatial domain-based method is more widely used and influential than the transform domain-based methods, and they are the adaptive LSB hiding method [3], [4], the spatial adaptive steganography algorithm HUGO [5], WOW [6], S-UNIWARD [7], HILL-CMD [8] and so on. The transform domain method is improved on the basis of the former to enhance the robustness against attacks, it includes quantization table (QT) [9], the hidden method in DWT (discrete wavelet transform) domain [10], DFT (discrete Fourier transform) domain [11], DCT (discrete cosine transform) domain [12] and IWT (integer wavelet transform) [13]. Embedding secret information in cover images requires intensive computations, designing steganography on hardware improves the speed and makes steganography widely used [14]. However, these modification caused by the embedding will be left in the cover image, which will make the detection technology for hidden information successful, this detection technology is also called steganalysis. The basic approach to steganalysis is based on feature extractors [15], such as SPAM [16], SRM [17], etc. The traditional machine learning classifiers are also important, such as SVM, decision trees, ensembles etc. With the optimization of deep neural networks recently, steganalysis methods based on neural networks [15] have been rapidly

The associate editor coordinating the review of this manuscript and approving it for publication was Chaker Larabi.

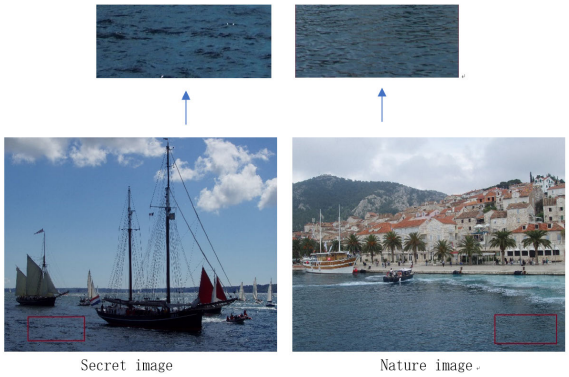


FIGURE 1. A natural image shares one similar patch with the secret image, information hidden in secret images can be transmitted through natural images.

developed. For example, using the deep convolutional neural network (CNN) [18] for steganalysis becomes very popular, and the experimental results show that CNN can significantly improve the classification accuracy by replacing usual classifiers.

As an emerging and challenging problem in information security, image steganography has recently become an active research field. In order to radically resist the detection of steganalysis and improve the robustness of steganography, Bilal et al. proposed ''Zero-steganography'' in 2013 [19]. In order to improve the security, Zhou *et al.* proposed the new concept of ''coverless'' in May 2014 [20]. Compared with the traditional image steganography, ''coverless'' still needs carriers. It emphasizes that it does not need other carriers but directly uses secret information as the driving force to ''generate/acquire'' cryptographic carriers. Zhou *et al.* have done some research on the algorithm of this idea. In 2015, they proposed image-based and text-based coverless information hiding, which opened the curtain of coverless information hiding technology in China. The image itself already contains a lot of feature information, such as pixel brightness value, color, texture, edge, contour and high-level semantics. It is possible to generate some relationships between these feature information and secret information for information hidden with a proper feature description. It can generate the same information data as the secret image by sending the generated image which is independent of the secret information. As shown in Fig.1, the transmitted image is only a nature image instead of the image embedded with any information of the secret image, and it has the same effect as transferring the secret image [21]. The method first establishes the mapping relationship between the cover image and the secret image, generates the current position label sequence according to the initial position label sequence, and finally retrieves the corresponding image in the image set according to the position label and mapping relationship. This method can effectively resist the attack of steganalysis tools, and significantly improves the security of the secret information.

The rest of this paper is organized as the following sections: Section II presents the related works including the traditional image steganography, steganalysis, coverless image steganography, the progress in the past five years and key challenges. We present some fundamental frameworks of the coverless image steganography in Section III. Some important sub-problems such as pre-processing, feature extraction, generation of hash sequence and mapping relationships will be described in Section IV. A summarization of performance of existing coverless steganography is given in Section V. Finally, the paper is discussed with several directions worthy of further development in Section VI.

II. BACKGROUND

A. TRADITIONAL IMAGE STEGANOGRAPHY

The most popular and easy-to-implement algorithm of information hiding is the Least Significan Bit (LSB) algorithm [15]. There are other algorithms for information hiding derived from the critical ideas of LSB algorithm: HUGO [5], WOW [6], UNIWARD [22], and others. Then many transform domain steganography methods have been proposed, such as the hidden method in DWT (discrete wavelet transform) [10], DFT (discrete Fourier transform) [11], DCT (discrete cosine transform) domain [12] and IWT (integer wavelet transform) [13]. As shown in Table.1, we briefly introduced the central idea of the above core steganography algorithm. Based on them, many steganography algorithms with higher security have been derived in recent years, the latest IWT steganography based on 3D sine chaotic map which has high security and acceptable robustness [23]. The combination with hardware, especially FPGA(field programmable gate array), can further improve the security [24]. Nevertheless, it is not difficult to find out that traditional image steganography modifies the content of the image to some extent so that it is hard to resist the detection of image steganalysis tools.

B. STEGANALYSIS

Steganalysis includes steganography detection and secret information extraction. Different from steganography, it reveals the existence of secret information in the image and tries to point out the secret information in the image [25]. The steganalysis methods are mainly aim at the LSB tools in spatial domain [3], [4] and the tools in DCT transform domain [12]. The method is based on the statistical anomaly of the carrier data caused by information embedding to determine whether the secret information exists. The steganalysis tools mainly analyze the influence of secret information embedding on the statistical characteristics of the image, such as SPAM [16], SRM [17], and its several variants [26], [27]. With the popularization of deep neural networks recently, new steganalysis methods based on neural networks arouse strong public concern [15]. The first attempt to use CNNs for hidden analysis was made by Qian *et al.* who proposed a Gaussian Neural Convolutional Neural Network (GNCNN) for image steganography in the spatial domain [28]. By using Gaussian function instead of ReLU or sigmoid in traditional

TABLE 1. Summarization of traditional image steganography methods.

Categories	Name	Ref.	Years	Content
Spatial Domain (Replaces secret information with the remainder of the carrier)	LSB	[3-4]	1993	It hides secret information in the lowest effective bit of the image with poor robustness, large hiding capacity, and simple algorithm.
	HUGO	[5]	2010	It protects the higher-order statistics of four adjacent pixels, HUGO allows the embedder to hide 7 × longer message with the same level of security level in comparison with LSB.
	WOW	[6]	2012	The directional filtering in wavelet analysis domain is used to define the embedded distortion of the spatial image.
	S-UNIWARD	[7]	2013	The universal wavelet relative distortion applied to the spatial domain with the best performance against the general steganalysis.
Transform Domain (Embedding secret information into a transformation space of the carrier)	DWT(discrete wavelet transform)	[10]	2007	It embeds secret information into the low-frequency coefficient of DWT domain and decomposes the image in multi-space, multi-scale and multi-frequency.
	DFT(discrete Fourier transform)	[11]	2007	It is resistant to geometric attacks such as translation, rotation, scaling to some extent. large hiding capacity, and simple algorithm.
	DCT(discrete cosine transform)	[12]	2010	It hides secret information in the non-zero coefficient of the image DCT domain with weak resistance to accumulation and change, good ompression effect, fast algorithm, and it is combined with JPEG image frequently.
	IWT(integer wavelet transform)	[13]	2017	Unlike DWT, it uses the integer coefficients for wavelet transformation, error percentage is very low in the extraction process in the integer wavelet.

CNNs as the activation function, GNCNN [29] achieves the same performance as SRM [17]. Xu *et al.* studied the CNN structure design for the application of image steganalysis, and the results showed that well-designed CNN has the potential to provide better detection performance in steganography analysis. Ye *et al.* developed a specific steganalysis application to supervise CNN model [30] which has achieved better experimental results than SRM. In addition, Chen *et al.* [31] proposed a phaseaware CNN for JPEG images steganalysis through designing phase separation model to achieve high-precision steganalysis detection. These methods are used to test the anti- steganalysis ability of coverless steganography will be described in Section V.

C. COVERLESS IMAGE STEGANOGRAPHY

Coverless image steganography, which has been proposed in recent years has great potential for development with not readable and invisible [32]. Coverless image steganography does not mean that the carrier is not used, but the carrier is not modified. Its own properties such as pixel brightness value, color, texture, edge, contour and high-level semantics are used to represent the secret information. It bypasses the process of forming the camouflage carrier in the traditional steganography method and directly passes the carrier (the secret information). The basic idea of coverless image steganography is to analyze the attributes of the carrier and map them to the secret information according to certain rules based on the characteristics of the attributes. In this way, the carrier can directly represent the secret information. In summary, the major contributions of coverless image steganography are as below:

- The covert communication can be realized without modifying the stego-image.
- Because the stego-image has not been modified, the existing steganalysis tools can not detect secret information.

D. PROGRESS IN THE PAST FIVE YEARS

Since the concept of ''coverless'' was proposed in 2014, coverless image steganography has developed rapidly. As shown in Fig.2, Zhou *et al.* proposed a new image steganography framework called coverless image steganography based on the image gray value in 2015 [33]. In this framework, they select a series of appropriate images that already contain secret data directly from a constructed database. In 2016, Zhou *et al.* proposed a coverless information hiding algorithm based on the BOW (Bag-of-words) model; this method extracts the visual keywords in the model to express the secret information [20]. Volkhonskiy *et al.* proposed a new model for generating image-like containers based on Deep Convolutional Generative Adversarial Networks (DCGAN [34]) called Steganographic Generative Adversarial Networks (SGAN), opened a new field for applications of GAN. This approach allows to generate more steganalysis message embedding and improve the security of the image using standard steganography algorithms in 2017 [15]. Two months later, Zhou et al. proposed a new steganography approach based on histograms of oriented gradients (HOGs)-based hashing algorithm [35]. In this method, the original images whose hash sequences equal to the secret information are directly selected from a large-scale database, which can be used as the stego-images for secret communication. In July of

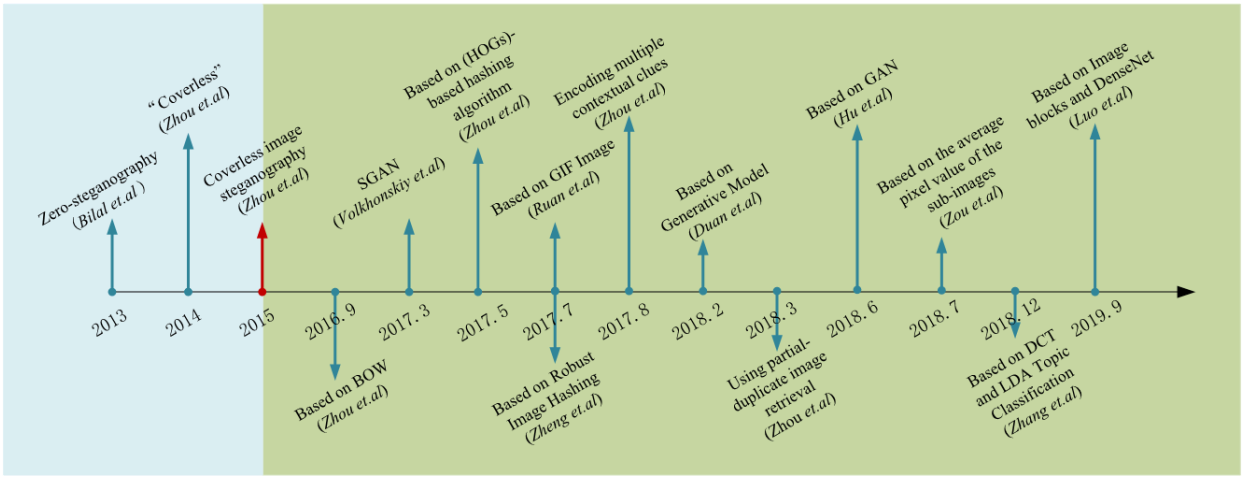


FIGURE 2. The summary of recent developments in coverless image steganography.

the same year, Ruan and Qin proposed a coverless covert communication method based on GIF (Graphics Interchange Format) image with a larger embedding rate due to it proposed a way to use multidimensional extension attributes of the image [36]. At the same time, Zheng *et al.* proposed a new coverless steganography method based on robust image hashing with higher capacity, robustness, and security than the method proposed in [33], which is also based on image hash [37]. In order to improve the discriminability of the BOW (Bag-of-words) model combined with contextual clues, Zhou *et al.* proposed a multiple contextual clue encoding approach for partial-duplicate image retrieval in August [38]. A new coverless information hiding method based on the generative model was proposed by Duan and Song in 2018, this is the first time to propose information hiding method based on the generative model that can effectively resist steganalysis tools [39]. In June, Hu *et al.* proposed a novel image steganography method via deep convolutional generative adversarial networks which is the first approach in this direction. The secret information is mapped into a noise vector and the trained generator neural network model is used to generate the carrier image based on the noise vector [40]. Zhou *et al.* proposed a novel coverless steganographic approach without any modification for transmitting the secret image by a set of partial duplicates of the given secret image as stego-images [20], which are retrieved from a natural image database in the meantime [21]. Zou *et al.* proposed a novel coverless information hiding method based on the average pixel values of sub-images to improve the information hiding capacity by generating hash sequences and realizing the secret information hiding [41]. A novel coverless image steganography algorithm based on discrete cosine transform and latent Dirichlet allocation (LDA) topic classification was proposed by Zhang *et al.* to improve the robustness and capability of resisting image steganalysis, which has great potential application in secure communication of big data environment [42]. On the basis of Zhou *et al.* [21], Luo *et al.* proposed a coverless real-time image information hiding based on dense convolutional network in September 2019 [43]. The capacity, accuracy and robustness have been improved in the this method, which proves that convolutional neural network can be well applied to coverless steganography.

E. KEY CHALLENGES

Coverless image steganography aims at secret information hiding which needs to be noted in some aspects: Firstly, the feature used cannot be single. Otherwise, the capacity and efficiency of data transmission will be insufficient. Secondly, to send messages quickly and accurately, the sender usually need to prepare image data set formed by a large number of nature images in advance, and these images have a wide range of sources and cannot correctly meet the ideal situation. Thirdly, the method must be able to resist steganalysis tools and attackers. To sum up, coverless image steganography faces three challenges: high capacity, high accuracy, and security, as illustrated in the following sections.

1) HIGH CAPACITY

Although the image itself already contains a lot of feature information, such as pixel brightness value, color, texture, edge, contour and high-level semantics. The standard method of information hiding is constructing a robust image hash with SIFT [44] to protect the secret data. Since the hash sequence is generated based on the mapping relationships between secret information and nature images, the capacity of coverless steganography is limited by the length of the image hash. This phenomenon is inevitable so that the size of coverless steganography is distinctly less than that of traditional image steganography. The capacity is usually measured by bits per pixel which means the average number of bits concealed into each pixel of the cover image.

2) HIGH ACCURACY

High accuracy requires that the whole image steganography process can transfer complete and correct information smoothly. Information hidden in secret images can

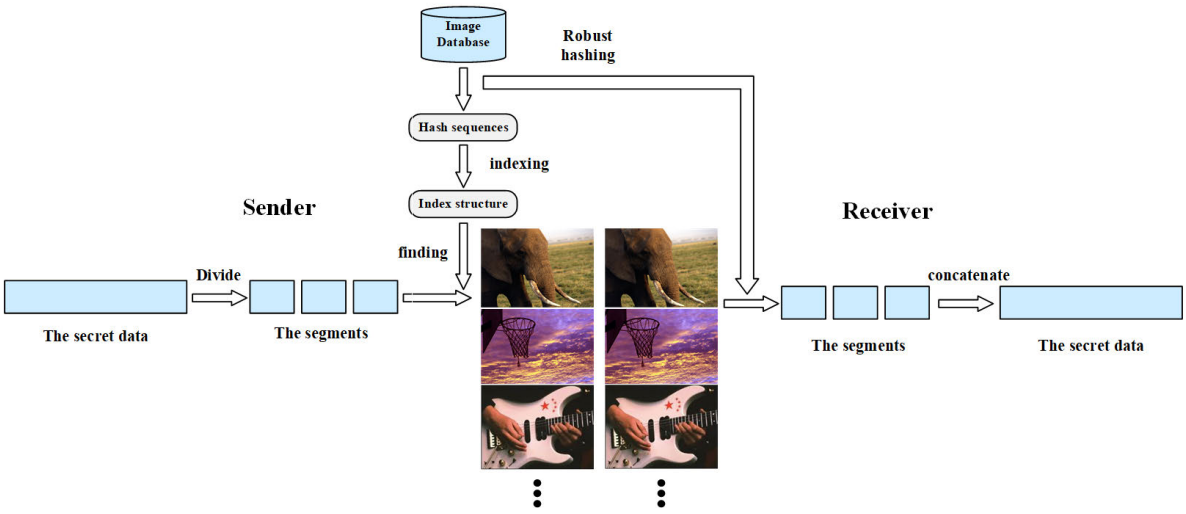


FIGURE 3. The flow chart of the steganography framework based on the gray image.

be transmitted through natural images by coverless image steganography. To send messages accurately, it is usually necessary to prepare an image data set composed of a large number of natural images, which are from a wide range of sources and cannot accurately meet the needs. Deviation of natural image selection or imperfect establishment of the inverted index structure will both lead to incomplete or wrong information transmission. At the same time, if the image is attacked in the process of transmission, the accuracy will be reduced.

3) SECURITY

The security of steganography includes two aspects: resistance to steganalysis tools and the security to attackers. As everyone knows, most of the different steganalysis tools are working on the modification traces and the ideal image steganography method has complete resistance to all types of steganalysis tools. Meanwhile, since coverless image steganography directly uses secret information as the driving force to ''generate/acquire'' cryptographic carriers. The most methods is to establish the mapping relationship between the cover image and the secret image, generate the current position label sequence according to the initial position label sequence, and finally retrieve the corresponding image in the image set according to the position label and mapping relationship. If the attackers discover these mapping and image data set, they can quickly get the information we want to hide, and the secret information will be easily destroyed, even wrong information will be transmitted. Therefore, the resistance to the attackers is also important. These evaluation indexes are essential to evaluate the existing coverless image steganography and will be used in Section V.

III. FUNDAMENTAL FRAMEWORK

There has been steady progress in coverless image steganography, as evidenced by Fig.2, The ''coverless'' strategy will be the mainstream. Therefore, the design of efficient fundamental framework plays a key role. At the same time, the feature extraction algorithm and the mapping relationship are also very important, which will be explained in detail in section 4. In this section, we describe the fundamental milestone framework of coverless image steganography since ''coverless'' entered the field, as listed in Fig.2 and summarized in Table.2. Nearly all frameworks proposed are based on Zhou'method and attempt to improve it in one or more areas.

A. COVERLESS IMAGE STEGANOGRAPHY BASED ON THE GRAY IMAGE

As we know, any original image contains a lot of information. In this framework, the image database is first constructed by collecting a large number of images from the network. And then, for each image in the database, its hash sequence is generated by a robust hashing algorithm. Afterward, all of these images are indexed according to their hash sequences to build an inverted index structure. To communicate the secret information, the sender first transforms the secret information to a bit string and divides it into several equal length segments. Afterward, a series of images associated with the coverless image steganography through index structure, which can be regarded as stego-images, are acquired and then transmitted to the receiver. On the receiver, the hash sequence of these received images is generated by the same hash algorithm. Because these hash sequences are the same as the secret information segments, the receiver can concatenate them to recover the secret information [33]. The flow chart of the steganography framework can be found in Fig.3. Experimental results show this method has high safety. According to the above content, the main components of this framework mainly consists of the following four parts:

1) **Hash sequence generation by a robust hashing algorithm**. Firstly, the complete image would be transformed to the gray-level image and divide the image

TABLE 2. The main coverless image steganography.

Method.	Name	Ref.	Years	Key points
1	coverless image steganography based on the image gray value	[33]	2015	An image can represent 8 bits of information through hash sequence generation by a robust hashing algorithm, and an inverted index structure is built for all the hash sequences namely lookup table.
2	coverless information hiding based on the BOW (Bag-of-words) model	[20]	2016	The relational library is established for visual words and text keywords, and visible words of images can be used to represent text information.
3	coverless covert communication method based on GIF	[36]	2017	From Method. 1, this method proposes a way to use multidimensional extension attributes of the image, which quantifies each GIF image in the existing carrier image data set and extracts the attribute value of its extension with a larger embedding rate.
4	coverless Image Steganography Based on HOGs-Based Hashing Algorithm	[35]	2017	This method divides the image into several non-overlapping blocks and extracts the hash sequence of each image block through the HOGs histogram hash algorithm.
5	coverless steganography based on robust image hashing	[37]	2017	By improving the method of 1, this method can extract 18-bits binary sequences from each image as the strong hash value. A modified inverted index of quadtree structure is designed.
6	coverless steganography based on generative model	[39]	2018	This is the first time to propose the coverless image information hiding method based on a generative model. The secret image can be achieved by transmitting the meaning-normal image which is not
7	coverless steganography based on generative adversarial networks	[40]	2018	The secret information is mapped into a noise vector and the trained generator neural network model is used to generate the carrier image based on the noise vector.
8	coverless steganographic based on partial duplicates of a given secret image as stego-images	[21]	2018	Steganography uses a set of appropriate partial duplicates of a given secret image as stego-images, which are retrieved from a natural image database.
9	coverless information hiding based on the average pixel values of sub-images	[41]	2018	The secret information is segmented in accordance with the structure of a Chinese sentence, the position of each segment could be got according to the dictionary. Then the label information of the hash sequence in each part of the hash array is obtained to transmit information.
10	coverless image steganography based on discrete cosine transform and latent Dirichlet allocation (LDA) topic classification	[42]	2018	The first step is used LDA topic model to handle the image set and choose an image of a certain topic. Then, the feature sequences are generated through DC coefficients. Finally, the image corresponding to the information segment is transmitted to the receiving end.
11	coverless image information hiding based on image block matching and dense convolutional network	[43]	2019	This method transfers a set of stego-images which share one or several visually similar blocks with the given secret image. The supervised learning of deep learning and the DCT coefficient improve the retrieval accuracy and robustness.

into 3×3 secret image patches and the average gray value of each region is calculated. The second step is to concatenate the values in a zigzag order to a vector. The difference in gray value between the former region and the latter region (if the gray value of the former region is greater than that of the latter region, the value is 1, otherwise is 0) is the information so that an image can represent 8-bits of information.

2) **Establishment of the inverted index structure.** To speed up the search, we first index all the images from the database according to their hash sequence. Then, we build an inverted index structure, namely lookup table, for all the hash sequences. The lookup table contains all possible 8-bits hash sequences as entries. From Fig.4, it is simple to find that each entry points to a list storing of all the IDs that the images whose hash sequence is the same as the entry.
3) **Finding appropriate images by searching for the segments in the inverted index structure**. The sender first transforms the secret data needed to be sent to a bit string and then divides it into plenty of segments with the equal length, namely 8-bits. Each segment is then used as a query to obtain all images with the same hash sequence as that segment. Only one image is randomly selected from the image for transmission. Finally, a series of images are obtained by searching the inverse index structure.
4) **Transmission of the secret information**. For the sender, those images are transmitted to the receiver one by one in order. For the receiver, all of the images are also received in order. Depending on the order of the received images, the receiver concatenates all the hash sequences of the images to restore the secret information.

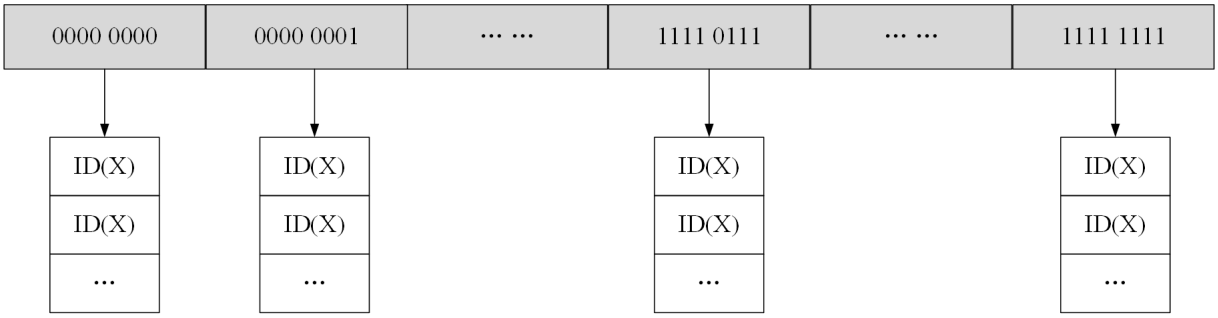


FIGURE 4. The inverted index structure.

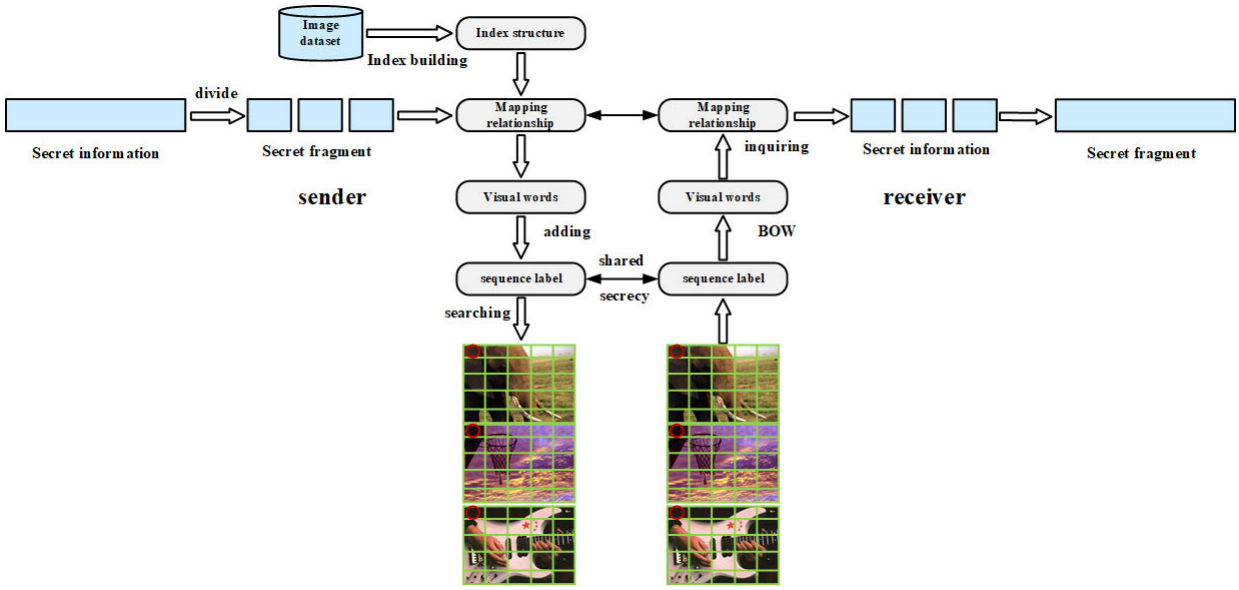


FIGURE 5. The flow chart of the steganography framework based on the BOW.

B. COVERLESS INFORMATION HIDING BASED ON THE BOW (BAG-OF-WORDS) MODEL

This method uses the bag-of-words model (BOW) [45] model to extract the visual words (VW) to express the text information to be hidden, so as to realize the hiding of text information in the image. As shown in Fig.5, firstly, the BOW model is used to extract visual words in the image set, and the mapping relationship between keywords and visual words in text information is established. Then each image is divided into several sub-images. For each sub-image, a histogram of visual words is calculated, and the visual words have the largest values in the histogram selected to represent the sub-image. According to the mapping relation, a set of sub-images with visual words related to the text information is found. The images containing these sub-images are used as stego-images [20] for secret communication. The experimental results and analysis show that the method has good performance. The coverless information hiding method based on image BOW model mainly consists of the following parts:

1) **Mapping relational library for visual words and text keywords**. The first step is to extract the SIFT features [44] of images, then uses the K-means [46] cluster to receive the codebook. The location of visual words in the dictionary is numbered and used as the ordinal ID of visual words. Then the mapping relation library of visual words and text keywords are established so that the text information to be hidden can be converted into the corresponding visual words lists [47]. As shown in Fig.6, text information and visual words can be inversely converted by both using the same mapping relational library.
2) **Multilevel inverted index structure**. To ensure efficient and accurate search, a multilevel inverted index containing visual words ordinal number information, sub-image location, and visual words frequency information is established.
3) **Hiding and extraction algorithms**. When hiding, the secret information shall be cut into words or words existing in the dictionary. Then, through the hash function agreed by the receiver and the sender in advance, the shared position tag sequence is randomly processed with the identity ID of the receiver to generate the position tag sequence for this communication. According to the location tag and mapping relation, the sub-image

The mapping relationship between dictionary and visual words

keywords	the corresponding visual word ID
南(Nan)	1
南京(Nanjing)	2
南京大学(Nanjing university)	3
…	…

FIGURE 6. Relationship between visual words and text keywords.

set which has mapping relation with the secret information fragment is retrieved from the image library. Only the image containing these sub-images is passed as a dense image without any other information. When extracting, the direction and scale of the received image are normalized. Then the position label sequence of this communication is obtained by the same method. Then, visual words sequences corresponding to sub-image positions are extracted by the same feature extraction method. Finally, the text information hidden in the image is obtained according to the mapping relationship. The specific process is shown in Fig.7.

C. COVERLESS COVERT COMMUNICATION METHOD BASED ON GIF

For the purpose of hiding the secret information, this method quantifies each GIF image in the existing carrier image library and extracts the attribute value of its extension [36]. For an extension attribute, according to the distribution of its attribute values in the carrier image library, the attribute values can be divided into multiple categories, and each category can be mapped into a secret message. Then, the classification of attribute values of multiple extension is combined and their completeness is verified to obtain a mapping space of secret messages to GIF images and the classification of multi-dimensional attributes that can be used as keys. The receiver analyzes the GIF image and the received key, reconstructs multiple categories of multi-dimensional extension attribute values, and then extracts the hidden information. Compared with [33], the proposed method has a larger embedding rate, because it proposes a way to use multidimensional extension attributes of the image. At the same time, since the cover library is combined with the theoretical analysis, this method can adapt to different cover libraries in the actual use of the scene. Fig.8 shows the flow of this method based on GIF which mainly consists of the following parts:

1) **Data hiding**. First of all, it is necessary to quantify the L extension attributes of each GIF image and obtain the extension attribute values of each GIF image in image data set. Ma center values are generated between its maximum and minimum values. Based on these central values, categories can be constructed on the attribute value domain and then mapped to bits. So, an extension attribute has M categories, and any class is generated into a set of multi-dimensional class combinations so that at least one image attribute value in the carrier image library is in this combined category set. Finally, the sender generates the secret key, and the package contains the maximum and minimum values of L extension attributes, as well as the set of values of L extension attributes.
2) **Data extraction**. The key is used to identify the classification of each dimension of the generality of a multidimensional extension attributes. What the receiver received is a collection of GIF images and critical information. According to these central values, the attribute range is divided into M categories, so that bits can be extracted for the extension attributes properties of each image. The bit-string extracted from each attribute is spliced in order to obtain the transmitted information.

D. COVERLESS COVERT COMMUNICATION METHOD BASED ON HOGS-BASED HASHING ALGORITHM

This method is similar to Method.1, as shown in Fig.9, its flow chart is the same as Method.1. It can be regarded as an extended version of the previous works of Method.1. Since histograms of oriented gradients (HOGs) have shown the robustness to a variety of common image attacks. Unlike Method.1, for each image in this approach, we divide it into some non-overlapping blocks and obtain each block' hash sequence by using the HOG-based hashing algorithm. Therefore, this coverless image steganography approach for secret communication includes three main components:

1) **Generation of hash sequences by using the HOGs-based hashing algorithm**. There are three steps in the procedure of the generation of the hash sequences. For each database image, it is transformed to gray-level image and then it is split into 4×4 non-overlapping blocks. For each block, we compute its HOGs. In the histogram, the value of each entry is compared to the mean value of all the entries to generate the hash sequence of the block (if the HOGs of the block is greater than the mean value, the value is 1, otherwise is 0). This method is described in detail in section 4.
2) **Construction of two-level inverted index file**. Similar to Method.1, each image is firstly indexed according to their blocks' hash sequences. A two-level inverted index file is constructed by using all blocks' hash sequences of the images. The first level of table T contains all possible block positions as entries, and the second level includes all possible sequences.
3) **Communication of secret information**. To implement the communication of secret information, the HOGs-based hashing algorithm and a block position P are needed to be shared between the sender and receiver beforehand. In Method.1, the block position P is set as a constant value. As shown in Fig.9, to improve the security, we use the one-time pad technique, which shifts

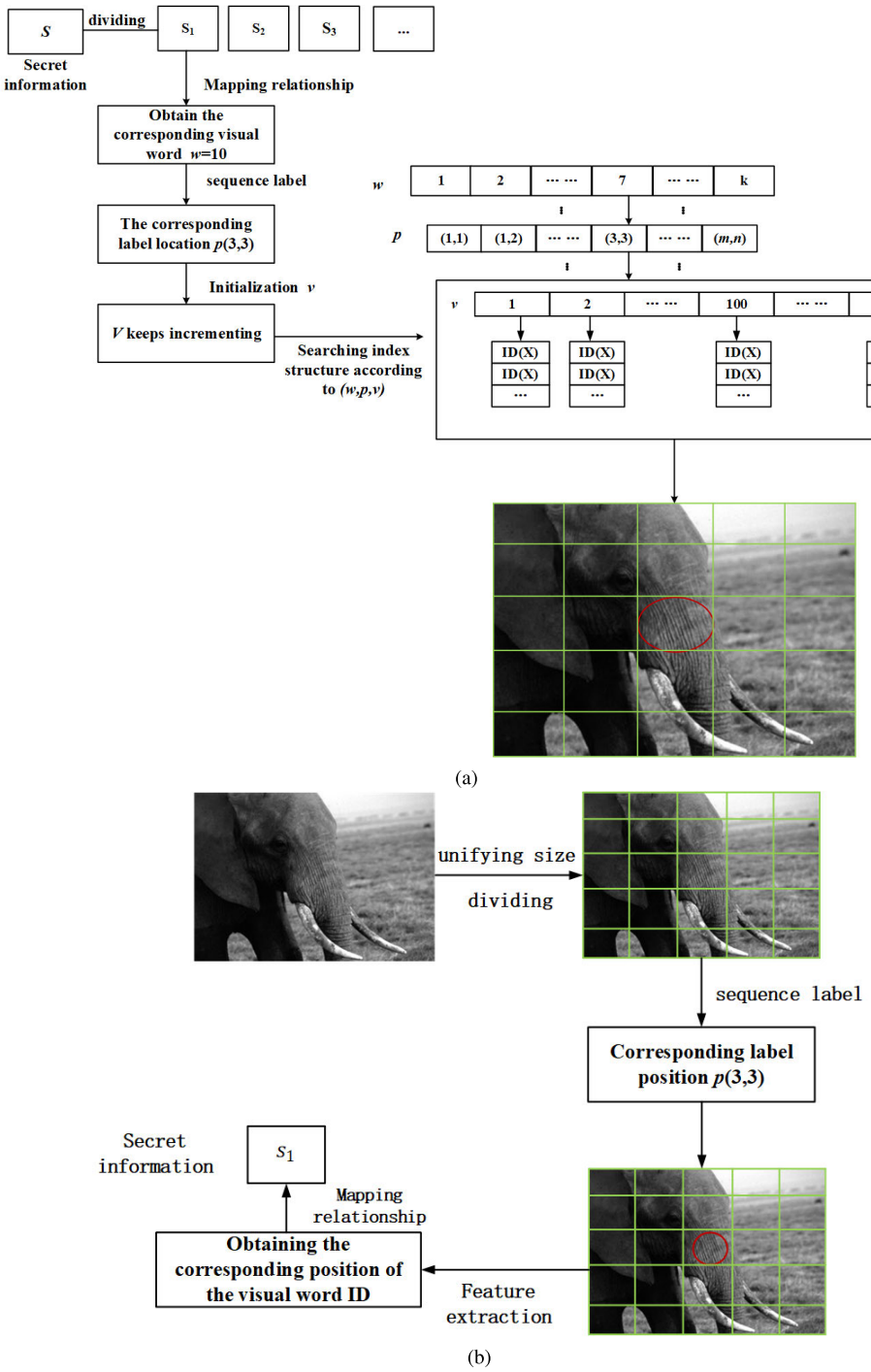


FIGURE 7. (a) Secret information hiding (b) Secret information extraction.

the block position P for each transmission, to choose the images.

E. COVERLESS STEGANOGRAPHY BASED ON ROBUST IMAGE HASHING

As shown in Fig.10, an effective and stable image hash by using the orientation information of the SIFT feature [44] points is proposed which can extract 18-bits binary sequences as the robust hash value from each image. Then a local image database is created and the corresponding hash values of these images in the database are calculated. Secondly, the secret message is divided into segments of the same length as the hash sequence. A series of images are selected from the image database by matching all the secret information segments and hash sequences. The information containing the binary secret data length and the direct correspondence between the hash sequences and the image is hidden in another image picked randomly from the image database with LSB

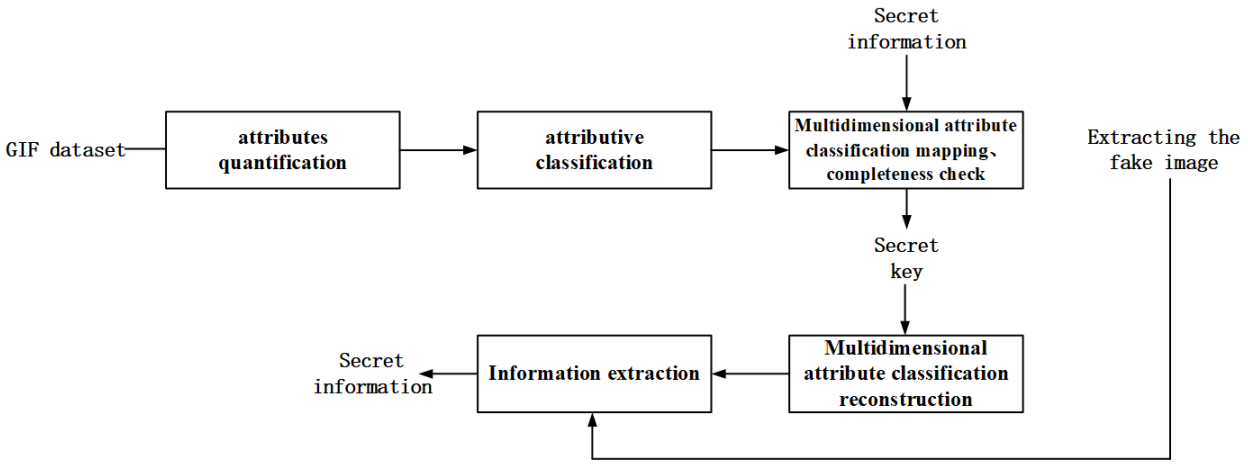


FIGURE 8. The flow chart of the steganography framework based on GIF.

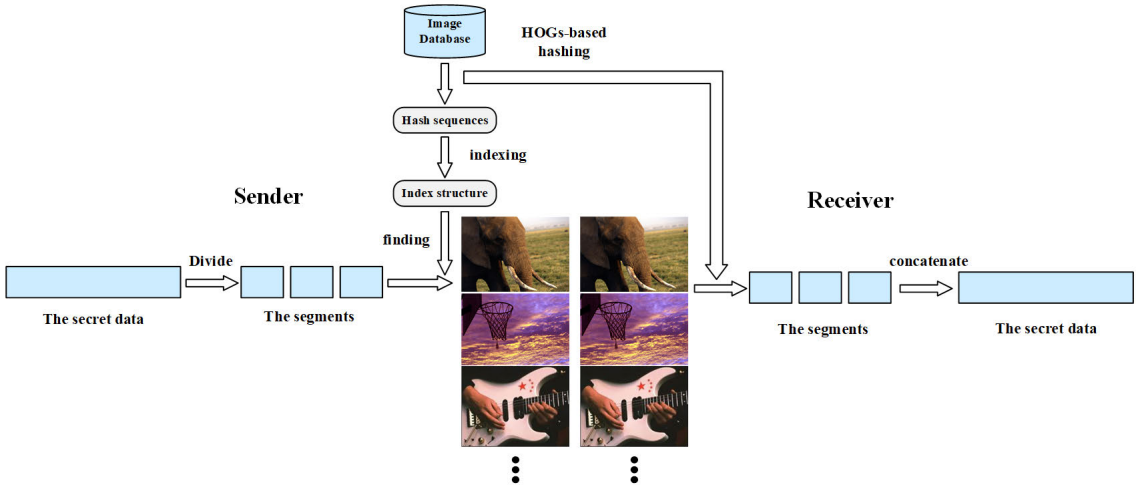


FIGURE 9. The flow chart of the steganography framework based on HOGs-based hashing algorithm.

algorithm [3], [4]. Thus, the images containing secret information are actually a series of images without being embedded and one containing additional information. The receiver accepts these images in turn and gets some necessary information from the additional image firstly, then extracts secret information from the other images using the shared hash algorithm. In addition, the proposed method is compared with state-of-art coverless steganography method proposed in [33] which also based on image hash, and results show that this method has higher capacity, robustness, and security. This method consists of the following three parts:

1) **Robust image hashing method**. 512 × 512 image is obtained through image normalization, and then the images are divided into 3 × 3 segments. For each stable point extracted on image blocks which are chosen by adjusting the threshold, the appropriate size of a window is selected around the point (circular area). The gradient directions of all the sampling points in the window are accumulated to form a histogram. Then if the max one locates between 0° and 90°, the hash value of this image block is set to 00; otherwise, if the max one locates between 90° and 180°, the hash value is set to 01, and so on. The orientation information of selected obvious extreme points of each block is counted, and the hash values of nine segments are connected in certain order to form final image hash. This method is described in detail in section 4.
2) **Inverted index of quadtree structure**. In order to improve the retrieval and matching efficiency of this hash system, an inverted index of quadtree structure is designed. As shown in Fig.11, the height of this quadtree is 10; each node has four child nodes whose values are 00, 01, 10, 11, corresponding to four different values of image blocks. There is a list in each leaf node respectively to store the information about images, the hash sequences of which in some particular order of blocks is the same as the current secret information segment. To Protect secret information from

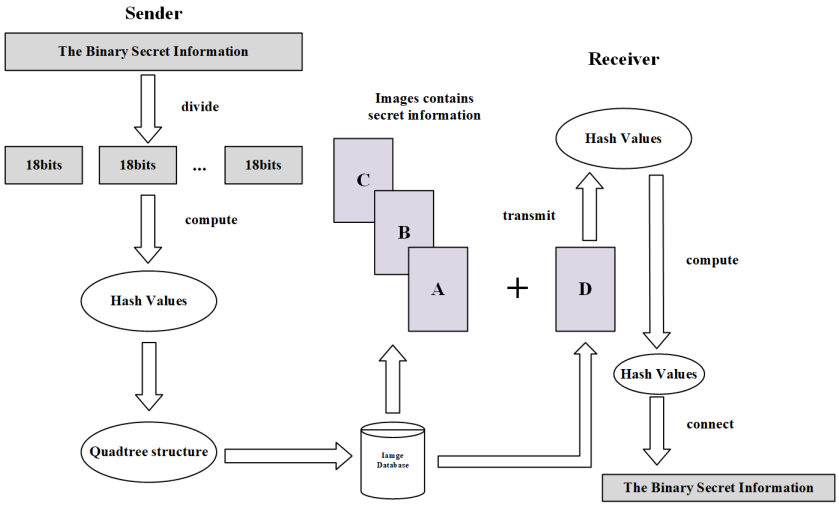


FIGURE 10. The flow chart of the steganography framework based on robust image hashing.

detection, there are at least two images information stored in each leaf node.

3) **The process of information hiding and extraction**. The secret information must be divided into n secret binary segments. Supposing the length of secret data is L, n, and L satisfy the following relationship:

$$n = \begin{cases} \dfrac{L}{18}, & if \quad L\%18 = 0 \\ \dfrac{L}{18} + 1, & otherwises \end{cases} \tag{1}$$

Matching inverted index quadtree structure with every two bits of these secret data segments computed by the previous step. Each segment will correspond to a leaf node. And each image stored in the leaf node can be used as the carrier of this confidential data segment. An image is selected randomly from the image database, then the information that the block orders of these images matched from Step 2 and the length of secret information L is hidden into this image with LSB replacement steganography. The receiver sequentially receives all images and extracts the information that the order of image blocks and the length of secret data from the additional image. To protect the integrity of confidential information, the data length s contained in the last image is computed as follows:

$$n = \begin{cases} 18, & if \quad L\%18 = 0 \\ L\%18, & otherwises \end{cases} \tag{2}$$

F. COVERLESS IMAGE INFORMATION HIDING BASED ON GENERATIVE MODEL

This framework is the first one to propose the coverless image information hiding method based on a generative model [39]. The secret image is input into the generative model database to generate normal and independent images with different

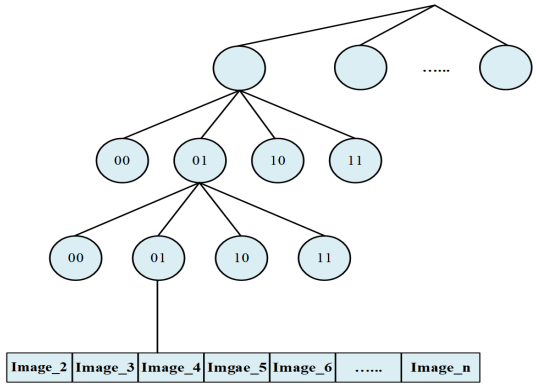


FIGURE 11. The inverted index of the quadtree structure.

meanings from the secret image. Then the receiver receives the generated image, and the generated image is input into the model database to generate another image that is the same as the secret image. The sender and receiver share the same data set and the same parameters. Therefore, we only need to transmit the mean standard image which is independent of the secret image to achieve the same effect as the secret image transmission.

1) **Training process**. WGAN (Wasserstein Generative Adversarial Networks) [48] is an improvement to GAN and is used to guarantee trainning stability. We keep train the generative model database through the WGAN until it can generate a meaning-normal and independent image which is not related to the secret image. We feed the secret image into the generative model, creating the meaning-normal and independent IMG which has nothing to do with the hidden image that we're transmitting. For example, as shown in Fig.12, we choose ''guitar'' as the secret image, it can generate the IMG visually the same as ''elephant'' that we want to transmit. In the meantime, we also train ''elephant'' to generate

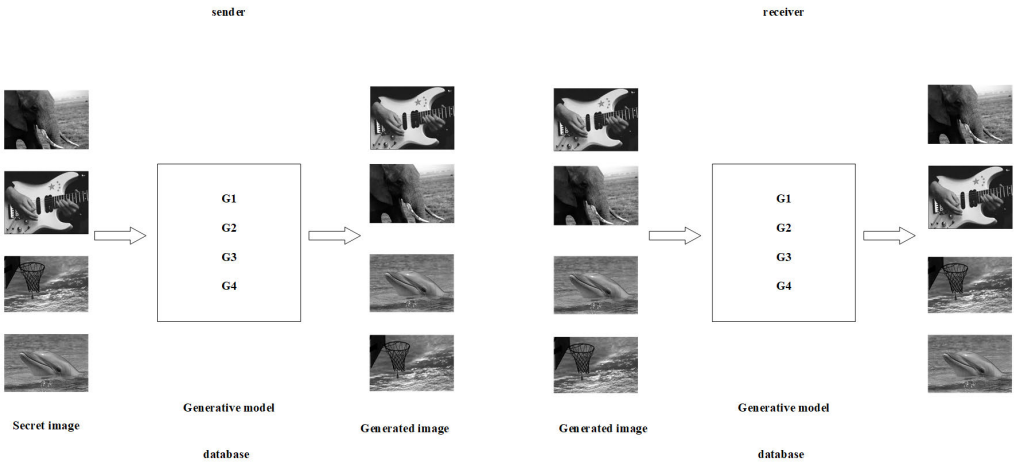


FIGURE 12. The training process.

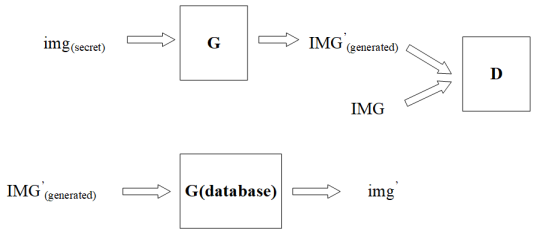


FIGURE 13. The flow chart of the steganography framework based on generative model.

the IMG visually the same as "guitar" through the WGAN. We save the corresponding generated models G1 and G2 as visual generated "elephant" and "guitar" respectively. Using the same method, we take more images as the secret images to experiment respectively and save the corresponding generative models, and apply them to the next experiment. We put the generative model G1, G2, etc. of generating visually the same as "elephant", "guitar", etc. in a database respectively so that the generative model database is built.

2) **The process of information hiding and extraction**. Since both the sender and the receiver train well the generative model database, the sender and receiver share the same dataset and the same parameters. As shown in Fig.13, we only need to deliver a meaning-normal image which is not related to the secret image to the receiver, so that the receiver can generate an image visually the same as the secret image.

G. COVERLESS INFORMATION HIDING BASED ON DCGAN

Volkhonskiy et al. proposed a new model for generating image-like containers based on Deep Convolutional Generative Adversarial Networks (DCGAN [34]) called Steganographic Generative Adversarial Networks (SGAN), which opened a new field for applications of GAN. As shown in Fig.14, in DCGANs, the Stego image is generated by the generator based on the pre-processed secret information, and no information is embedded in the Stego image. While in this method, DCGANs is used to generate a cover image which itself already contained secret information. DCGANs is carefully trained at different stages, and the confidential information is mapped into a noise vector [40]. To sum up, the mapped noise vectors are trained into stego images, then the DCGANs is trained to extract the noise vectors from the stego images. This coverless steganography method based on DCGAN consists of the following parts:

1) **Stego image generation**. The secret information is divided into segments, and several bits (two or three) of the segment are mapped to a noise vector. DCGANs is trained on an image set generator and G is obtained after DCGANs convergence. The stego images are produced by G which consists of a fully connected layer and four deconvolution layers and it is used to fit the data distribution of the real images in the training set to generate an artificial image. Note that the larger the dimension of noise, the richer the details in the generated image.
2) **Training of the extractor**. Based on the recovery errors from a large number of random noise vectors, a CNNs model is trained, called the extractor E. E has four convolutional layers, a leak-Relu activation function and batch normalization are used in each layer which has no pooling layer or dropout operation. In addition, a fully connected layer is used after the last convolutional layer. The generator produces stego images as the input of the extractor with dimensions of $64 \times 64 \times 3$ according to these noise vectors, and the output is a noise vector with dimensions of 1×100. When the loss of training is sufficiently small or the network of E is convergent, E is used to recover the secret data from the stego image generated by G.
3) **Secret communication**. After training the DCGANs, the sender and receiver own the network and parameters of G and E respectively. The sender divides

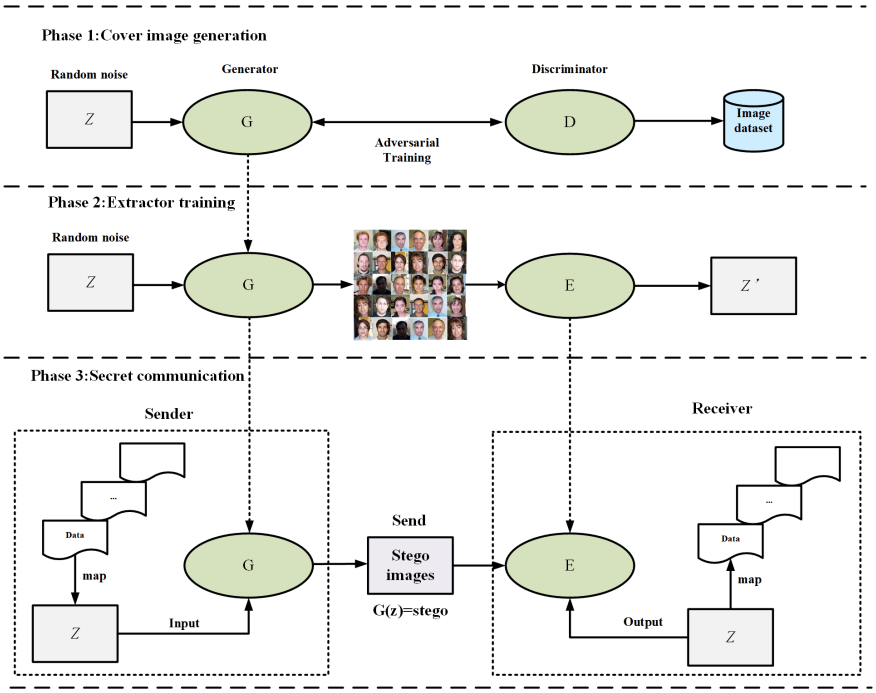


FIGURE 14. The flow chart of the steganography framework based on DCGANs.

the secret information into segments according to the capacity of steganographic image, and generates stego images through G according to noise z. The receiver receives the stego images, extracts the noise vectors by E, and restores the secret data according to the reverse mapping rules.

H. COVERLESS STEGANOGRAPHIC BASED ON PARTIAL DUPLICATES OF A GIVEN SECRET IMAGE AS STEGO-IMAGES

This approach needn't any modification for transmitting the secret color image [21]. Firstly, a large number of images is collected to construct a large-scale database from networks. Then, each image is divided into a number of non-overlapping patches and the label for each patch of a given image is computed by using a robust hashing algorithm. The label is used as the location information. The location information of the image indicates which patch of the image is used for hiding secret information. Note that the location information is shared as a secret key between the sender and receiver. The feature from each image patch is extracted and an inverted index structure is built by using the hierarchical BOW. To conceal the secret image, the first step is to divide the image into several patches with the same size. Then, for each patch, the partial-duplicate image that contains the similar patches with the secret image is retrieved by using the inverted index files. Afterward, a number of partial-duplicate images are obtained, which can be considered as stego-images. Then, those stego-images are transmitted to the receiver. At the receiver end, the location information is used to extract those patches from the stego-images. Because the secret image and its partial duplicates share one or several similar patches, the secret image can be recovered by stitching those patches together. As shown in Fig.15, there are three important steps, which are as follows:

1) **Index process**. The sender divides the image into several image patches and extract feature for each image patch construction of inverted index files.
2) **Hiding process**. For a given secret image needed to be hidden, similar to the process of the database image, it can be divided into a set of patches. The location information is employed for hiding the secret message. At the same time, the secret key is defined, which is shared between the sender and receiver in advance. Then, according to the new set of the label, many partial duplicates from the image database can be obtained by using the index files. The final partial-duplicate image which represent the secret image patch is the smallest distance between the secret image patch and candidates by computing Euclidean distance. The partial duplicates should be displaced by their original images.
3) **Extracting process**. The receiver can obtain the same location information as the embedding process through the key. The first sub-step is that receiver obtain each partial duplicate from the original image using the location information. The second is to give a blank area which size is the same to the secret image. Finally, the receiver place these duplicates in the blank area one by one to generate an image, which is visually similar to the secret image.

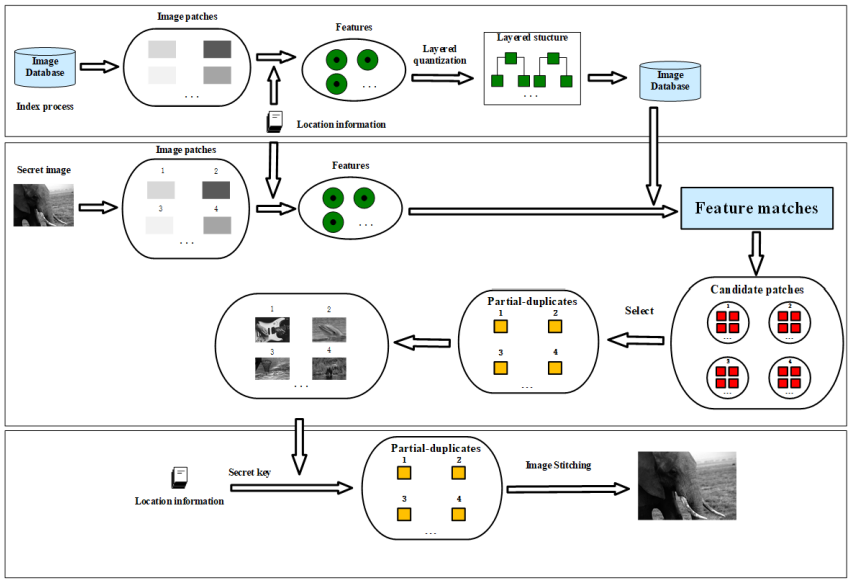


FIGURE 15. The flow chart of the steganography framework based on partial duplicates.

I. COVERLESS INFORMATION HIDING BASED ON THE AVERAGE PIXEL VALUES OF SUB-IMAGES

In the framework of coverless information hiding based on the average pixel values of sub-images, the secret information is segmented according to the structure of a Chinese sentence. And then the position of each segment could be got according to the dictionary and the label information of the hash sequence in each part of the hash array of images can be received according to the position of each secret information segment. Then the stego-image that randomly selected from the corresponding images is sent to the receiver. After receiving the image, the receiver blocks the image and gets the sub-images at first and generates a hash sequence by the average pixel values of sub-images. Then the receiver partitions the hash sequence and gets the label information according to each segment of the hash sequence and the hash array. And then the receiver can get corresponding position information of each segment of secret information. Finally, each segment of secret information can be extracted according to the position information [41]. The complete flow of this method can be seen in Fig.16, the entire process consists of the following parts:

1) **Building data sets**. A Chinese dictionary and a 50968 $\times$ 80 hash array are built. The Chinese dictionary is composed of four parts, including the subjects, the predicates, the objects, and the prepositions and each part of them contains 8 different Chinese words. To communicate with the dictionary, the 50968 $\times$ 80 hash array is divided into four 50968 $\times$ 20 hash arrays, which are marked as {M1,M2,M3,M4} respectively.
2) **Mapping the secret information**. Equivalent to the Chinese dictionary, every 20-bit hash sequence in the four 50968 $\times$ 20 hash arrays is labeled according to its decimal of 20-bits hash. The position information of each segment in the dictionary can be received through segmenting the secret information into four segments. The label information will be obtained because of the match between the position information and the hash label. Then the corresponding hash sequences which according to the relationship of the label and corresponding binary decimal and the decimal of 20-bits hash can be got.
3) **Multi-level index structure**. As shown in Fig.15, the 20-bits hash sequences marked as B1s, B2s, B3s, and B4s respectively are obtained from the previous label information section. Then, the Image1s that contain B1s in M1 are retrieved from the image database. Then, the Image2s that contain B2s in M2 are searched from Image1s. Then, the Image3s that contain B3s in M3 are retrieved from Image2s. Finally, the Image4s that contain B4s in M4 are searched from Image3s is the Stego-images.
4) **Information hiding**. The secret information is divided into four segments according to the subject, the predicate, the object, and the proposition. The label of the 20-bits hash sequence can be obtained through the mapping relationship and the corresponding 20-bits hash sequences can be obtained through the hash array. With the index method in Fig.17, it is easy to get all the corresponding stego-images from the image database.
5) **Information extraction**. The receiver divides the received stego-image into 80 sub-images and calculates the average pixel value of each sub-image firstly, and then obtain an 80-bits hash sequence by the hashing algorithm. The 80-bits hash sequence will be segmented into four 20-bits hash sequences, and the label is added to each 20-bits hash sequence. The receiver could get the position information of secret information

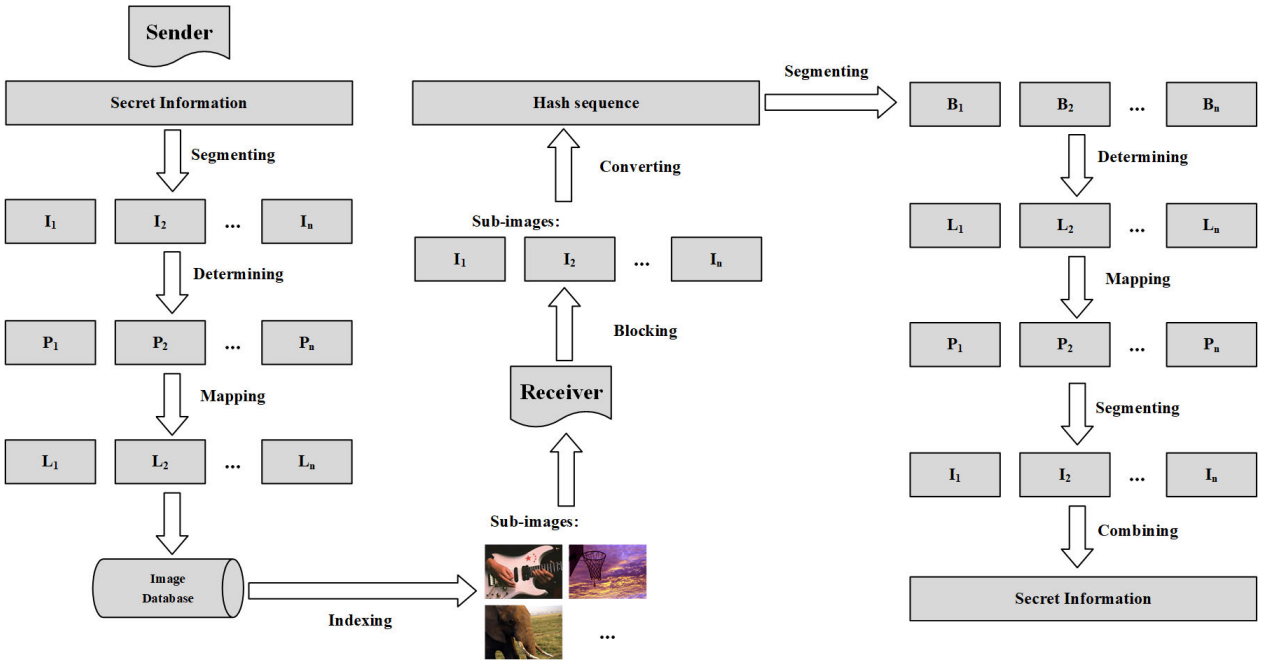


FIGURE 16. The flow chart of the steganography framework based on the average pixel values of sub-images.

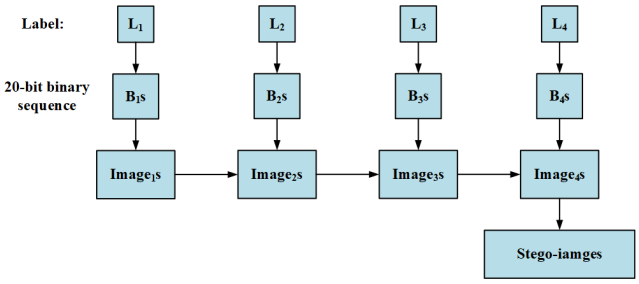


FIGURE 17. Multi-level index structure.

according to the mapping relationship. Finally, the four segments of secret information can be extracted according to the position information and be combined to get the secret information.

J. COVERLESS IMAGE STEGANOGRAPHY BASED ON DISCRETE COSINE TRANSFORM AND LATENT DIRICHLET ALLOCATION (LDA) TOPIC CLASSIFICATION

The sender extracts the features of the images in the database by processing the image set using the LDA topic model, picks a topic at random and processes these images by 8×8 blocks discrete cosine transform. Then the robust feature sequence is generated by the relation between direct current coefficients in the adjacent blocks. Finally, create an inverted index containing the feature sequence, DC, location coordinates, and image path. At the same time, the secret information is converted into binary sequence and divided into segments, and the image whose feature sequence is equal to the secret information segment is selected as the cover image according to the index. At the receiver, the feature sequence is calculated by the DC coefficients, and all feature sequences are concatenated to obtain the secret information [42]. The specific steps can be found in Fig.18 which consist of the following four parts:

1) **Topic classification of image database**. BOF is used to extract features and calculate word frequency of all images in the image database. LDA topic model is used to calculate the topic distribution of images. After obtaining the topic probability distribution, by comparing the probability of each topic, the image is classified into the topic prompt with the maximum probability, and the above steps are repeated until all images are classified into the corresponding topic.
2) **Extraction of feature sequence**. The size of all images is unified to $N \times N$ and partitioned to $\frac{N \times N}{l \times l}$ blocks. For each block, it is converted to YUV color space and Y channel of the blocks are partitioned into 16 sub-blocks. 8×8 DCT transform is done to each sub-block. Each bit of feature sequence belonged to each block is obtained according to the adjacent sub-blocks. Repeat the above steps to obtain the feature sequences of all sub-blocks in each image.
3) **Establishment of an inverted index**. As seen from Fig.19, each feature sequence corresponds to a list of three columns in an inverted index. DC is used to determine the order in which images are received will be computed and stored in the first column of the list. Details on defining DC values can be found in [30]. For example, if DC of the previous image is 804.78, the selected item can be the first one and the image path is ID (fish). The sub-block location coordinate p is (1,1)

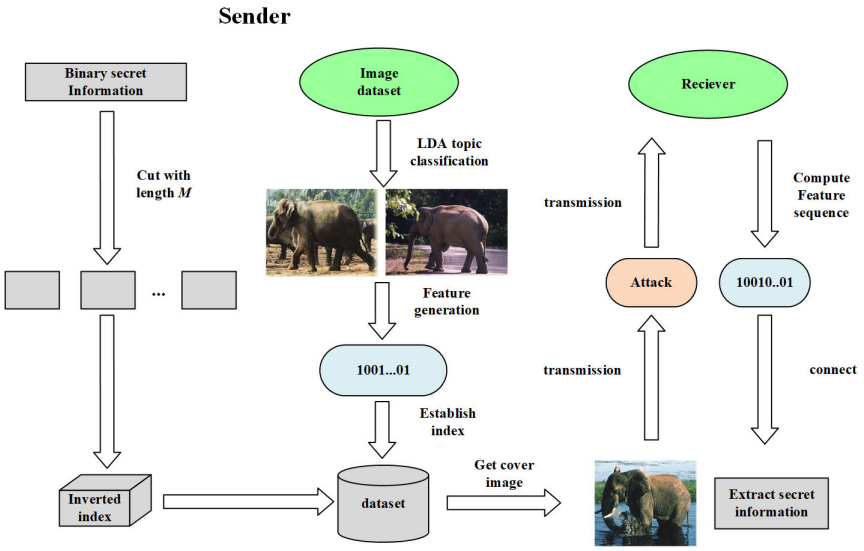


FIGURE 18. The flow chart of the steganography framework based on the discrete cosine transform and LDA topic classification.

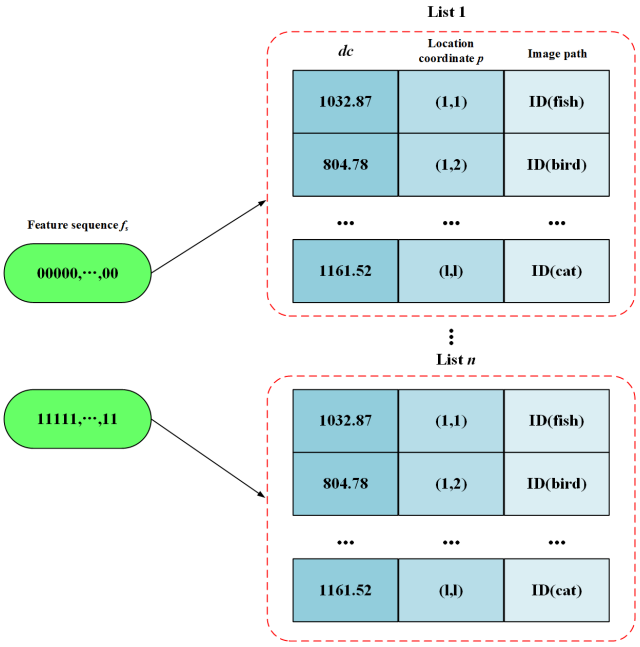


FIGURE 19. The flow chart of the inverted index.

and DC is 1032.87. Record the location coordinate p as side information and select the image according to the image path. Edge information and images are then sent to the receiver for secret communication. Each bit of the feature sequence belonged to the sub-block is calculated. Repeat the above steps to obtain the feature sequence of each image, so that the secret information is obtained.

4) **Coverless image steganography and secret information extraction**. Secret information is first divided into M binary information segments. The image database is trained and the feature sequences of each sub-block are obtained. Then, a reverse exponential is established. For the secret information segment, the image with the same feature sequence is retrieved. Repeat the previous step to get all overwritten images of the secret information. Record the number of zeros filled in the last paragraph and convert it to a binary sequence with M bits. Search the image using the previous method and add it to the cover image. If feature sequences are extracted at full size, there is no need to record location information. Otherwise, all position coordinates are stored in the matrix and encrypted by the AES encryption algorithm. At the receiver, the geometric calibration algorithm is used to correct the image, and then the order of the image containing secret information is obtained from the dc. Decrypt the matrix to obtain the sub-block position coordinates of each image. 8×8 DCT transform is performed to sub-block.

K. COVERLESS IMAGE INFORMATION HIDING BASED ON IMAGE BLOCK MATCHING AND DENSE CONVOLUTIONAL NETWORK

Similar to Method.8, this method transfers a set of stego-images which share one or several visually similar blocks with the given secret image [43]. It is worth mentioning that in this method, the supervised learning of deep learning is used to extract the high-level semantic features of image blocks. At the same time, DC coefficient is used to generate robust hash sequences. The secret information can be transmitted by matching and splicing the image blocks. As shown in Fig.20, this method includes three important steps:

1) **Pre-processing**. A large number of images on the network are searched and downloaded. Then, these images are divided into several non-overlapping blocks, the high-level semantic features of each block are

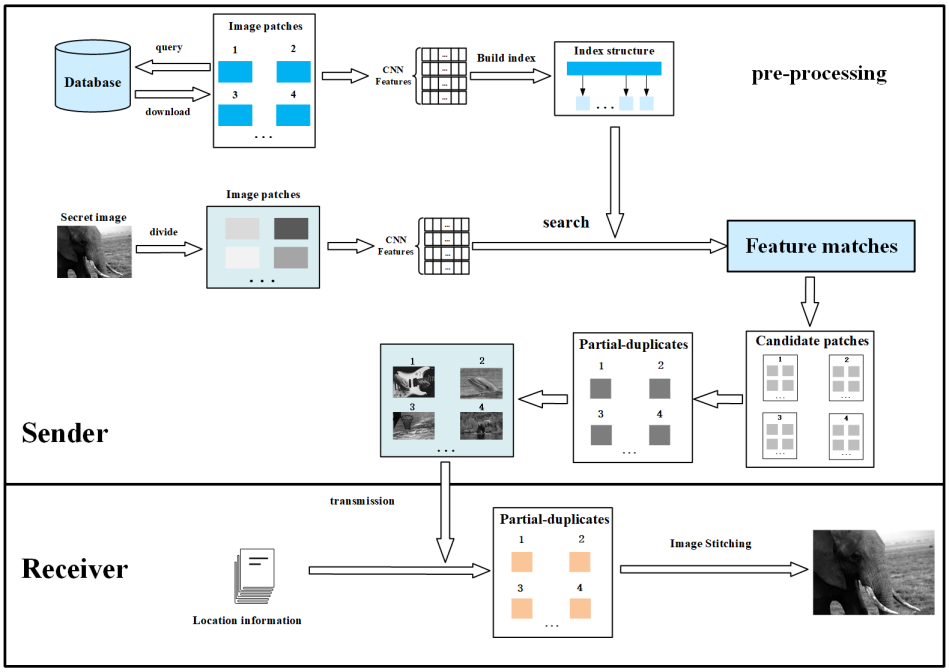


FIGURE 20. The flow chart of the steganography framework based on image block matching and dense convolutional network.

extracted with DenseNet for feature matching. Based on this, DCT is used to generate a robust hash sequence through the DC coefficients between adjacent image blocks and the inverted index structure is constructed.

2) **Sending processing**. The secret image is divided into blocks and the similar image blocks are matched by index structure. The most suitable image block are selected through Euclidean distance for transmission.
3) **Receiving processing**. The location information is used to obtain similar blocks from the stego-image. Next, a blank area of the same size as the secret image is given. Finally, these image blocks are placed into the blank area to generate an image similar to the secret image.

IV. SUB-PROBLEMS

In this section, we will introduce some sub-problems in coverless image steganography. All steganography frameworks are developed around these problems, and better steganography methods can be improved based on them, including pre-processing, feature extraction, generation of hash sequence and mapping relationships.

A. PRE-PROCESSING

In order to transmit secret information more accurately and quickly, the pre-processing is essential. Recently, in addition to segmenting secret information in advance, coverless image steganography approaches often construct own image data set by processing a large number of images in the same way to realize pre-processing. For example, Method.8 and Method.11 collected a lot of natural images, unified their size, cut them in the same proportion, extracted features from each image block to generate sequence codes, and when matching them, it could quickly match the required images only through the index instead of the natural data set without treatment [21], [43]. The mentioned feature extraction, generation of hash sequence and mapping relationships will be introduced detailedly in the following sections. Meanwhile, coverless steganography based on convolutional neural network often train deep learning net-work fully in advance to ensure the accuracy. In Method.11, the DenseNet121 network is trained previously through ImageNet data set [43]. As for the generative adversarial networks network, because of the antagonism between the generative model and the discriminative model [40], it is more necessary to train it. Method.6 and Method.7 both trained the generative model and the extractive model [39]. Because the pre-processing is very important, it is the key to innovation in the field of coverless steganography.

B. FEATURE EXTRACTION

As shown in the Fig.21, the most common method of processing image is to divide the image evenly into 9 small blocks and extract features of each block separately which are recorded as $\{fk^1, fk^2, \ldots, fk^9\}$. Then the final feature of the image is recorded as $FK = \{fk^1, fk^2, \ldots, fk^9\}$. As the most commonly used feature, average gray value is used in Method.1 and Method.4. As we all know, SIFT features [44] are more stable, which is usually detected at the edges of the image and some areas where brightness changes significantly. Therefore, SIFT features will not disappear with these

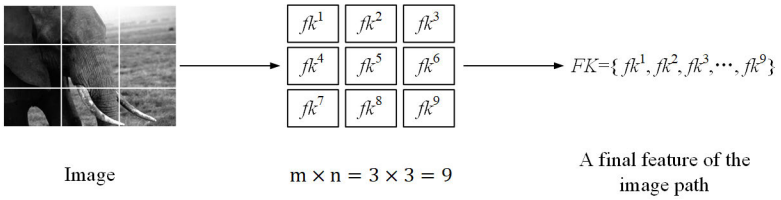


FIGURE 21. The process of extracting the final feature of an image.

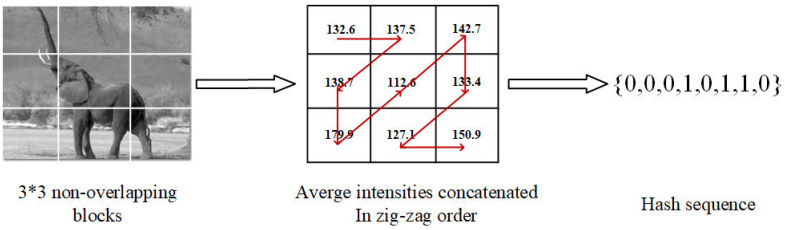


FIGURE 22. The initial process of generating the hash sequence.

various content changes, such as luminance change, contrast enhancement, JPEG compression, rescaling, noise adding and so on. In Method.2 and Method.5, the SIFT features are used to form the robust image hash. At the same time, Method.2 uses the bag-of-words model (BOW) [45] model to extract the visual words (VW) of the image to express the text information to be hidden. Encoding contextual clues into the BOW representation is a common technique to improve its recognition [38]. In addition, DCT [49] is also a common way to form the robust image hash in Method.10 and Method. 11.

C. GENERATION OF HASH SEQUENCE

During the communication, the stego-images could be manipulated by various typical attacks, such as rescaling, luminance change, contrast enhancement, JPEG compression and noise. Therefore, in order to ensure the secret information can be transmitted reliably and correctly with few losses and differences, the hash sequence of the image should not be changed during the communication. In this section, we mainly describe three common robust hash algorithms for steganography.

Fig.22 shows a simple hash algorithm. The complete image is converted into gray, and the image is divided into m x n patches and the average gray value of each region is calculated according to $\{V_1, V_2, \ldots, V_{m*n}\}$. These values are zigzagged to a hash sequence which denoted as $\{W_1, W_2, \ldots, W_{m*n-1}\}$through the following formula (3). Therefore, an image can represent m*n-1-bits information.

$$\begin{cases} W_i = 1 & if \quad V_i \geq V_{i+1} \\ W_i = 0 & otherwises, \end{cases} \quad 1 \leq i \leq n-1 \tag{3}$$

On this basis, an improved robust hash algorithm is proposed in Method.5, which can extract 18-bits binary sequences from each image. As shown in Fig.23, the local extreme point of each layer of the image difference pyramid is calculated on each image block. For each stable point extracted on image blocks by adjusting the threshold, the appropriate size of a window is selected around the point (circular area). The gradient directions of all the sampling points in the window are accumulated to form a histogram. Then the max one of all the values are selected and marked as M_i, the hash value of this image block (V_i) is set through (4). The max value in this histogram is illustrated by the red arrows. We count the orientation information of selected obvious extreme points of each block and connect the hash values of nine segments in order to form final image hash.

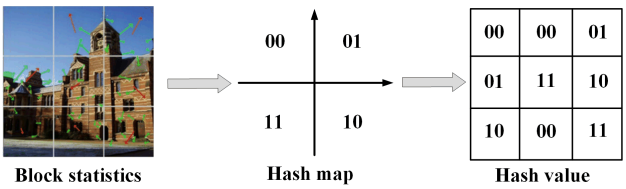


FIGURE 23. The improved process of generating the hash sequence.

$$\begin{cases} V_i = 00 & if \quad 0° \leq M_i \leq 90° \\ V_i = 01 & if \quad 90° \leq M_i \leq 180° \end{cases} \tag{4}$$

At the same time, there is another algorithm to generate the hash sequence code, as shown in Fig.24. This algorithm is used in Method.4. For each image, the algorithm transformed it to gray-level image and then split it into 3 × 3 nonoverlapping blocks. For each block, we compute its HOGs. Specially, each pixel's gradient magnitude and gradient orientation in the block are computed, and these gradient magnitudes are put into a histogram with N orientation bins. Then, each histogram bin is normalized. The histogram is denoted as $H_{(i,j)} = \{h^1_{(i,j)}, h^2_{(i,j)}, \ldots, h^N_{(i,j)}\}$. In the histogram $H_{(i,j)}$ of block $b_{(i,j)}$, the value of each entry $h^k_{(i,j)}$ is compared to the mean value of all the entries to generate the hash sequence of

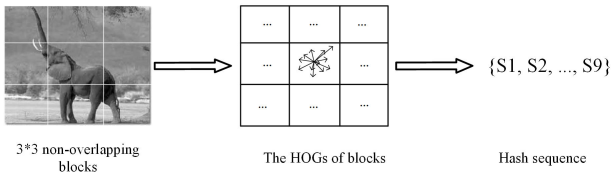


FIGURE 24. Generation of hash sequence by Using the HOGs.

the block denoted as {S1,S2, . . .,Sk, . . .,SN} by Eq.(5) where $\overline{H_{(i,j)}}$ is the mean value of all the entries of $H_{(i,j)}$.

$$\begin{cases} S_k = 1 & if \quad h^k_{(i,j)} \geq \overline{H_{(i,j)}} \\ S_k = 0 & otherwises, \end{cases} \quad 1 \leq k \leq N \tag{5}$$

D. MAPPING RELATIONSHIPS

This is well known that the reason why coverless image steganography is so popular is that it skips the steps of modifying images and directly uses natural images as carriers to convey secret information. Therefore, it is a key to establish the mapping relationship between secret information and natural images. Since the advent of coverless image steganography, the most common mapping relationships are based on the hash sequences and the secret information segments. The establishment process is described in the previous section. For example, by matching the secret information fragments and hash sequences, a series of images are selected from the image database in Method.1 and Method.5.

At the same time, there is another special dictionary-based mapping relationship. As shown in Method.2, the location of visual words in the dictionary is numbered and used as the ordinal ID of visual words so that the text information to be hidden can be converted into the corresponding visual words lists. A novel mapping relationship based on the binary sequences is proposed in Method.9 which includes a dictionary and a hash array. Then, the dictionary and the hash array are matched through the mapping relationship. To be specific, the secret information is divided according to the structure of a Chinese sentence, and the location of each part can be obtained according to the dictionary. Then the hash sequence label information in each part of the hash array in images is obtained for information transmission.

V. PERFORMANCE EVALUATION

In this section, we compared coverless image steganography with traditional image steganography and analyzed the performance of the existing coverless image steganography. The existing coverless image steganography can be divided into two categories: One is to transfer secret information based on the original attributes of the image through feature mapping such as Method.1, Method.2, Method.3, Method.5, Method.8, Method.9 and Method.10, the other is to transmit secret information by training Generative Adversarial Networks such as Method.6 and Method.7. The performance of these two types of steganography will be analyzed respectively. The results will be presented as pictures and tables.

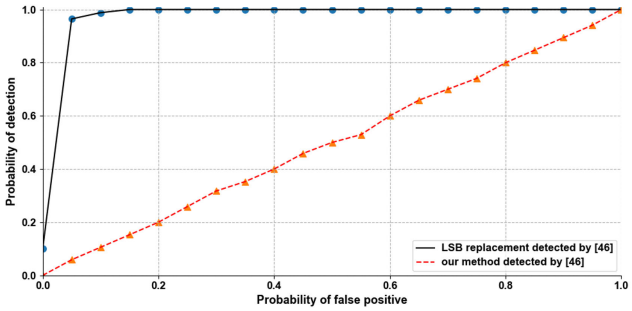


FIGURE 25. The ROC curves of LSB replacement and method.1.

A. COMPARISON WITH TRADITIONAL STEGANOGRAPHY

In this section, we compared coverless steganography with traditional steganography mainly from the advantages and disadvantages. It is worth noting that the robustness of image steganography should be evaluated according to its specific framework and algorithm, which is not directly related to whether the method is coverless steganography. Therefore, robustness is not discussed too much in this section, and details are listed in the next section.

Advantages: The greatest advantage of coverless image steganography is that it can resist the existing steganalysis tools to a great extent. The Receiver Operating Characteristic (ROC) curve is used to test the resistance of different methods to steganalysis tools. Since most of the existing coverless steganography is improved on Method.1, the performance of them are better than that of method.1. Therefore, we apply standard steganalysis tools, namely improved standard Pairs method to detect LSB replacement and Method.1 respectively. We use 10% LSB steganography (i.e. 0.1 bits per pixel). From Fig.25, we can see visually that coverless method has a perfect performance for resisting the standard steganalysis tools, and outperforms the traditional steganography methods.

Disadvantages: Since the direct mapping from secret information to hash sequences is the key of coverless image steganography, its greatest disadvantage is that the steganographic capacity is limited by the length of the image hash [33]. Although longer hash sequences can be generated to hide more secret information, the image database needs to be expanded at the same time. Otherwise, the larger the capacity, the worse the security. Among all coverless steganographies, Luo's method [43] has a relatively large capacity of 800(bits per pixel) which is very small volume compared to traditional steganography. High embedding capacity is one of traditional steganography advantages [13].

B. COVERLESS IMAGE STEGANOGRAPHY BASED ON THE ATTRIBUTES OF THE IMAGE

Since anti-steganography analysis was discussed in the previous section. For the coverless image steganography based on the attributes of the image, we mainly analyze it from two indicators: capacity and robustness.

TABLE 3. The capacity of proposed methods.

Methods	Years	Capacity(bits pixel^{-1})
1.Zhou's method	2015	8
2.Zhou's method	2016	1-1.57-1.86(0-2-3words.sub-image^{-1})
3.Ruan's method	2017	7(1D)/11(2D)/13(3D)/14(4D)
4.Zhou's method	2017	16
5.Zheng's method	2017	18
8.Zhou's method	2018	384
9.Zou's method	2018	80
10.Zhang's method	2018	15
11.Luo's method	2019	800

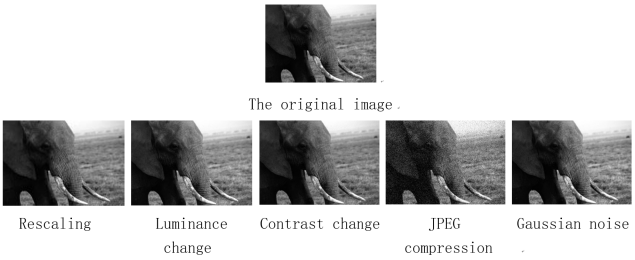


FIGURE 26. The attack ways on the image.

Capacity: Although coverless steganography has many advantages, compared with traditional steganography, it is still short of capacity. Meanwhile, according to KEY CHALLENGE, capacity has become a very important evaluation index. As can be seen from the Table.3, with the development of coverless image steganography, its capacity is also gradually increasing. It is worth noting that the mapping relation of the Method.2 is based on text and image so that its capacity is the sub-image corresponding to the length of words. Its specific parameter information can be seen in Table.4. At the same time, since the Method.3 expands the image attributes, its capacity increases with the expanded dimensions.

Robustness: In the transmission process, images will inevitably encounter some kinds of content damage, such as image noise, JPEG compression, rescaling, luminance change, contrast enhancement, and so on. The information extracted from the image must be able to resist these attacks. The success rate of secret data extraction is used to evaluate the robustness [50] of these methods. The results are shown in Table.4. Due to the particularity of Method.2 and Method.3, their main attacks are slightly different from those attacks, so they are listed separately in Table.5 and Table.6. Fig.26 shows the attacks on images, and their parameters are as follows:

1) Rescaling to 25 %;
2) Luminance change by adding the intensity of image pixels with 15;
3) Contrast change by multiplying the intensity of the image pixels with a fact of 1.4;
4) JPEG compression with a fact of 90%;
5) Gaussian noise adding with 0.5.

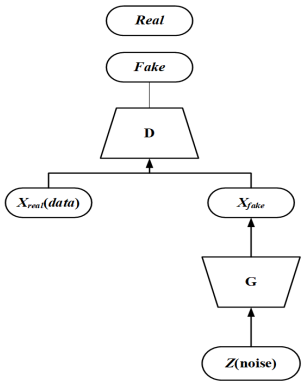


FIGURE 27. The structure of GAN.

C. COVERLESS IMAGE STEGANOGRAPHY BASED ON GENERATIVE ADVERSARIAL NETWORKS

Generative Adversarial Networks (GAN), proposed by Goodfellow *et al.* [51] in 2014, is a recent approach to deep unsupervised learning. The structure of GAN is shown in Fig.27. Two-person zero-sum game is the core idea of GAN. Any differentiable function can represent the main structure of GAN: generator (G) and discriminator (D) [52]. The main idea of this method is to train two neural networks simultaneously.

Method.6 is based on the WGAN model(Wasserstein Generative Adversarial Networks) [48] which is an improvement to GAN and is used to guarantee training stability. As shown in Method.6, we have successfully achieved the effect of coverless image information hiding based on a generative model. By feeding a secret image in WGAN, we can generate a meaning-normal image and transmit it.

Method.7 is based on GAN, it maps the secret information into a noise vector and the trains generator neural network model to generate the carrier image based on the noise vector. The index to evaluate the capacity is relative capacity (Relative capacity = Absolute capacity / The size of the image). In this method, the absolute capacity is greater than 37.5 (bits per pixel), the size of the image is 64×64 so that the relative capacity is $9.16e^{-3}$ (bits per pixel) which can meet the need.

In the proposed methods, there is no way to reveal the secret information except the extractor. Assuming that an attacker suspects that the transmitted image contains confidential information, it is difficult to extract the confidential information from the image, because it does not have the same GAN model as the communicator. In Method.7, even if the attacker accidentally extracts the secret information, it does not know the combination of dictionary and category label so that it can not decode the secret information into the original text information. Method.6 and Method.7 can resist detection of all the existing steganalysis tools.

It is worth noting that these two methods embed information directly in the generated image based on GAN. They guarantees the complexity of image texture, but not the integrity and quality of image.

TABLE 4. The success rate of extraction of different methods.

Method	Rescaling	Luminance change	Contrast enhancement	JPEG compression	Gauss noise
LSB replacement	54%	98.67%	98.03%	88.45%	70.05%
1	100%	100%	100%	93.64%	99.98%
4	100%	100%	100%	100%	99.96%
5	100%	100%	100%	100%	100%
8	100%	99.63%	98.37%	99.07%	40.67%
9	100%	100%	100%	100%	100%
10	100%	100%	100%	100%	100%

TABLE 5. The success rate of method.2.

Attack	rotation	translation	zoom	affine transformation	crop
Success rate of extraction	98.99%	100%	98.32%	97.85%	69.68%

TABLE 6. The success rate of method.3.

JEPG compression	20	40	60	80
Success rate of extraction	100%	100%	100%	100%
Time frame compression ratio	10%	20%	30%	40%
Success rate of extraction	100%	98.6%	97.9%	94.7%

VI. CONCLUSION AND FUTURE WORK

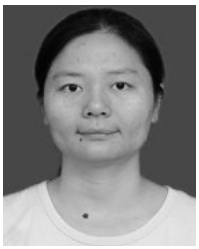
Image steganography is an important and challenging problem in information security and has received considerable attention. Thanks to ''coverless'', coverless image steganography can resist the developed steganalysis algorithm. As a comprehensive survey on coverless image steganography, this paper focuses on the stego image which hides secret information, highlights the recent achievements, provides a description for the framework of these methods and discusses performance for the most representative methods. Despite the tremendous successes achieved of coverless image steganography in the past several years, there remains a huge margin for improvement, especially in terms of the following domains:

1) **Improve the capacity**. We can divide the image into more sub-images according to actual needs, or produce a longer hash sequence to achieve higher information hiding capacity. However, when generating longer hash sequence to hide more secret information, the image database should be enlarged at the same time. Otherwise, the capacity is larger, the security and the invisibility will be worse. We should focus on how to enlarge the capacity of steganography while ensuring the retrieval accuracy and security. At the same time, we can construct a more effective inverted index, such as quadtree etc.
2) **Extend the image's other attribute information**. The embedding rate can be improved by using multi-dimensional image attributes. For example, multiple images can be connected to produce a GIF, however, the attributes of the GIF image extension module are limited. The next work will not only focus on the extension of GIF image modules, but also consider other attributes of images and increase the efficiency.
3) **Use a more sophisticated Wasserstein GAN(WGAN)**. In this paper, we described a new generative coverless stenography method based on generative adversarial networks. The resulting information resolution is still need to be improved. The next step is focused on creating a more realistic and clear natural information with a more sophisticated Wasserstein GAN (W-GAN) [48].
4) **Extend the nature image database**. In order to better convey secret information through natural images, our future work should extend our image database to improve the quality of the recovered image. We also pay more attention to the increase of hiding capacities.
5) **Improve robustness to against attack**. Existing coverless image steganography based on DCT transform can achieve good performance in subjective and objective detection resistance, and it is robust against most image processing attacks and geometric attacks. The following work will be devoted to the robustness against attacks of rotation and the content loss.
6) **Strengthen the implementation on hardware**. The reconfigurable device is becoming increasingly important for image processing and computer vision tasks. There are few hardware contribution techniques on steganography, these work on FPGA implementation of steganography offer higher speed and better throughput [14]. It is necessary to strengthen the implementation of coverless image steganography on hardware to increase the throughput of steganography system and protect more information.

REFERENCES

[1] J. Ni, J. Ye, and Y. I. Yang, ''Deep learning hierarchical representations for image steganalysis,'' *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 11, pp. 2545–2557, Nov. 2017.
[2] M.-M. Liu, M.-Q. Zhang, J. Liu, Y. Zhang, and Y. Ke, ''Coverless information hiding based on generative adversarial networks,'' *J. Appl. Sci.*, vol. 36, pp. 371–382, Mar. 2018.
[3] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne, ''A digital watermark,'' in *Proc. 1st Int. Conf. Image Process.*, Nov. 1994, pp. 86–90.
[4] C. H. Yang, C. Y. Weng, S. J. Wang, and H. M. Sun, ''Adaptive data hiding in edge areas of images with spatial LSB domain systems,'' *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 3, pp. 488–497, Sep. 2008.

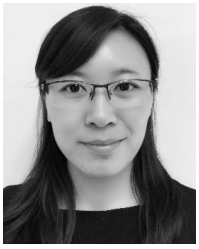
[5] T. Pevný, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Information Hiding*, vol. 6387. Berlin, Germany: Springer, Jan. 2010, pp. 161–177.
[6] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Proc. IEEE Int. Workshop Inf. Forensics Security*, Dec. 2012, pp. 234–239.
[7] V. Holub and J. Fridrich, "Digital image steganography using universal distortion," in *Proc. 1st IEEE Inf. Hiding Multimedia Secur. Workshop*, Jun. 2013, pp. 59–68.
[8] B. Li, S. Tan, M. Wang, and J. Huang, "Investigation on cost assignment in spatial image steganography," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 8, pp. 1264–1277, Aug. 2014.
[9] X. Li and J. Wang, "A steganographic method based upon JPEG and particle swarm optimization algorithm," *Inf. Sci.*, vol. 177, no. 15, pp. 3099–3109, Aug. 2007.
[10] W.-Y. Chen, "Color image steganography scheme using set partitioning in hierarchical trees coding, digital Fourier transform and adaptive phase modulation," *Appl. Math. Comput.*, vol. 185, no. 1, pp. 432–448, Feb. 2007.
[11] R. T. McKeon, "Strange Fourier steganography in movies," in *Proc. IEEE Int. Conf. Electrio/Inf. Technol.*, May 2007, pp. 178–182.
[12] I. J. Cox, J. Kilian, T. Shamoon, and F. T. Leighton, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 1673–1687, Dec. 1995.
[13] M. Y. Valandar, P. Ayubi, and M. J. Barani, "A new transform domain steganography based on modified logistic chaotic map for color images," *J. Inf. Secur. Appl.*, vol. 34, pp. 142–151, Jun. 2017.
[14] K. S. Shet, A. R. Aswath, M. C. Hanumantharaju, and X.-Z. Gao, "Novel high-speed reconfigurable FPGA architectures for EMD-based image steganography," *Multimedia Tools Appl.*, vol. 78, no. 13, pp. 18309–18338, Jul. 2019.
[15] D. Volkhonskiy, I. Nazarov, B. Borisenko, and E. Burnaev, "Steganographic generative adversarial networks," in *Proc. Workshop Adversarial Training (NIPS)*, Barcelona, Spain, Mar. 2017, pp. 1–15.
[16] T. Pevný, P. Bas, and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 215–224, Jun. 2010.
[17] J. Fridrich and J. Kodovský, "Rich models for steganalysis of digital images," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 868–882, Jun. 2012.
[18] J. Wang, J. H. Qin, X. Y. Xiang, Y. Tan, and N. Pan, "APTCHA recognition based on deep convolutional neural network," *Math. Biosci. Eng.*, vol. 16, no. 5, pp. 5851–5861, 2019, doi: 10.3934/mbe.2019292.
[19] M. Bilal, S. Imtiaz, W. Abdul, and S. Ghouzali, "Zero-steganography using DCT and spatial domain," in *Proc. ACS Int. Conf. Comput. Syst. Appl. (AICCSA)*, May 2013, pp. 1–7.
[20] Z.-L. Zhou, Y. Cao, and X.-M. Sun, "Coverless information hiding based on bag-of-words model of image," *J. Appl. Sci. Electron. Inf. Eng.*, vol. 34, no. 5, pp. 527–536, Sep. 2016.
[21] Z. Zhou, Y. Mu, and Q. J. Wu, "Coverless image steganography using partial-duplicate image retrieval," *Soft Comput.*, pp. 1–12, Mar. 2018.
[22] V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," *EURASIP J. Inf. Secur.*, vol. 2014, no. 1, pp. 1–13, Dec. 2014.
[23] M. Y. Valandar, M. J. Barani, P. Ayubi, and M. Aghazadeh, "An integer wavelet transform image steganography method based on 3D sine chaotic map," *Multimedia Tools Appl.*, vol. 78, no. 8, pp. 9971–9989, Apr. 2019.
[24] W. Liang, J. Long, A. Cui, and L. Peng, "A new robust dual intellectual property watermarking algorithm based on field programmable gate array," *J. Comput. Theor. Nanosci.*, vol. 12, no. 10, pp. 3959–3962, Oct. 2015.
[25] H. Gao, "Summary of research on key technologies of information hiding," *Electron. World*, vol. 9, pp. 146–148, 2016.
[26] V. Holub and J. Fridrich, "Random projections of residuals for digital image steganalysis," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 12, pp. 1996–2006, Dec. 2013.
[27] T. Denemark, V. Sedighi, V. Holub, R. Cogranne, and J. Fridrich, "Selection-channel-aware rich model for Steganalysis of digital images," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2014, pp. 48–53.
[28] Y. Qian, J. Dong, W. Wang, and T. Tan, "Deep learning for steganalysis via convolutional neural networks," *Proc. SPIE*, vol. 9409, pp. 94090J-1–94090J-10, Mar. 2015.
[29] P. Bas, T. Filler, and T. Pevný, "'Break our steganographic system': The ins and outs of organizing BOSS," in *Information Hiding*, vol. 6958. Cham, Switzerland: Springer, May 2011, pp. 59–70.
[30] G. Xu, H.-Z. Wu, and Y.-Q. Shi, "Structural design of convolutional neural networks for steganalysis," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 708–712, May 2016.
[31] M. Chen, V. Sedighi, M. Boroumand, and J. Fridrich, "JPEG-phase-aware convolutional neural network for steganalysis of JPEG images," in *Proc. 5th ACM Workshop Inf. Hiding Multimedia Secur*, Jun. 2017, pp. 75–84.
[32] Sh. Wang, X. Zhang, and W. Zhang, "Recent advances in image based steganalysis research," *Chin. J. Comput.*, vol. 32, no. 7, pp. 1247–1263, Jul. 2009.
[33] Z. Zhou, H. Sun, R. Harit, X. Chen, and X. Sun, "Coverless image steganography without embedding," in *Cloud Computing and Security*. Cham, Switzerland: Springer, Jan. 2016, pp. 123–132.
[34] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," in *Proc. ICLR*, Jan. 2016, pp. 1–16. [Online]. Available: https://arxiv.org/pdf/1511.06434.pdf
[35] Z. L. Zhou, Q. M. J. Wu, C.-N. Yang, X. Sun, and Z. Pan, "Coverless image steganography using histograms of oriented gradients-based hashing algorithm," *J. Internet Technol.*, vol. 18, no. 5, pp. 1177–1184, Sep. 2017.
[36] S. Ruan and Z. Qin, "Coverless covert communication based on GIF image," *Commun. Technol.*, vol. 50, no. 7, pp. 1506–1510, Jul. 2017.
[37] S. Zheng, L. Wang, B. Ling, and D. Hu, "Coverless information hiding based on robust image hashing," in *Intelligent Computing Methodologies*. Cham, Switzerland: Springer, Jul. 2017, pp. 536–547.
[38] Z. Zhou, Q. M. J. Wu, and X. Sun, "Encoding multiple contextual clues for partial-duplicate image retrieval," *Pattern Recognit. Lett.*, vol. 109, pp. 18–26, Jul. 2018.
[39] X. Duan and H. Song, "Coverless information hiding based on generative model," in *Proc. Comput. Vis. Pattern Recognit.*, Feb. 2018, pp. 1–4. [Online]. Available: https://arxiv.org/abs/1802.03528
[40] D. Hu, L. Wang, W. Jiang, S. Zheng, and B. Li, "A novel image steganography method via deep convolutional generative adversarial networks," *IEEE Access*, vol. 6, pp. 38303–38314, 2018.
[41] L. Zou, J. Sun, M. Gao, W. Wan, and B. B. Gupta, "A novel coverless information hiding method based on the average pixel value of the sub-images," *Multimedia Tools Appl.*, vol. 78, pp. 7965–7980, Apr. 2019.
[42] X. Zhang, F. Peng, and M. Long, "Robust coverless image steganography based on DCT and LDA topic classification," *IEEE Trans. Multimedia*, vol. 20, no. 12, pp. 3223–3238, Dec. 2018.
[43] Y. Luo, J. Qin, X. Xiang, Y. Tan, Q. Liu, and L. Xiang, "Coverless realtime image information hiding based on image block matching and dense convolutional network," *J. Real-Time Image Process.*, pp. 1–11, Sep. 2019, doi: 10.1007/s11554-019-00917-3.
[44] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
[45] J. Yang, Y. Jiang, A. Hauptmann, and C. Ngo, "Evaluating bag-of-visual-words representations in scene classification," in *Proc. Int. Workshop Workshop Multimedia Inf. Retr.*, Jan. 2007, pp. 197–206.
[46] J. A. Hartigan and M. A. Wong, "A K-means clustering algorithm," *Appl. Statist.*, pp. 100–108, Jun. 2013.
[47] Y. Long, Y. L. Liu, Y. Zhang, X. Ba, and J. Qin, "Coverless information hiding method based on Web text," *IEEE Access*, vol. 7, pp. 31926–31933, 2019, doi: 10.1109/ACCESS.2019.2901260.
[48] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," Jan. 2017, *arXiv:1701.07875*. [Online]. Available: https://arxiv.org/abs/1701.07875
[49] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 886–893.
[50] Y. Tan, J. Qin, X. Xiang, W. Ma, W. Pan, and N. N. Xiong, "A robust watermarking scheme in YCbCr color space based on channel coding," *IEEE Access*, vol. 7, pp. 25026–25036, 2019, doi: 10.1109/ACCESS.2019.2896304.
[51] I. Goodfellow, J. Pouget-Abadie, M. Mirza, D. Warde-Farley, B. Xu, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Conf. Adv. Neural Inf. Process. Syst.*, vol. 2, Jun. 2014, pp. 2672–2680.
[52] K. Wang, C. Gou, Y. Duan, Y. Lin, X. Zheng, and F.-Y. Wang, "Generative adversarial networks: The state of the art and beyond," *Acta Autom. Sinica*, vol. 43, no. 2, pp. 321–332, Mar. 2017.



JIAOHUA QIN received the B.S. degree in mathematics from the Hunan University of Science and Technology, China, in 1996, the M.S. degree in computer science and technology from the National University of Defense Technology, China, in 2001, and the Ph.D. degree in computing science from Hunan University, China, in 2009. She was a Visiting Professor with the University of Alabama, Tuscaloosa, AL, USA, from 2016 to 2017. She is currently a Professor with the College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. Her research interests include network and information security, machine learning, and image processing.

YUANJING LUO received the B.S. degree in automation from Hainan Normal University, China, in 2018. She is currently pursuing the M.S. degree in computer technology with the College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. Her research interests include deep learning and image processing.

XUYU XIANG received the B.S. degree in mathematics from Hunan Normal University, China, in 1996, the M.S. degree in computer science and technology from the National University of Defense Technology, China, in 2003, and the Ph.D. degree in computing science from Hunan University, China, in 2010. He is currently a Professor with the College of Computer Science and Information Technology, Central South University of Forestry and Technology, China. His research interests include network and information security, image processing, and machine learning.



YUN TAN received the M.S. and Ph.D. degrees from the Beijing University of Posts and Telecommunications (BUPT) of China, in 2004 and 2016, respectively. She is currently a Lecturer with the College of Computer Science and Information Technology, Central South University of Forestry and Technology. Her research interests include image security, compressive sensing, and signal processing.

HUAJUN HUANG received the B.S. degree in applied physics from Yunnan University, China, in 2001, and the M.S. and Ph.D. degrees in software engineering from Hunan University, China, in 2004 and 2007, respectively. He is currently a Faculty Member with the School of information Technology and Management, Hunan University of Finance and Economics. His current research interests include information hiding and hidden information detection, blockchain security, anti-phishing, privacy protection in cloud computing, and big data.