IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 5 August 2022, accepted 24 August 2022, date of publication 29 August 2022, date of current version 8 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3202905

SURVEY

Digital Image Steganalysis: Current Methodologies and Future Challenges

WAFA M. EID[1], SARAH S. ALOTAIBI[2], HASNA M. ALQAHTANI[1], AND SAHAR Q. SALEH[3]

[1]Department of Information Technology, College of Computing and Informatics, Saudi Electronic University, Riyadh 11673, Saudi Arabia
[2]Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 145111, Saudi Arabia
[3]Department of Computer Science, Faculty of Computing and Information Technology, Taiz University, Taiz 6803, Yemen

Corresponding author: Wafa M. Eid (w.eid@seu.edu.sa)

ABSTRACT With the growing use of the internet and social media, data security has become a major issue. Thus, researchers are focusing on data security techniques such as steganography and steganalysis. Steganography is the approach of concealing the existence of secret messages in digital media for secure transmission. Steganalysis techniques aim to detect the existence of concealed messages and extract them. Digital image steganography and steganalysis techniques are classified into the spatial and transform domains. In this paper, we provide a detailed survey of the state-of-the-art works that have been performed in two-dimensional and three-dimensional image steganalysis. We present the most popular datasets and explain some steganographic methods for embedding hidden data. Steganalysis is a very difficult task due to the lack of information about the characteristics of the cover media that can be exploited to detect hidden messages. Therefore, we review studies performed on image steganalysis in the spatial and transform domains using classical machine learning and deep learning approaches. Additionally, we present open challenges and discuss some directions for future research.

INDEX TERMS Steganography, steganalysis, deep learning, machine learning.

I. INTRODUCTION

A vast amount of digital media, such as image, video, and audio, are published on the internet every day. These digital media might contain hidden messages that can be embedded in plain sight using a well-known method called ''information hiding'' [1], a technique that allows users to conceal information via digital data with no perceptible effect on the data. Thus, people cannot detect whether there is a hidden message inside the digital data. The term steganography is formed of two Greek terms, ''steganos'', which means ''covered'', and ''graphein'', which means ''writing'' [2]. Therefore, steganography is a method that hides specific information inside digital data [3]. The input data are called the cover object, and the output is the stego object which contains the hidden message.

There are major distinctions between steganography and other information-embedding methods. The main difference between steganography and cryptography, for example, is that cryptography hides the message, while steganography hides the presence of the message. Watermarking is used to protect owners' property rights, and its aim is to add additional information to the source (cover data). Currently, many steganography techniques exist in the spatial and transform domains [4], [5], [6]. With the increased development and use of steganography techniques, there is a need to detect hidden messages. Steganalysis is an approach that distinguishes whether a message is hidden by steganography inside a certain media [7]; it is categorized into passive and active types [7]. The hidden data in the spatial domain are embedded directly by adjusting the value of the pixels in the cover image. In contrast, the hidden data in the transform domain are embedded in the coefficients of the cover image. Thus, Passive steganalysis detects the existence of hidden messages, while active steganalysis retrieves the hidden messages, and they are further classified into the spatial and transform domains. Image steganalysis methods utilize feature-based approaches to extract the discriminative attributes from images, such as local binary patterns (LBP) [8] and the subtractive pixel adjacency model

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

(SPAM) [9]. Steganalysis methods can be designed to detect a specific embedding algorithm, as in [10], [11] and [12]; they can also act universally to detect the existence of hidden data regardless of the embedding algorithm, as in [13], [14] and [15]. Many methods have been proposed for steganalysis applications based on machine learning and deep learning algorithms.

Classical machine learning consists of two parts: the first part is the feature extractor while the second part is the trainable classifier. The feature extractor is used to obtain distinguishing features from input data that are used by the classifier for training. Machine learning includes many techniques, such as the support vector machine (SVM) [16], a machine learning tool introduced by Vapnik that can be used for classification and regression problems [17]. Other popular machine learning techniques include linear regression, in which the output is predicted by using a known parameter, principal component analysis (PCA), which reduces the dimensionality of a dataset [18], nearest neighbor [19], K-mean clustering [20], etc.

In the last few years, deep learning models have received a significant amount of attention in many fields. In 2015, the first steganalysis technique was developed using a convolutional neural network (CNN), which is a deep learning technique [21]. Deep learning is considered a subset of machine learning that can be seen as a black-box framework. It integrates the feature extraction and classification steps into one process, thus allowing an end-to-end learning process for the machine. Deep learning models use forward processes to learn feature extractions and perform the classification directly from the input data. Then, in the backward direction, an updating of the extracted features based on the decision of the classifier is performed. This process is automatically repeated until the model's error is decreased. Fig. 1 illustrates the difference between the concepts of classical machine learning and deep learning. Fig. 1 (a) presents the classical machine learning concept. Fig. 1 (b) presents the deep learning concept.

Steganalysis is a very difficult field due to the lack of information about the characteristics of the cover media that can be exploited to detect hidden messages. There are many algorithms for the steganalysis field that use three common cover media: image, video, and audio. In this survey paper, we focus on steganalysis algorithms for two-dimensional (2D) and three-dimensional (3D) images. The main difference between 2D and 3D image steganography is in the cover image. For steganography in 2D images, the cover is an image where a message will be hidden within pixel intensities, while in 3D image steganography, the cover is a 3D mesh consisting of points or vertices in 3D geometry, and it will be manipulated to hide information.

Studying current methodologies and understanding future challenges in digital image steganalysis helps researchers achieve better outcomes in steganography and steganalysis. Thus, in this survey, we present a summary of the different types of algorithms for digital image steganalysis that have been developed using classical machine learning and deep learning technologies. We mainly focus on algorithms for images in the spatial and transform domains. Developing and adapting steganalysis techniques begins with a good understanding of steganography. Therefore, in Section II, we begin the survey by providing an overview of the steganography algorithms used on 2D and 3D images. In Section III, we outline the most commonly used datasets in steganalysis and categorize them into 2D and 3D datasets. In Section IV, we analyze various steganalysis methods for 2D and 3D images that have been performed using machine learning and deep learning. In Section V, we highlight some open research challenges in the steganalysis field. Finally, we conclude the paper in Section VI.

II. STEGANOGRAPHY

Any steganography technique can be defeated once its steganalysis technique is determined [22]. This section provides an introduction to digital image steganography and some steganography schemes in the spatial and transform domains.

Steganography is the science of communicating secretly by hiding multimedia data inside an appropriate multimedia carrier, such as an image, text, file, or video [23]. These multimedia carriers are called cover objects. The first steganography technique was developed in ancient Greece, and the importance of steganography has increased recently due to the increase in data exchange in social media networks. Image steganography techniques have been developed for information concealed exclusively in images. The secret message is hidden in a cover image and sent to a receiver in such a way that only the sender and the receiver are aware of its existence. Both the secret message and cover image constitute the input of the steganographic encoder. The stego image is obtained by embedding the secret message in a cover image. In the end, the stego and cover images are very similar and show no visible changes. The receiver must input a stego image into a steganographic decoder to read the secret message. A stego key is used for encoding and decoding the secret message.

There have been many steganographic techniques presented in the literature. All these techniques must satisfy at least three requirements to be applied correctly: The maximum amount of information that can be concealed inside the cover image (embedding capacity) must be considered; the visual quality of the stego image must remain unchanged (imperceptibility), and it must be robust against noise [3]. There are a number of methods to hide information inside a 2D image. These embedding methods can be in either the spatial or transform domain [24]. The idea of spatial domain embedding techniques is to use the actual physical location of a pixel of information in the image. These techniques are considered easy to implement because of the simplicity of their algorithms and mathematical analysis. Spatial domain techniques provide high embedding capacity; however, their robustness is weaker than their counterparts [25]. The most commonly used technique is least significant bit (LSB).

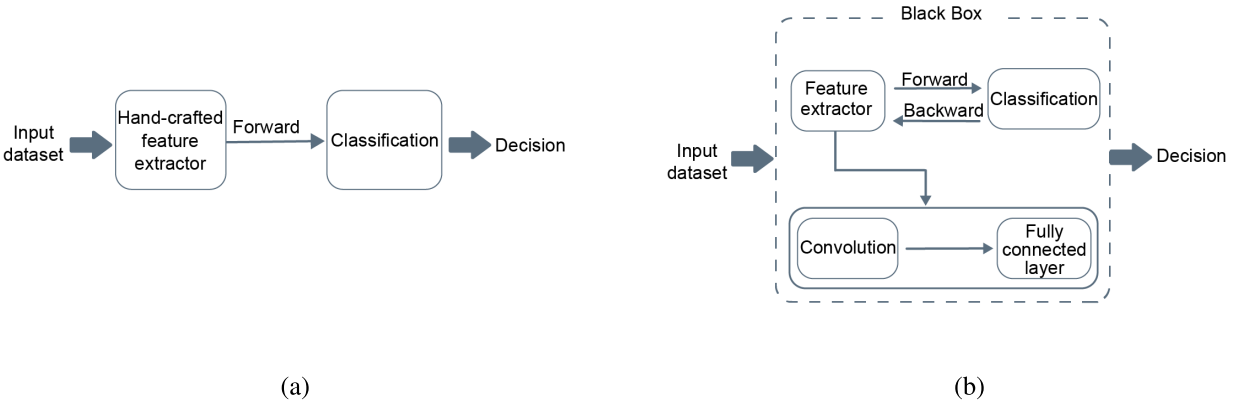


FIGURE 1. Comparison between deep learning and classical machine learning concepts. (a) Classic machine learning method. (b) Deep learning method.

LSB insertion is a simple data hiding approach that can be effortlessly applied in the spatial or transform domain over any digital cover media [26]. It is easily understood by users, and the data are hidden by replacing the least significant bit of each pixel. However, maximization of the hidden data capacity into a cover image by using LBS is limited because it may lead to more noise and distortion, affecting the resolution and quality of the image and thus making it easier to detect the hidden message. This method is easily detected by exposing the data to strong feature-based image steganalysis and recovering the least significant bit out of the pixels in the image.

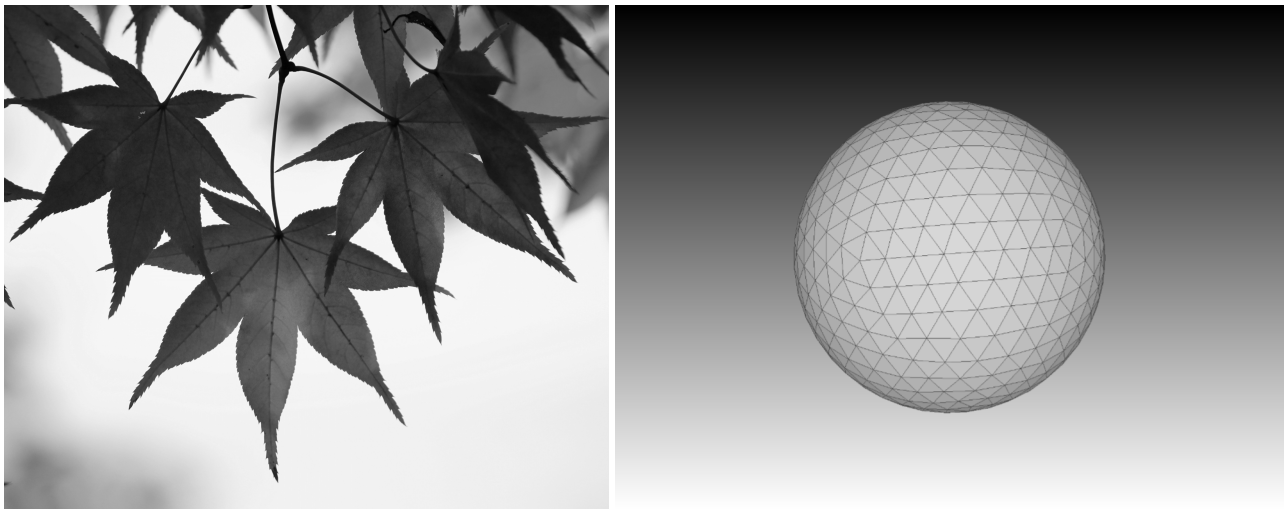
The need to remove noise from embedded data is essential to increase its security. Highly undetectable steGO (HUGO) was developed in 2010, and it is considered the first adaptive schema proposed in the literature [27]. HUGO is a nontraditional and content-adaptive steganographic technique that adaptively changes the embedding locations according to the image content. Therefore, this method achieves high security in the embedding process by covering the noise with the inhered noise.

Several adaptive steganography algorithms for the spatial domain have been proposed that follow the same embedding model of HUGO, for instance, an embedding algorithm for the break our steganography system (BOSS) competition [28], spatial-universal wavelet relative distortion (S-UNIWARD) [29], wavelet obtained weights (WOW) [30], high-pass and two low-pass (HILL) [31], and minimizing the power of optimal detector (MiPOD) [32]. WOW is a high-security steganography algorithm, as it changes the pixels of texture regions while maintaining the edge of the cover image [30]. S-UNIWARD steganography obtains advanced security by using the directional high-pass filter [29]. In the HILL algorithm [31], the embedding changes mainly in the textural areas; the algorithm uses a high-pass filter and two average low-pass filters to ensure that all pixels within the textural regions have relatively low costs. In contrast to all the above discussed adaptive steganography algorithms, the MiPOD algorithm minimizes the impact of embedding on the cover model in addition to obtaining a superior security level [32].

Transform domain embedding is a method for representing the signal in another form. However, the information content present in the image is not changed. Wavelet transform (WT) is a mathematical procedure used to transform a spatial domain into a frequency domain [33]. The main idea of using WT in image stenographic techniques is based on separating the high-frequency and low-frequency information on a pixel-by-pixel basis. Transformation techniques use JPEG compression due to the significant increase in available steganalysis tools. Discrete cosine transform (DCT), discrete wavelet transform (DWT), and discrete Fourier transform (DFT) in the embedding process are the utilized transform steganography techniques. The DCT domain embedding technique is very popular because it is the core of the lossy image compression algorithm known as JPEG, which is the format used for digital cameras [34]. In comparison to DCT, DWT shows high robustness, and the embedded secret image can be extracted with a high visual quality [35].

Another classification for steganographic methods is based on the coded formats of images. Image steganography can be applied in different formats for cover images, such as BMP, JPEG and GIF. High color quality JPEG format images are the most mainstream images in modern communications. The most efficient JPEG steganographic techniques are based on Syndrome Trellis Coding (STC) [36] and Uniform Embedding Distortion (UED) [37], which uses only nonzero DCT coefficients of different magnitudes with equal probability. This schema possibly leads to minimal artifacts for the statistics of all the DCT coefficients, which makes them naturally content adaptive [38].

Early steganography algorithms focused mainly on 2D images, videos, and audios. However, due to the rapid growth in digital media, the use of 3D images as input media in steganography algorithms has been consistently established in the past decade. A 3D image is a geometric setting that requires three coordinate axes to represent the position of a point. Steganography algorithms hide the secret data bits



(a) (b)

FIGURE 2. An example of 2D and 3D images. (a) 2D image of leaves (b) 3D image of sphere mesh.

inside the points of a 3D mesh. Fig. 2 shows the differences between 2D and 3D images.

The advantage of using 3D images is that they have a data structure with a high carrying capacity that provides extra security for embedding larger amounts of secret information. However, there are currently fewer 3D image steganography algorithms than 2D image steganography algorithms due to some challenges and complexities in 3D images. These algorithms can also be classified into either the spatial or transform domain. Due to additional efforts required to transfer the image in and out of the frequency domain, most work is done in the spatial domain.

Researchers have used 3D data hiding techniques for embedding data into 3D images. These techniques are detected by some of the steganalysis methods that will be discussed later in this paper. Cayre and Macq proposed a substitutive procedure-based steganography algorithm for 3D images that geometrically quantizes the 3D object into a two-state geometric [39]. The main drawbacks of this method lie in embedding capacity and distortion. Wang and Cheng [40] improved the method in [39] by employing a multilevel embedding procedure and using an advance jump strategy to increase the embedding capacity to three bits per vertex and reduce the distortion. Chao *et al.* [41] also presented an algorithm with a high capacity and a reduced distortion algorithm on a multilayer embedding. The problem with this method is that the rapidly increasing distortion limits the number of embedded layers. To balance the increasing embedding capacity and reducing distortion, Yang and Ivrissimtzis [42] designed an algorithm based on computing an appropriate quantization level for the 3D vertices and replacing the unused LSB with watermark bits. However, the large amount of embedded noise increases the error significantly, which may lead to malicious attacks. Tsai [43] proposed an adaptive steganography algorithm that achieved a high accuracy of the complexity estimation for each embedded vertex and the embedding capacity. A steganography method that combines both the spatial and representation domains is presented in [44]. A number of 3D data hiding schemes have been investigated in steganalysis for example adaptive-steganography [45], 3D wavelet-based high capacity and 3D wavelet-based fragile steganography [46], shifting and truncated steganography [47], distortion-free steganography [48], permutation steganography [49], the maximum expected level tree data-hiding approach [50], and a data hiding approach for polygon meshes [51]. Some researchers have tested their steganalysis techniques using other data hiding techniques, such as watermarking. The Laplacian coordinate-based watermarking method [42], the two variants of robust watermarking [52], frequency-based watermarking [53] and steganalysis-resistant 3D watermarking [54] are examples of watermarking algorithms. These techniques are detected by some of the steganalysis methods discussed later in this paper.

III. IMAGE STEGANALYSIS DATASETS

Steganography can be used on various types of media, such as images, videos, audio, etc. Therefore, researchers need to evaluate their steganalysis techniques on large datasets. We classify datasets into two categories, 2D and 3D datasets, and 2D datasets can be further categorized into grayscale and colored images. This section explains the commonly used datasets in the field of steganalysis.

A. TWO DIMENSIONAL DATASETS

1) GRAYSCALE DATASETS

A challenge called the Breaking Our Watermarking System (BOWS) [55] was created in 2007 by the International Challenges in Information Forensics and Security community to remove watermarks from three images. The second edition

TABLE 1. Commonly used 2D grayscale datasets in steganalysis.

Dataset	Number of Images	Format	Image Size	Steganographic Scheme	Year
BOWS-2 [56]	10,000	PGM	512 × 512	N/A	2008
BOSSbase v1 [28]	10,000	PGM	512 × 512	HUGO [27]	2010
BURSTbase [57]	7 × 9,310	JPEG	512 × 512	J-UNIWARD [29], J2-UNIWARD [57], and SI-UNIWARD [29]	2017
LIRMMBase 256×256 [58]	1,008	PGM	256 × 256	N/A	2015
LIRMMBase 512× 512 [58]	15,245	PGM	512 × 512	N/A	2015

TABLE 2. Commonly used 2D colored datasets in steganalysis.

Dataset	Number of Images	Format	Image Size	Steganographic Scheme	Year
Istego100K [63]	208,109	JPEG	1024×1024	J-UNIWARD [29], nsF5 [66], and UERD [62]	2019
Alaska 2 [60]	300,000	JPEG	512 ×512	J-UNIWARD [29], UERD [62], and J-MiPOD [61]	2020
ImageNet [64]	More than 15 million	JPEG	Different sizes	N/A	2009
RAISE [65]	8,156	Raw	Different sizes	N/A	2015
LIRMMBaseColor [58]	15,320	Raw	Not specified	N/A	2015
Alaska 1 [59]	80,000	Raw, JPEG	Different sizes	nsF5 [66], UED-JC, EBS, and J- UNIWARD [29]	2018

of the challenge (BOWS-2) was presented in 2008 [56]. This challenge was the inspiration behind creating the BOSS challenge in 2010 [28]. The BOSSbase dataset, which includes 10,000 grayscale images, has become one of the most commonly used datasets in the field of steganalysis. Although BOSSbase has had a great impact on steganalysis research, it has some limitations, especially with the emergence of new technologies. In 2017, the BURSTbase dataset was created [57]. It contains 7 × 9.310 images in JPEG format taken with a camera in burst mode. Table. 1 presents the details of the grayscale datasets.

2) COLORED DATASETS

In 2018, a new competition began under the name Alaska1 [59]. It aimed to provide a large dataset of images of different sizes taken by different cameras. Images were compressed with different quality factors. As part of the Alaska1 challenge, participants were given codes for steganographic schemes to build their own training sets without having to worry about cover-source mismatches. Alaska2 [60] is a follow-up competition to Alaska1. It contains a total of 300,000 colored images divided into 75,000 sets of cover, stego images (images with steganographic schemes applied to them such as J-UNIWARD [29], J-MiPOD [61] and Uniform Embedding Revisited Distortion (UERD) [62]) and 5,000 test images. IStego100K is a dataset created in 2019 [63]. It contains 208,104 colored images of size 1024 × 1024. Images were separated into 100,000 images for cover and 100,000 images for stego, and the remaining 8,104 images were used for testing. This dataset takes into consideration the mismatch between the training set and test set in a real environment. ImageNet is one of the most popular datasets used in classification problems [64]. It was created to promote research in computer vision. Some researchers use ImageNet as a dataset for steganalysis by creating stego images using a steganography scheme of their choice; it has been used in creating new models for steganalysis, as it contains more than 14 million colored images of various sizes. RAISE is another dataset created to support research in the image forensics and image processing fields [65]. Table. 2 illustrates the details of the colored datasets.

3) GRAYSCALE AND COLORED DATASETS

Few datasets contain both colored and grayscale images. One of the few that does is Steganalysis Real Test Version 1 (STEGRT1) [67]. STEGRT1 is a new dataset created in 2020 to evaluate steganalysis systems on real-world scenarios. STEGRT1 contains 8,000 cover and stego images of various sizes and properties. The Large Scale Steganalysis Database (LSSD) is a combination of different datasets [68]. The idea behind combining datasets is to increase the diversity and present real-world scenarios. Table. 3 shows more details of the grayscale and colored datasets.

B. THREE DIMENSIONAL DATASETS

Currently, with advancements in 3D technology and hardware, it has become easy to obtain 3D models for natural objects. These models have become commonly used in different fields, such as medical imaging, virtual reality, augmented reality, games, movies, and many more areas. 3D steganography has become one of these fields due to its high embedding capacity in 3D meshes, which can be excellent data carriers. Table. 4 provides more details on 3D datasets.

IV. STEGANALYSIS

Most steganalysis techniques have been formulated as a binary classification problem. Rich model-based steganalysis is one of these methods that achieves better detection accuracy than most other steganalysis algorithms. The method first extracts various handcrafted features from the filtered digital images in the training phase. Then, an ensemble classifier is trained to distinguish cover images from stego images. The trained classifier is used in the testing phase to determine whether a new input image includes concealed data. In steganalysis using classical machine learning, the features are extracted by handcrafted methods and are separated from the

TABLE 3. Commonly used 2D grayscale and colored datasets in steganalysis.

Dataset	Number of Images	Format	Image Size	Steganographic Scheme	Year
STEGRT1 [67]	8,000	BITMAP, JPEG	Different sizes	For JPEG: RLSBR [69], F5 [70], CE [71], OutGuess 0.2 [72], nsF5 [66], J-UNIWARD [29], and SI-UNIWARD [29] For BITMAP: RLSBR [69], LSBM [73], PVD [74], HUGO [27], WOW [30], HILL [31], S-UNIWARD [29], and MiPOD [32]	2020
LSSD [68]	2 million	JPEG	256×256	J-UNIWARD [29]	2020

TABLE 4. Commonly used 3D datasets in steganalysis.

Dataset	Number of Objects	Website
Princeton Segmentation Benchmark (PSB) [75]	354	http://segeval.cs.princeton.edu/
Princeton ModelNet (PMN) [76]	12,311	http://modelnet.cs.princeton.edu/
The Stanford 3D Scanning Repository [77]	9	http://graphics.stanford.edu/data/3Dscanrep/

classification stage. Therefore, the accuracy of the classifier relies on the effectiveness of the feature extraction method. These extracted features are fed forward to the classification stage. In contrast, in deep learning, the feature extraction stage is blended with the classification stage, and the decision of the classifier is used to update the extracted features. Many feature extraction techniques are used in steganalysis in machine learning. However, the large number of features in images causes the curse of dimensionality (CoD) and time complexity, especially when working on universal steganalysis. Extracting the discriminative features can help to improve the steganalysis accuracy. Previous studies suggested some algorithms to reduce the dimensionality of the data. Applying feature subset selection in the context of steganalysis presents many advantages as follows:

1) The accuracy of image steganalysis methods relies on sensitive features that can detect the presence of hidden messages in all types of steganography methods, and it does not rely on a large dimensional feature set.
2) By selecting the vital features, the redundant features are removed, and the discriminant features are preserved to train the classifier.
3) The computation time complexity is reduced for the feature extraction stage and training of the classifier. This will help to detect the hidden messages in different real-time applications where security is important.

A new method for processing features in two phases of optimization was proposed in [78]. The first optimization model is the eigenvalue of the scatter matrix within a class. The second phase of optimization is the employment of the random subspace Fisher linear discriminant (FLD). Kulkarni and Gorkar [79] investigated the presence of hidden malware in images. They used PCA based on eigenvalues to reduce the number of dimensions and keep all the vital features needed for the classifier. A universal image steganalysis technique focusing on feature selection was presented by Desai and Patel [80]. This method is a feature grouping based on PCA. Desai *et al.* computed the eigenvalue of the covariance matrix and then clustered the feature using the K-means method.

Some techniques can be used to accelerate feature extraction. Li *et al.* [81] found that the cost of the classical divide-and-conquer method depends on the updating of singular vectors, which includes two matrix multiplications. As a result, they concluded that the singular vector matrices of a broken matrix are Cauchy-like matrices and have off-diagonal and low-rank properties, so they can be estimated by hierarchically semiseparable (HSS) matrices. They introduced an accelerated DC algorithm where a structured low-rank estimation method is used. Their study showed that ADC can be three times faster than DC. On the other hand, Liao *et al.* [82] proposed a parallel structured divide-and-conquer aiming to reduce the computational cost. Their method builds the local matrices by employing Cauchy-like matrix generators without any communication and then reduces the computation costs by utilizing a structured low-rank approximation method.

These ADC methods showed that the computational cost of the methods massively decreased and will help significantly in steganalysis problems.

In the following sections, we present the main contribution of this survey, which is to highlight the works that have been performed using classical machine learning and deep learning techniques in the spatial and transform domains in the image steganalysis field.

A. TWO DIMENSIONAL IMAGE STEGANALYSIS METHODS BASED ON CLASSICAL MACHINE LEARNING TECHNIQUES

Different methods for 2D image steganalysis using machine learning techniques have been proposed. These methods use two phases to solve the steganalysis problem. The first phase is a handcrafted feature extraction, which has the capability of modeling the embedding distortions in the image using any steganographic algorithm. The second phase is the classification process that uses an integrated classifier for feature training. Different classifiers can be used for image steganalysis, such as the SVM and ensemble classifiers. The following sections discuss the steganalysis methods in the spatial and transform domains.

1) SPATIAL DOMAIN STEGANALYSIS

The spatial domain is used for ease of implementation and its high capacity for hidden information. Some methods in the spatial domain include HUGO [27], pixel-pair matching

scheme [83], [84], MiPOD [32], HILL [31], Gibbs construction [85], spatial rich models (SRM) [86], and WOW [30].

A method for detecting steganographic least significant bit matching (LSBM) was presented by Pevny *et al.* [9]. This method identifies deviations caused by steganographic embedding by modeling the differences between adjacent pixels in the images. A filter is used in the steganalysis to suppress the image content while exposing the stego noise. First- and second-order Markov chains are used to model the dependencies between the neighboring pixels in the filtered image. Then, the sample transition probability matrix is used to obtain a feature vector in a machine learning-based steganalyzer. Although the presented feature set was designed to detect spatial domain steganography, it was also able to detect algorithms that are hidden in the block DCT domain.

The SRM proposed by Fridrich and Kodovsky [86] is a widely used, state-of-the-art image steganalyzer. SRM extracts residual features by applying nonlinear and linear high-pass filters. The model detects stego images by acquiring the noise pattern's discontinuity in adjacent pixels in the tampered and nontampered areas. The authors proved that a medium-dimensional feature fed into a Gaussian SVM and a high-dimensional feature fed into an ensemble classifier can improve the detection accuracy for all the tested methods.

Veena and Arivazhagan [87] proposed a universal quantitative steganalyzer using reduced instances and features in which both the local and global features are considered for feature space. The global features are co-occurrence features from the Markov model, while the local features constitute the local filter pattern (LFP) [88]. The extracted features are concatenated using the greedy randomized adaptive search procedure (GRASP) [89]. Then, discretized all condensed nearest neighbor (D-AllCNN) is applied for instance reduction, and RFE is applied to reduce the feature dimensionality based on the divide and conquer principle. The AdaBoost estimator with regression trees is used to predict the payload in the stego image. The proposed blind quantitative steganalyzer is suitable for spatial LSB-based methods and can be used to improve the existing multimodel steganalytic features. As steganography algorithms negatively affect the correlations among gradient amplitudes of color channels, Kang *et al.* [90] presented a steganalysis method using channel gradient amplitude correlation for color images. The extracted features are the cooccurrence matrix from the gradient amplitude residuals that describe the correlation of the different color channels and then these features are combined with the existing features as in [91] and [92] for color image steganalysis. The proposed method was tested on the BOSSbase dataset. The dimension of the features is 5,404, which consumes a great deal of space and time when features are extracted and saved. Due to the high dimensionality of the steganalysis features, the proposed algorithm uses an ensemble classifier, which is a common learning technique for image steganalysis. To improve the detection of perturbations of the local patterns in stego images, features are used in texture classification tasks such as LBP. LBP features can characterize local structure changes, and they seem to be promising. LBP can effectively summarize the local structures of an image by comparing pixels with their neighbors. Inspired by this idea, Gui *et al.* [93] proposed extracting multiscale rotation invariant LBPs from smooth pixels as unique textural features, which are then fed into the linear SVM. The experimental results showed that the method performed well in detecting stego images and had a high accuracy.

Liu *et al.* [94] presented a blind image steganalysis method based on a nature-inspired feature selection method. The features are extracted for image steganalysis using SPAM. Then, the ideal feature subset is chosen from the original features using the binary bat method (BBM) [95]. The classifiers used to verify the proposed method are KNN, RF, AdaBoost, DCA, NB and SVM. The proposed method was tested using the BOSSBase v1.01 dataset, and the accuracy was 68.08% with the SVM classifier.

2) TRANSFORM DOMAIN STEGANALYSIS

The transform domain embeds the hidden messages in the coefficients of the cover image. Therefore, the transform domain has an advantage over the spatial domain, where the hidden messages in the transform domain are not affected by image processing, compression, or cropping. Methods in the transform domain include UED [37], UERD for JPEG steganography [62], statistical features of contourlet transform [96] and block-based image steganalysis based on DCT and Markov features [97]. These steganography methods leave minimal traces of hidden data, so it is necessary to extract independent features from the image to proceed to the next phase. Therefore, efficient features for the steganalysis process include the Markov transition probabilities of pixels, histogram of residuals, cooccurrence matrices, LBP operators, etc. The next phase is the classification process in which integrated classifiers for feature training are used. Classifiers that can be used for image steganalysis include the SVM and ensemble classifiers.

Liu *et al.* [98] presented a new method based on feature mining, the DCT domain and SVM for JPEG image steganalysis. They extracted features using both the intr-block and intera-block neighboring joint density from the DCT coefficient; then, they fed these features into SVM for detection. To predict the hidden amount in JPEG steganography, the authors applied a neural-fuzzy inference system. Their experimental results showed that their method performed better than the well-known Markov process-based method.

Holub and Fridrich proposed a novel feature set for JPEG steganalysis called discrete cosine transform residual (DCTR) [99]. These features are low in complexity and small in dimension, and they are created as histograms of the residuals achieved using 64 DCT bases. The authors used the Fisher Linear Discriminant (FLD) [100] ensemble as a binary classifier. The results show that DCTR achieved competitive detections over many JPEG methods.

Song *et al.* [101] proposed a steganalysis method by applying 2D Gabor filters for the feature extraction phase to detect

the embedded changes constrained in the complicated texture regions of the JPEG images. The results showed that the proposed features for image steganalysis obtained a competitive performance when compared to the achieved steganalysis features UED [37], J-UNIWARD [29] and SI-UNIWARD [29].

Shankar and Azhakath [102] proposed a blind feature-based image steganalysis for the JPEG file format. A total of 274 image features were extracted using a DCT from the first-order (dual, global, and individual histograms), second-order (co-occurrence, variance, and blockiness), and special Markov features. The classifiers were SVM and SVM-particle swarm optimizations (SVMPSO). The classifiers were adapted with 10% embedding and 10-fold cross-validation. The kernels used in the classification process were linear, multiquadratic, Epanechnikov, radial, polynomial and ANOVA. The two image datasets used for the suggested blind steganalysis were INIRA holidays [103] and UCID [104]. It was shown that the PSO classifiers achieved better performance than SVM for all the kernels and samples.

Liu *et al.* [105] presented a framework based on fusing SVM classifiers; it consists of three stages, training the subclassifier, training the fusing classifier and testing the fusing classification. The author used rich model features proposed in [106] that are divided into different groups based on the correlation features. The fusion classifier is able to learn the correlation of the detection results for the subclassifiers, and the accuracy is enhanced when the classifiers are increased. Lu *et al.* [107] presented an improvement framework for steganalysis based on feature selection and preclassification. The features are extracted using a dependency analysis of the adjacent image data. The K-means algorithms are applied to preclassify the images that have various content and texture complexities from the image dataset. Then, the optimal features from each cluster were chosen for a final decision aimed at improving the overall performance of the steganalysis. Shankar and Azhakath [108] explored four feature extractions for the steganalysis, which were first order, extended DCT, second order, and Markov features. They used the LSBM method [73] and F5 [70] for the spatial domain and transform domain, respectively. They employed six different kernels and four kinds of SVM samplings with cross-validation. Their study concluded that the transform domain provided better accuracy of classification than the spatial domain. Table. 5 provides more details on the work performed in the spatial and transform domains.

B. THREE DIMENSIONAL IMAGE STEGANALYSIS METHODS BASED ON CLASSICAL MACHINE LEARNING TECHNIQUES

In this section, we describe some steganalysis algorithms that have been used on 3D images. As the goal of 3D steganalysis is to find the concealed data in 3D images, it is a challenging problem compared to 2D image steganalysis because 3D images are 3D complex objects that have an arbitrary topology and irregular geometry.

Yang and Ivrissimtzis [115] presented the first 3D steganalysis features (YANG208) for detecting hidden messages in triangle meshes. For each mesh, they calculated the characteristic feature vector that captured the geometric information from its Cartesian and Laplacian coordinates. They then applied a calibration technique on the extracted feature vector by computing the difference between the mesh and the reference mesh to extract the discriminative features. The extracted features were then fed into the supervised learning method based on quadratic discriminant analysis (QDA). The method was tested on six well-known steganographic frameworks and showed satisfactory accuracy rates.

Li and Bors in [116] proposed a method (LFS52) that extracted a 52-D local feature vector for the 3D steganalysis problem. The 52-D feature vector combined three components, a 40-D feature vector consisting of the most effective features in YANG208 [115], a 4-D vertex normal feature vector and an 8-D local shape curvature feature vector. The combined features were used as input to the FLD ensemble and a quadratic classifier to distinguish the 3D stego-objects from the cover objects. The proposed method was tested on the PSB dataset, where stego objects were created using two different steganography techniques that hide messages in the 3D objects. The results showed that the method provides better performance for the 3D steganalysis process, where local shape curvature features and vertex normal features have better discriminability.

Li and Bors [117] proposed the Robustness and Relevance based Feature Selection (RRFS) algorithm as a solution for the cover-source mismatch problem in 3D steganalysis. A feature set (LAY252) is extracted using a combination of LFS52 features [116] and YANG208 features [115]. The proposed selection algorithm selects the features based on their robustness and correlation. The selected features are fed into the FLD ensemble. The proposed algorithm chooses better features than other algorithms. However, this algorithm is limited to a set of transformations in the cover-source mismatch problem.

Kim *et al.* [118] proposed the local feature set (LFS64). They used mean, total curvature, and edge normal in addition to features presented in [115] and [116], and they mapped the features using a homogeneous kernel map to help the FLD ensemble classifier detect setgo meshes. The proposed method outperformed LFS52 [116].

Li and Bors [119] proposed a method (LFS76) extended to LFS52 features [116] to identify the small variances between the cover and stego 3D graphical objects. The proposed method extracts and combines various 3D features, such as vertex normal, local curvature, and a local geometric representation of the vertex in spherical coordinates. The statistics of the sets of extracted features with the 76-D feature vector are fed into the SVM classifier, FLD ensemble, and QDA. The authors used the PSB dataset that contains 354 3D mesh cover objects. Stego objects were created using three different steganographic methods for information hiding. The experimental results showed that the FLD ensemble provided the

TABLE 5. Summary of work performed on 2D images using classical machine learning.

Paper	Method	Domain	Steganography Algorithm	Dataset
Fridrich *et al.* [86]	SRM extracts residual features. Medium - dimensional feature fed into a Gaussian SVM while a high-dimensional feature fed into the ensemble classifier.	Spatial	HUGO [27], edge-adaptive [109], and LSBM [73]	BOSSbase 0.92
Pevny *et al.* [9]	First and second order Markov chains are used. Then, the sample transition probability matrix is used for the feature vector and it is fed into SVM.	Spatial	LSBM [73]	BOWS2, NRCS, and JOINT
Veena *et al.* [87]	Global and local features are combined using GRASP. D-AllCNN is applied for instance reduction while RFE is applied to reduce features dimensionality. AdaBoost estimator with regression trees is used for classification.	Spatial	LSBR, LSBMR [110], LSBM [73], Modulo LSB (LSBRmod5) [111], and two bit LSBR (LSBR2) [112]	BOSSbase 1.01
Kang *et al.* [90]	The extracted features are the cooccurrence matrix from the gradient amplitude residuals. The features are fed into ensemble classifier.	Spatial	WOW [30], and S-UNIWARD [29]	BOSSbase 1.01
Gui [93]	Extracting multiscale rotation invariant LBPs as unique textural features that are fed into SVM.	Spatial	LSBM [73]	BOSSbase 1.01
Liu *et al.* [94]	The features are extracted using SPAM. Then, the ideal feature subset is chosen using BBM. The classifiers used are KNN, RF, AdaBoost, DCA, NB and SVM.	Spatial	HILL [31], WOW [30], and HUGO [27]	BOSSbase 1.01
Liu *et al.* [98]	The features are extracted using both intr-block and intera-block neighboring joint density from the DCT coefficient. The features are fed into SVM.	Transform	CryptoBola, JPHS, Steghide [113], F5 [70], MB1 [114], and MB2 [114]	Their raw data
Holub *et al.* [99]	Features are created as histograms of residuals achieved using 64 DCT bases. FLD ensemble is used for classification.	Transform	J-UNIWARD [29]	BOSSbase 1.01
Song *et al.* [101]	2D Gabor filters are applied to extract features and fed into ensemble classifier.	Transform	J-UNIWARD [29]	BOSSBase 1.01
Shankar and Azhakath [102]	Features were extracted using DCT from first order, second order and the special Markov feature. The classifiers were SVM and SVM-PSO.	Transform	LSBM [73], PVD [74], and LSBR	NRIA holiday [103] and UCID [104]
Liu *et al.* [105]	The rich model features are divided into different groups based on the correlation features. Fusing SVM classifiers are used for classification.	Transform	LSBR, J-UNIWARD [29], and LSBRmod5 [111]	BOSSBase 1.01
Lu *et al.* [107]	K-means were applied to preclassify the images that have various content and texture complexities.	Transform and Spatial	J-UNIWARD [29], UED [37] , nsf5 [66], S-UNIWARD [29], and MiPOD [32]	BOSSbase 1.01
Shankar and Azhakath [108]	The four extracted features are: first order, extended DCT, second order, and Markov features. Six different kernels and four kinds of samplings of SVM are used for classification.	Transform and Spatial	LSBM [73], and F5 [70]	NRIA holiday [103] and UCID [104]

best results for the steganalysis process when the mean-based watermarking steganographic method [52] was used to identify the information embedded in 3D objects.

Li *et al.* [120] proposed a 3D feature extraction technique that uses edge vectors to capture the local features, resulting in a 124-D feature vector (LFS124). The absolute differences between the edge lengths of the 3D components of the vector were computed in the Cartesian coordinate system. Then, the difference norm between the two vectors was computed, and two different features derived from the absolute differences and the angle between them were defined. Finally, six features were computed in the same way in the Laplacian coordinate system, all of which together formed 12 features. Then, the newly extracted feature set was combined with the existing feature set of LFS76 to obtain the 124-D feature vector. These features were fed into an FLD ensemble. The proposed method was tested on 354 cover 3D mesh objects from the PSB dataset. The 3D stego meshes were produced by six 3D information hiding techniques. The experiment showed that the proposed method is efficient in implementation and concluded that the edge vector plays a significant role in steganalysis.

Zhou *et al.* [45] proposed a specific steganalysis method using the PCA transform-targeted feature to differentiate between stego and cover 3D mesh objects. The transformation matrix of a stego mesh is close to the identity matrix after a PCA transform, while the transformation matrix of a cover mesh is far from the identity matrix on most occasions. The one-dimensional feature is defined by the norm between the two transformation matrices. This method

was tested on the PMS and PMN datasets. The proposed steganalysis method was only efficient for steganographic methods based on the PCA transform.

Zhou *et al.* [121] presented a 3D steganalytic scheme (NVT+) using a tensor voting model that collects the local shape context to distinguish a stego 3D mesh object from the cover object. First, three normal voting tensors with different neighbor definitions were performed. Second, three eigenvalues were computed from every tensor, where the absolute value of the difference between eigenvalues was regarded as a feature. Three tensor models that each extract three eigenvalue differences produced nine features. Third, several statistical moments of features processed by means of nonlinear mapping were extracted to form 36 features. The 36 obtained features were combined with the features of the LFS64 method in [118] to obtain a 100-D feature vector. The combined feature set was fed into the FLD ensemble for classification. The proposed method was examined on the PMS and PMN datasets. The experiment showed that the proposed method enhances the detection performance. However, the time taken and the complexity of this method are very high due to the calculation of each feature for the adjacent face.

Li and Bors [122] proposed WFS228, a novel set of 228-D steganalysis features extracted using multiresolution 3D wavelet analysis [123]. The features are extracted from transformations between an input mesh and its corresponding higher and lower resolutions. For an input mesh, its corresponding higher and lower graph resolutions are computed using the 3D wavelet algorithm. The method was trained using the FLD ensemble, and the experiments showed that the 3D wavelet feature provided the best performance for the steganalysis task. Table. 6 provides more details about the work performed in 3D image steganalysis.

C. TWO DIMENSIONAL IMAGE STEGANALYSIS METHODS BASED ON DEEP LEARNING

Over the last few years, deep learning has been widely used in steganalysis to extract appropriate features for classification. Convolutional neural networks (CNNs) have enhanced the performance of steganalysis; however, the memory space and the computational complexity cost of the models are still obstacles due to the large amount of training data. In this section, we present deep learning models that aim to reduce the learning cost by extracting the key features.

Ghosh *et al.* [125] presented an ANN model, a new hybrid ANN deep neural network based on eigenvalues (more specifically PCA) and Haralick features. They computed the co-occurrence matrix of the grayscale input image for four pixel pair directions and then computed the average. Then, two-dimensional reduction is applied: PCA and Haralick. Their method was promising and achieved enhanced accuracy. Zang *et al.* [126] proposed extracting key texture features by employing a learnable local histogram layer based on multiquadratic kernel modeling. The histogram layer used two convolutions to learn the center and width of the bin. They used an RBF neural network to update the bin center and width of the model, and eigenvalues were used to find the minimum and maximum values of the RBF. The method showed significant improvement in texture classification. Abazar *et al.* [127] presented a novel framework to reduce the learning cost by using a divide and conquer technique. The dataset is split into five disjoint clusters by employing k-means. Each cluster is fed into a distinct CNN. The networks are combined leveraging a fast weighting process. The proposed model is able to reduce the size of the training data for each model. The experimental results showed that the proposed framework reduces the time complexity while maintaining the accuracy.

The following sections provide a summary of the state-of-the-art works that have been performed in 2D image steganalysis using deep learning techniques in the spatial and transform domains.

1) SPATIAL DOMAIN STEGANALYSIS

As we mentioned previously, in the spatial domain steganography, the payload bits are hidden in a cover image by changing the pixel intensity values directly in the spatial domain. Knowing this, researchers have begun to take advantage of applying deep learning for spatial domain steganalysis. The first attempt to use an unsupervised deep learning method for steganalysis was carried out by Tan and Li [128]. The authors used stacked convolutional autoencoders (SCAEs) [129]. The weights of the kernels and filters in the CNN were randomly initialized. The authors believed that a well-trained CNN must perform comparably to the well-known and successful SRM. They used a nine-layer, three-stage CNN based on a blind steganalyzer.

Qian *et al.* [21] were the first to propose using supervised learning with CNNs for steganalysis. Their network consists of three steps, a high-pass filter used as a preprocessing layer, a convolutional layer for feature extraction and then a fully connected layer for classification. The high-pass filter layer is used because the stego has a weaker signal than the content of the image. This model achieved reasonable results compared to traditional models using handcrafted features. Wu *et al.* [130] proposed a new feature extraction framework that can learn joint features from input images and their corresponding residual images. Their feature fusion process in CNN is completely unsupervised. To minimize data dimensions, the method chooses feature maps from the middle three hidden layers and concatenates them into a 1D vector that is passed into the fully connected layers to obtain the classification result. The aim is to decrease the negative impact of the high-pass filter to guarantee that the network remains convergent.

Rezaei *et al.* [67] tested more than 40 CNN architectures and found that the best shape consists of two convolutional layers followed by three fully connected layers. The input image of the CNN is filtered first by high pass, as is done in the work of Qian *et al.* [21]. The CNN is evaluated on two scenarios, the first of which is a clairvoyant scenario in

TABLE 6. Summary of works performed in 3D Steganalysis.

Paper	Method	Data Hiding Algorithm	Dataset
Yang and Ivrissimtzis [115]	Extracted characteristic feature vector captures geometric information, then calibration technique was applied on the extracted features, and fed into QDA.	High-capacity steganography [41], the Laplacian coordinate-based watermarking method [42], mean-based watermarking [52], the variance-based watermarking [52], the frequency-based watermarking [53], and LSB modification [124]	PSB
Li and Bors [116]	The feature vector combined three components: a 40-D feature vector that consists of the most effective features in YANG208 [115], a 4-D vertex normal feature vector, and an 8D local shape curvature feature vector. FLD ensemble and quadratic classifier were used for classification purposes.	Mean-based watermarking [52], and high-capacity steganography [41]	PSB
Li and Bors [117]	RRFS selects discriminated features from LAY252 features set and the selected features are fed into the FLD ensembles classifier.	High-capacity steganography [41]	PSB
Kim *et al.* [118]	LFS52 feature set is combined with edge normal vector, mean, and total curvature. A homogeneous kernel map is used with the FLD ensembles classifier.	High-capacity steganography [41], and mean-based watermarking [52]	PSB
Li and Bors [119]	Features using the spherical coordinates are combined with LFS52 features and fed into the SVM classifier, FLD ensemble, and QDA.	High-capacity steganography [41], mean-based watermarking [52], and steganalysis-resistant watermarking [54]	PSB
Li *et al.* [120]	Features using edge vectors were combined with LFS76 features and fed into the FLD ensembles classifier.	High-capacity steganography [41], 3D wavelet-based fragile watermarking [46], 3D wavelet-based high capacity watermarking [46], mean-based watermarking [52], the variance-based watermarking [52], and steganalysis-resistant watermarking [54].	PSB
Zhou *et al.* [45]	PCA transform-targeted feature.	Adaptive-steganography-based method [45]	PSB, and PMN
Zhou *et al.* [121]	Features based on tensor voting model were combined with LFS64 features and fed into the FLD ensembles classifier.	High-capacity steganography [41], adaptive-steganography-based method [45], and steganography using a shifting strategy and a truncated space [47]	PSB and PMN
Li and Bors [122]	WFS228 features extracted using multiresolution 3D wavelet analysis and fed into the FLD ensemble.	High-capacity steganography [41], 3D wavelet-based fragile watermarking [46], 3D wavelet-based high capacity watermarking [46], mean-based watermarking [52], The variance-based watermarking [52], and steganalysis-resistant watermarking [54]	PSB

which it is assumed that the same embedding key is applied on different images. The authors compared this scenario to the ensemble classifier with SRM features and found that CNN reduced classification errors by three times. In the second scenario, a cover-source mismatch was assumed in which the source model used in steganography is different from the source model assumed for steganalysis. The classification errors decreased compared to the rich models and ensemble classifier.

Xu *et al.* [131] proposed using a CNN with statistical modeling to avoid network convergence. They employed a high-pass filter as a layer to gain the noise residuals of the original images and then fed them into five convolution and pooling layers. The 128-dimensional features are fed into the fully connected layer and then the softmax layer to classify the input. The main contribution of this technique is that it uses the absolute layer (ABS) after the convolution layer to obtain positive values. The output is then fed into the batch normalization layer to guarantee that the network did not become stuck in local minima. The hyperbolic tangent (tanH) nonlinear activation function was used in the first group of convolution layers, and rectified linear units (ReLUs) were used in the remaining layers. The authors trained their CNN model using minibatch gradient descent, and the results outperformed the traditional SRM ensemble classifier.

Ye *et al.* [132] introduced YeNet, which has a new truncated linear unit (TLU), in the CNN steganalysis model. The network contains 10 convolutional layers, and 30 high-pass kernels were initiated using SRM and used as a preprocessing layer. In the first convolution layer, the authors used TLU, and in the remainder of the layers, they employed the ReLU activation function. The output from 144-dimensional feature vectors was fed into one fully connected layer, followed by a softmax layer. YeNet achieved lower detection error rates in comparison with the SRM and maxSRMd2 steganalyzers.

Yedroudj *et al.* [133] presented a CNN model by incorporating one preprocessing layer consisting of 30 high-pass layers from SRM kernels followed by five convolutional layers and, finally, one softmax layer in the spatial domain. Their CNN model is similar to Xu's net [131] and Ye's Net [132]. YedroudjNet employed batch normalization and the ABS layers as Xu net [131], but they used shallower convolution layers compared to Ye's Net [132]. Finally, Yedroudj *et al.* used three fully connected layers.

In summary, studies performed in deep learning have concluded that taking into account knowledge of the domain in steganalysis can improve the performance of CNNs. As CNNs adopt a feature extraction step, domain knowledge should be taken into account when designing network architectures.

2) TRANSFORM DOMAIN STEGANALYSIS

Steganography approaches on the JPEG domain work in the transform domain by changing the coefficients obtained after applying DCT. In the past, most JPEG steganalysis techniques extracted features from decompressed images. However, researchers are now motivated to study JPEG steganographic algorithms using CNNs.

Xu [134] transformed the input JPEG images into the spatial domain and fed them into a set of specified DCT filters of sizes of $2\times2, 3\times3, 4\times4, 5\times5$, and 8×8 as a preprocessing step. The best result is obtained when a 4×4 filter is used. These features were used in a CNN architecture composed of 20 convolution layers with batch normalization and ReLU function layers. The output – 384 feature vectors – was fed into a fully connected layer followed by a softmax layer. The results showed that the proposed CNN network had decreasing classification error.

Chen *et al.* [135] developed a novel JPEG-phase awareness feature with two CNN architectures to increase detection accuracy. The JPEG phase is a statistical property collected using an 8×8 pixel neighborhood separately by obtaining noise residuals. Their CNN model relied on Xu's model [131]; however, they incorporated phase awareness into XuNet [134] and disabled the pooling layer from the first two layers. Each feature map that gains from the second layer is subsampled on 64 sublattices and then used in the phase-split layer. Depending on the phase splits, they implemented two networks called the PNet and the VNet. In PNet, the output feature maps with a size of 16×16 are split into 64 groups, resulting in 16 feature maps. Thus, each group has its specific processing layers. The resulting feature maps are then concatenated to form an 8,192 D feature vector. (This approach is not performed in VNet). The final output vector has a dimension of 512. The experiment showed that PNet outperformed VNet.

Zeng *et al.* [136] employed a hybrid deep learning framework for state-of-the-art JPEG steganography approaches, J-UNIWARD [29], UED [39] and UERD [67], and used handcrafted quantization and truncation (Q & T) phases of rich models with CNN. The CNN model has two stages. In the first stage, a 25, 5×5 DCT base was used to compute 25 residual maps from uncompressed and nontruncated JPEG images. These maps were then handed over to three Q & T phases. In the second stage, the three Q & T phase outputs were fed into three independent subCNNs. The output feature maps from each subCNN were flattened, and a 512-feature vector was obtained for each subCNN. The final output feature length of 1,536 was fed into four fully connected layers. The CNN model was trained using stochastic gradient descent (SGD).

Yousfi *et al.* [137] won the ALASKA steganalysis challenge in 2019 by using SRNet [138] to train different combinations of three input channels, luminance Y and chrominances Cr and Cb. SRNet used residual skip connections, and the filter size was 3×3. All the convolutional layers were followed by a batch normalization and the ReLU activation function. The first eight convolutional layers did not incorporate the pooling layer since average pooling is assumed to be a low-pass filter, while steganalysis is concerned with high-pass content where stego data are found. The output of these convolutions was fed to a fully connected layer that produced two outputs and was fed to a binary classifier.

Inspired by the idea of using the transfer learning of pretraining neural networks on unrelated tasks and refining steganalysis, Yousfi *et al.* [139] investigated pretrained deep learning networks such as EfficientNet [140], MixNet [141] and ResNet [142] for steganalysis. They concluded that removing pooling and stride in the first layers allowed for better performance. Xiancheng Wu *et al.* [143] explored the effects of applying compression in eight-bit calculations and floating-point quantizations to XuNet. The model achieved higher accuracy than Xu's model. Their results showed that the two CNN models based on quantization schemes are useful in steganalysis. Table. 7 provides details of the common techniques developed for the spatial and transform domains in steganalysis.

In 2D image steganalysis using classical machine learning, it seems that SVM and SRM are the most popular binary classifiers, while FLD is the most popular ensemble classifier for 3D image steganalysis. In deep learning, 2D CNN architectures are commonly used by researchers to implement steganalysis models for image steganalysis. It is known that 3D meshes have a higher embedding capacity than 2D images. However, many steganalysis studies target 2D images. Therefore, it is important to investigate the possibility of detecting 3D mesh steganography using deep learning techniques.

V. OPEN CHALLENGES

While steganalysis has received considerable attention in the past decade, some challenges remain unsolved. First, the different CNN models presented in this survey are designed to be suitable for specific datasets. To date, there is no generalized CNN model that can detect hidden messages in unseen data. Second, none of the currently available deep learning models take into account the use of generative adversarial networks (GANs). It is worth investigating whether the generator of GAN models can learn from stego and cover images and generate reasonable outputs to distinguish between the two. This will help to simplify the task of detecting steganography. Third, as discussed in Section III, many datasets are available with different specifications, such as the data domain. However, the current steganalysis deep learning models use specific datasets. Therefore, there is a

TABLE 7. Summary of work performed on 2D images using deep learning.

Paper	Method	Domain	Steganography Algorithm	Dataset
Tan and Li [128]	2 Conv, 3 max pooling, 2 FC, FC(softmax).	Spatial	HUGO [27]	BOSSbase 1.01
Qian *et al.* [21]	Preprocessing HPF, 5 conv, 3 avg-pooling, 2 FC, FC(softmax).	Spatial	HUGO [27], WOW [30], and S-UNIWARD [29]	BOSSbase 1.01
Wu *et al.* [130]	Preprocessing (HPF), 3 conv, 2 FC, FC (softmax).	Spatial	S-UNIWARD [29]	BOSSbase 1.01
Pibre *et al.* [62]	Preprocessing (HPF), 2 conv, 3 FC, FC (softmax).	Spatial	S-UNIWARD [29]	BOSSbase, and LIRMMBase 1.01
Xu *et al.* [131]	Preprocessing (HPF),5 Conv with ABS, BN, pooling, (tanH) the first group of convs and ReLU in the remaining layers,1 FC, FC (softmax).	Spatial	S-UNIWARD [29] and HILL [31]	BOSSbase 1.01
Ye *et al.* [132]	Preprocessing (30 HPF), 1 conv with TLU, 2 conv with ReLU,7 conv with ReLU and avg pooling, one FC, FC (softmax).	Spatial	S-UNIWARD [29], WOW [30] and HILL [31]	BOSSbase, and LIRMMBase 1.01
Yedroudj [133]	Preprocessing (HPF), 5 conv, 5 ABS, BN, 2 FC, FC (softmax).	Spatial	S-UNIWARD [29] and WOW [30]	BOSSbase 1.01
Guanshuo Xu [134]	Preprocessing (HPF), 1 conv with ABS, BN, TanH,avg pooling; 1 Conv with BN, TanH, avg pooling; 3 Conv with BN, ReLU,avg pooling; 1 FC, FC (softmax).	Transform	J-UNIWARD [29]	BOSSbase 1.01
Chen *et al.* [135]	PNet: Conv with ABS, BN, TanH; conv with BN, TanH, 8×8 phase split. 6 Conv with BN, ReLU, avg pooling, 1 FC, FC (softmax). VNet: 2 FC with ABS, BN, TanH; FC with BN,TanH, 8×8 phase split; FC with BN, ReLU, avg pooling; FC with BN, ReLU, avg pooling, 1 FC, FC (softmax).	Transform	J-UNIWARD [29] and UED-JC	BOWS2
Zeng *et al.* [136]	Stage 1: 25 DCT basis of size 5×5 and the output is 25 residual maps followed by three Q T phases. Stage 2: three independent subCNNs, the output of which the three subCNN is flatten; 4 FCs with softmax.	Transform	J-UNIWARD [29], UED [37] and UERD [62]	ImageNet
Yousfi *et al.* [137]	The pretrained models: MixNet [141], EfficientNet [140], and ResNet [142].	Transform	UERD [62] and J-UNIWARD [29]	ALASKA I
Yousfi *et al.* [139]	SRNet [138] with residual skip connections and the filters size of 3×3. All conv layers followed by BN, then ReLU. The first eight conv layers without pooling, 1 FC, FC (softmax).	Transform	J-UNIWARD [29], UED [37], and EBS [144]	ALASKA
Wu *et al.* [143]	8-bit calculation and floating-point quantization to XuNet.	Transform	J-UNIWARD [29]	ImageNet

need for an explicit steganalysis deep learning model that can learn from different datasets with different specifications to use the available data efficiently. Fourth, there are some fundamental open questions regarding the use of 3D datasets in steganalysis with deep learning. Deep learning methods have shown promising results in 2D image steganalysis, but there are two potentially challenging questions: Do 3D steganalytic methods based on deep learning provide better performance? Are there enough 3D data that can be used to train the steganalysis CNN models? Finally, steganography is designed to pass hidden messages through media such as the internet, and the data might be exposed to manipulation during the transmission process (e.g., by rotation or corruption). Thus, an interesting direction of research would be to build deep learning models that can learn to predict the hidden messages correctly. Feature dimensionality is still a problem specially when performing real-time steganalysis. Therefore, it is necessary to find an appropriate accelerating algorithm to speed up the learning process even in deep learning methods without compromising the accuracy.

VI. CONCLUSION

In this survey, we reviewed the works that have been performed in the digital image steganalysis field. We have analyzed the steganalysis methods available for 2D and 3D images. A decade ago, studies on traditional steganalysis methods focused on classical supervised machine learning, such as SVM and SRM. Recently, with the success of CNNs, different architectures have been developed to detect steganographic messages in the spatial and transform domains. These CNNs have achieved prominent performances compared to the classical machine learning methods in the field of steganalysis. Detecting stego images from CNN models is still in the early stages, and the deep learning models need to be robust against steganographic algorithms. Further research needs to explore how well the generative adversarial network architecture helps develop steganalysis algorithms for images in the wild.

REFERENCES

[1] D. L. Parnas, "On the criteria to be used in decomposing systems into modules," *Commun. ACM*, vol. 15, no. 12, pp. 1053–1058, Dec. 1972.

[2] T. Moerland, "Steganography and steganalysis," May 15, 2003. [Online]. Available: https://www.liacs.nl/home/tmoerlan/privtech.pdf

[3] D. Artz, "Digital steganography: Hiding data within data," *IEEE Internet Comput.*, vol. 5, no. 3, pp. 75–80, May 2001.

[4] N. F. Johnson, Z. Duric, and S. Jajodia, *Information Hiding: Steganography and Watermarking-Attacks and Countermeasures: Steganography and Watermarking: Attacks and Countermeasures*, vol. 1. New York, NY, USA: Springer, 2001.

[5] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding—A survey," *Proc. IEEE*, vol. 87, no. 7, pp. 1062–1078, Jul. 1999.

[6] S. Dumitrescu, X. Wu, and Z. Wang, "Detection of LSB steganography via sample pair analysis," in *Proc. Int. Workshop Inf. Hiding*. Berlin, Germany: Springer, 2002, pp. 355–372.

[7] N. F. Johnson and S. Jajodia, "Steganalysis: The investigation of hidden information," in *Proc. IEEE Inf. Technol. Conf., Inf. Environ. Future*, Sep. 1998, pp. 113–116.

[8] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.
[9] 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.
[10] N. Zaker and A. Hamzeh, "A novel steganalysis for TPVD steganographic method based on differences of pixel difference histogram," *Multimedia Tools Appl.*, vol. 58, no. 1, pp. 147–166, May 2012.
[11] I. Avcibas, N. D. Memon, and B. Sankur, "Image steganalysis with binary similarity measures," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, vol. 3, Sep. 2002, pp. 645–648.
[12] S. Dumitrescu, X. Wu, and Z. Wang, "Detection of LSB steganography via sample pair analysis," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1995–2007, Jul. 2003.
[13] G. Gul and F. Kurugollu, "SVD-based universal spatial domain image steganalysis," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 349–353, Jun. 2010.
[14] H. Farid, "Detecting hidden messages using higher-order statistical models," in *Proc. Int. Conf. Image Process.*, vol. 2, Sep. 2002, pp. 1–2.
[15] Y. Q. Shi, G. Xuan, D. Zou, J. Gao, C. Yang, Z. Zhang, P. Chai, W. Chen, and C. Chen, "Image steganalysis based on moments of characteristic functions using wavelet decomposition, prediction-error image, and neural network," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2005, pp. 269–272.
[16] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. 5th Annu. Workshop Comput. Learn. Theory (COLT)*, 1992, pp. 144–152.
[17] V. Vapnik, *The Nature of Statistical Learning Theory*, vol. 8. New York, NY, USA: Springer, 2000, pp. 1–15.
[18] I. Jolliffe, *Principal Component Analysis*. Berlin, Germany: Springer, 2011, pp. 1094–1096.
[19] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.
[20] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 129–137, Mar. 1982.
[21] Y. Qian, J. Dong, W. Wang, and T. Tan, "Deep learning for steganalysis via convolutional neural networks," *Proc. SPIE*, vol. 9409, Mar. 2015, Art. no. 94090J.
[22] H. Wang and S. Wang, "Cyber warfare: Steganography vs. steganalysis," *Commun. ACM*, vol. 47, no. 10, pp. 76–82, Oct. 2004.
[23] N. F. Johnson and S. Jajodia, "Exploring steganography: Seeing the unseen," *Computer*, vol. 31, no. 2, pp. 26–34, Feb. 1998.
[24] Y. J. Chanu, T. Tuithung, and K. M. Singh, "A short survey on image steganography and steganalysis techniques," in *Proc. 3rd Nat. Conf. Emerg. Trends Appl. Comput. Sci.*, Mar. 2012, pp. 52–55.
[25] M. Hussain, A. W. A. Wahab, Y. I. B. Idris, A. T. S. Ho, and K.-H. Jung, "Image steganography in spatial domain: A survey," *Signal Process. Image Commun.*, vol. 65, pp. 46–66, Jul. 2018.
[26] K. Ghazanfari, S. Ghaemmaghami, and S. R. Khosravi, "LSB^{++}: An improvement to LSB^{+} steganography," in *Proc. IEEE Region 10 Conf.*, Nov. 2011, pp. 364–368.
[27] T. Pevný, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in *Information Hiding*, R. Böhme, P. W. L. Fong, and R. Safavi-Naini, Eds. Berlin, Germany: Springer, 2010, pp. 161–177.
[28] P. Bas, T. Filler, and T. Pevný, "Break our steganographic system: The ins and outs of organizing boss," in *Information Hiding*, T. Filler, T. Pevný, S. Craver, and A. Ker, Eds. Berlin, Germany: Springer, 2011.
[29] 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, 2014.
[30] V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2012, pp. 234–239.
[31] B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 4206–4210.
[32] V. Sedighi, R. Cogranne, and J. Fridrich, "Content-adaptive steganography by minimizing statistical detectability," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 2, pp. 221–234, Feb. 2016.
[33] S. M. Thampi, "Information hiding techniques: A tutorial review," ISTE-STTP Netw. Secur. Cryptogr., LBSCE, Tech. Rep., 2008.
[34] Q. Cheng and T. S. Huang, "An additive approach to transform-domain information hiding and optimum detection structure," *IEEE Trans. Multimedia*, vol. 3, no. 3, pp. 273–284, Sep. 2001.
[35] A. A. Abdelwahab and L. A. Hassaan, "A discrete wavelet transform based technique for image data hiding," in *Proc. Nat. Radio Sci. Conf.*, Mar. 2008, pp. 1–9.
[36] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 920–935, Sep. 2011.
[37] L. Guo, J. Ni, and Y. Q. Shi, "Uniform embedding for efficient JPEG steganography," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 5, pp. 814–825, May 2014.
[38] L. Guo, J. Ni, and Y. Q. Shi, "An efficient JPEG steganographic scheme using uniform embedding," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2012, pp. 169–174.
[39] F. Cayre and B. Macq, "Data hiding on 3-D triangle meshes," *IEEE Trans. Signal Process.*, vol. 51, no. 4, pp. 939–949, Apr. 2003.
[40] C.-M. Wang and Y.-M. Cheng, "An efficient information hiding algorithm for polygon models," mphComput. Graph. Forum, vol. 24, pp. 591–600, Sep. 2005.
[41] M. W. Chao, C. H. Lin, C. W. Yu, and T. Y. Lee, "A high capacity 3D steganography algorithm," *IEEE Trans. Vis. Comput. Graph.*, vol. 15, no. 2, pp. 274–284, Mar. 2009.
[42] Y. Yang and I. Ivrissimtzis, "Polygonal mesh watermarking using Laplacian coordinates," *Comput. Graph. Forum*, vol. 29, pp. 1585–1593, Jul. 2010.
[43] Y.-Y. Tsai, "An adaptive steganographic algorithm for 3D polygonal models using vertex decimation," *Multimedia Tools Appl.*, vol. 69, no. 3, pp. 859–876, Apr. 2014.
[44] Y.-M. Cheng and C.-M. Wang, "A high-capacity steganographic approach for 3D polygonal meshes," *Vis. Comput.*, vol. 22, nos. 9–11, pp. 845–855, Sep. 2006.
[45] H. Zhou, K. Chen, W. Zhang, Y. Yao, and N. Yu, "Distortion design for secure adaptive 3-D mesh steganography," *IEEE Trans. Multimedia*, vol. 21, no. 6, pp. 1384–1398, Jun. 2019.
[46] K. Wang, G. Lavoue, F. Denis, and A. Baskurt, "Hierarchical watermarking of semiregular meshes based on wavelet transform," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 4, pp. 620–634, Dec. 2008.
[47] N. Li, J. Hu, R. Sun, S. Wang, and Z. Luo, "A high-capacity 3D steganography algorithm with adjustable distortion," *IEEE Access*, vol. 5, pp. 24457–24466, 2017.
[48] A. Bogomjakov, C. Gotsman, and M. Isenburg, "Distortion-free steganography for polygonal meshes," *Comput. Graph. Forum*, vol. 27, no. 2, pp. 637–642, Apr. 2008.
[49] N. C. Huang, M. T. Li, and C. M. Wang, "Toward optimal embedding capacity for permutation steganography," *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 802–805, Sep. 2009.
[50] S.-C. Tu and W.-K. Tai, "A high-capacity data-hiding approach for polygonal meshes using maximum expected level tree," *Comput. Graph.*, vol. 36, no. 6, pp. 767–775, Oct. 2012.
[51] S.-C. Tu, W.-K. Tai, M. Isenburg, and C.-C. Chang, "An improved data hiding approach for polygon meshes," *Vis. Comput.*, vol. 26, no. 9, pp. 1177–1181, 2010.
[52] J.-W. Cho, R. Prost, and H.-Y. Jung, "An oblivious watermarking for 3-D polygonal meshes using distribution of vertex norms," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 142–155, Jan. 2007.
[53] R. Ohbuchi, A. Mukaiyama, and S. Takahashi, "A frequency-domain approach to watermarking 3D shapes," *Comput. Graph. Forum*, vol. 21, pp. 373–382, Sep. 2002.
[54] Y. Yang, R. Pintus, H. Rushmeier, and I. Ivrissimtzis, "A 3D steganalytic algorithm and steganalysis-resistant watermarking," *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 2, pp. 1002–1013, Feb. 2017.
[55] A. Piva and M. Barni, "The first BOWS contest (break our watermarking system)," *Proc. SPIE*, vol. 6505, pp. 425–434, Feb. 2007.
[56] P. Bas and T. Furon, "The second BOWS contest (break our watermarking system)," Tech. Rep., 2007. [Online]. Available: http://bows2.ec-lille.fr
[57] T. Denemark and J. Fridrich, "Steganography with multiple JPEG images of the same scene," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 10, pp. 2308–2319, Oct. 2017.
[58] L. Pibre, J. Pasquet, D. Ienco, and M. Chaumont, "Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover sourcemismatch," in *Proc. Media Watermarking, Secur., Forensics, Int. Symp. Electron. Imag. (IS&T)*. San Francisco, CA, USA: SPIE, 2016.

[59] R. Cogranne, Q. Giboulot, and P. Bas, "The ALASKA steganalysis challenge: A first step towards steganalysis 'into the wild,'" in *Proc. ACM Workshop Inf. Hiding Multimedia Secur.*, Paris, France, Jul. 2019, pp. 125–137.

[60] R. Cogranne, Q. Giboulot, and P. Bas, "ALASKA#2: Challenging academic research on steganalysis with realistic images," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, New York, NY, USA, Dec. 2020, pp. 1–5.

[61] R. Cogranne, Q. Giboulot, and P. Bas, "Steganography by minimizing statistical detectability: The cases of JPEG and color images," in *Proc. ACM Workshop Inf. Hiding Multimedia Secur.*, Denver, CO, USA, Jun. 2020, pp. 161–167.

[62] L. Guo, J. Ni, W. Su, C. Tang, and Y.-Q. Shi, "Using statistical image model for JPEG steganography: Uniform embedding revisited," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 12, pp. 2669–2680, Dec. 2015.

[63] Z. Yang, K. Wang, S. Ma, Y. Huang, X. Kang, and X. Zhao, "IStego100K: Large-scale image steganalysis dataset," in *Digital Forensics and Watermarking*, H. Wang, X. Zhao, Y. Shi, H. J. Kim, and A. Piva, Eds. Cham, Switzerland: Springer, 2020, pp. 352–364.

[64] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.

[65] D.-T. Dang-Nguyen, C. Pasquini, V. Conotter, and G. Boato, "RAISE: A raw images dataset for digital image forensics," in *Proc. 6th ACM Multimedia Syst. Conf.*, New York, NY, USA, Mar. 2015, pp. 219–224.

[66] J. Fridrich, T. Pevný, and J. Kodovský, "Statistically undetectable JPEG steganography: Dead ends challenges, and opportunities," in *Proc. 9th Workshop Multimedia Secur.* New York, NY, USA: Assoc. Comput. Machinery, 2007, pp. 3–14.

[67] M. Rezaei, M. Riahi, and H. Hayati, "STEGRT1: A dataset for evaluating steganalysis systems in real-world scenarios," in *Proc. 28th Iranian Conf. Electr. Eng. (ICEE)*, Aug. 2020, pp. 1–5.

[68] H. Ruiz, M. Yedroudj, M. Chaumont, F. Comby, and G. Subsol, "LSSD: A controlled large JPEG image database for deep-learning-based steganalysis 'into the wild,'" in *Proc. Int. Conf. Pattern Recognit.*, 2021, pp. 470–483.

[69] S. Katzenbeisser and F. A. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, 1st ed. Norwood, MA, USA: Artech House, Inc., 2000.

[70] A. Westfeld, "F5—A steganographic algorithm," in *Proc. 4th Int. Workshop Inf. Hiding*. Berlin, Germany: Springer-Verlag, 2001, pp. 289–302.

[71] C.-L. Liu and S.-R. Liao, "High-performance JPEG steganography using complementary embedding strategy," *Pattern Recognit.*, vol. 41, no. 9, pp. 2945–2955, Sep. 2008.

[72] N. Provos, "Defending against statistical steganalysis," in *Proc. 10th Conf. USENIX Secur. Symp.*, vol. 10. Berkeley, CA, USA: USENIX Assoc., 2001, pp. 1–13.

[73] T. Sharp, "An implementation of key-based digital signal steganography," in *Proc. 4th Int. Workshop Inf. Hiding*. Berlin, Germany: Springer-Verlag, 2001, pp. 13–26.

[74] D.-C. Wu and W.-H. Tsai, "A steganographic method for images by pixel-value differencing," *Pattern Recognit. Lett.*, vol. 24, nos. 9–10, pp. 1613–1626, Jun. 2003.

[75] X. Chen, A. Golovinskiy, and T. Funkhouser, "A benchmark for 3D mesh segmentation," *ACM Trans. Graph.*, vol. 28, no. 3, pp. 1–12, Jul. 2009.

[76] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao, "3D ShapeNets: A deep representation for volumetric shapes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1912–1920.

[77] "The Stanford 3D scanning repository," 2014. [Online]. Available: https://graphics.stanford.edu/data/3Dscanrep/3Dscanrep.html

[78] L. Fan, W. Sun, and G. Feng, "Image steganalysis via random subspace Fisher linear discriminant vector functional link network and feature mapping," *Mobile Netw. Appl.*, vol. 24, no. 4, pp. 1269–1278, Aug. 2019.

[79] Y. Kulkarni and A. Gorkar, "Intensive image malware analysis and least significant bit matching steganalysis," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2020, pp. 2309–2317.

[80] M. B. Desai and S. V. Patel, "PFA-based feature selection for image steganalysis," *Int. J. Bioinf. Res. Appl.*, vol. 14, no. 1/2, p. 119, 2018.

[81] S. Li, M. Gu, L. Cheng, X. Chi, and M. Sun, "An accelerated divide-and-conquer algorithm for the bidiagonal SVD problem," *SIAM J. Matrix Anal. Appl.*, vol. 35, no. 3, pp. 1038–1057, Jan. 2014.

[82] X. Liao, S. Li, Y. Lu, and J. E. Roman, "A parallel structured divide-and-conquer algorithm for symmetric tridiagonal eigenvalue problems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 2, pp. 367–378, Feb. 2021.

[83] S. Arivazhagan, W. S. L. Jebarani, and S. Keerthinathan, "Extended adaptive pixel pair matching for data hiding in medical images," *J. Intell. Fuzzy Syst.*, vol. 29, no. 2, pp. 877–883, Oct. 2015.

[84] W. Hong and T.-S. Chen, "A novel data embedding method using adaptive pixel pair matching," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 1, pp. 176–184, Feb. 2012.

[85] T. Filler and J. Fridrich, "Gibbs construction in steganography," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 4, pp. 705–720, Dec. 2010.

[86] 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.

[87] S. T. Veena and S. Arivazhagan, "Quantitative steganalysis of spatial LSB based stego images using reduced instances and features," *Pattern Recognit. Lett.*, vol. 105, pp. 39–49, Apr. 2018.

[88] S. Veena and S. Arivazhagan, "Local descriptor based steganalysis of spatial LSB variant stego images," in *Proc. TEQIP 2nd Int. Conf. Comput., Commun. Innov.*, 2016, pp. 54–57.

[89] T. A. Feo and M. G. Resende, "Greedy randomized adaptive search procedures," *J. Global Optim.*, vol. 6, no. 2, pp. 109–133, 1995.

[90] Y. Kang, F. Liu, C. Yang, L. Xiang, X. Luo, and P. Wang, "Color image steganalysis based on channel gradient correlation," *Int. J. Distrib. Sensor Netw.*, vol. 15, no. 5, 2019, Art. no. 1550147719852031.

[91] H. Abdulrahman, M. Chaumont, P. Montesinos, and B. Magnier, "Color images steganalysis using RGB channel geometric transformation measures," *Secur. Commun. Netw.*, vol. 9, no. 15, pp. 2945–2956, 2016.

[92] M. Goljan, J. Fridrich, and R. Cogranne, "Rich model for steganalysis of color images," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2014, pp. 185–190.

[93] X. Gui, X. Li, and B. Yang, "Steganalysis of LSB matching based on local binary patterns," in *Proc. 10th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Aug. 2014, pp. 475–480.

[94] F. Liu, X. Yan, and Y. Lu, "Feature selection for image steganalysis using binary bat algorithm," *IEEE Access*, vol. 8, pp. 4244–4249, 2020.

[95] R. Y. M. Nakamura, L. A. M. Pereira, K. A. Costa, D. Rodrigues, J. P. Papa, and X.-S. Yang, "BBA: A binary bat algorithm for feature selection," in *Proc. 25th SIBGRAPI Conf. Graph., Patterns Images*, Aug. 2012, pp. 291–297.

[96] M. Sheikhan, M. S. Moin, and M. Pezhmanpour, "Blind image steganalysis via joint co-occurrence matrix and statistical moments of contourlet transform," in *Proc. 10th Int. Conf. Intell. Syst. Design Appl.*, Nov. 2010, pp. 368–372.

[97] S. Cho, B.-H. Cha, M. Gawecki, and C. J. Kuo, "Block-based image steganalysis: Algorithm and performance evaluation," *J. Vis. Commun. Image Represent.*, vol. 24, no. 7, pp. 846–856, 2013.

[98] Q. Liu, A. H. Sung, and M. Qiao, "Neighboring joint density-based JPEG steganalysis," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 2, pp. 1–16, Feb. 2011.

[99] V. Holub and J. Fridrich, "Low-complexity features for JPEG steganalysis using undecimated DCT," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 2, pp. 219–228, Feb. 2015.

[100] J. Kodovský, J. Fridrich, and V. Holub, "Ensemble classifiers for steganalysis of digital media," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 432–444, Apr. 2012.

[101] X. Song, F. Liu, C. Yang, X. Luo, and Y. Zhang, "Steganalysis of adaptive JPEG steganography using 2D Gabor filters," in *Proc. 3rd ACM Workshop Inf. Hiding Multimedia Secur.*, Jun. 2015, pp. 15–23.

[102] D. D. Shankar and A. S. Azhakath, "Minor blind feature based steganalysis for calibrated JPEG images with cross validation and classification using SVM and SVM-PSO," *Multimedia Tools Appl.*, vol. 80, no. 3, pp. 4073–4092, Jan. 2021.

[103] H. Jégou, M. Douze, and C. Schmid, "Hamming embedding and weak geometric consistency for large scale image search," in *Proc. Eur. Conf. Comput. Vis.*, vol. 5302, Oct. 2008, pp. 304–317.

[104] G. Schaefer and M. Stich, "UCID: An uncompressed color image database," *Proc. SPIE*, vol. 5307, pp. 472–480, Dec. 2003.

[105] P. Liu, F. Liu, C. Yang, and X. Song, "Improving steganalysis by fusing SVM classifiers for JPEG images," in *Proc. Int. Conf. Comput. Sci. Mech. Autom. (CSMA)*, Oct. 2015, pp. 185–190.

[106] J. Kodovsky and J. Fridrich, "Steganalysis of JPEG images using rich models," *Proc. SPIE*, vol. 8303, pp. 81–93, Feb. 2012.

[107] J. Lu, G. Zhou, C. Yang, Z. Li, and M. Lan, "Steganalysis of content-adaptive steganography based on massive datasets pre-classification and feature selection," *IEEE Access*, vol. 7, pp. 21702–21711, 2019.

[108] D. D. Shankar and A. S. Azhakath, "Small embed cross-validated JPEG steganalysis in spatial and transform domain using SVM," in *Advances in Machine Learning and Computational Intelligence*. Singapore: Springer, 2021, pp. 283–291.
[109] W. Luo, F. Huang, and J. Huang, "Edge adaptive image steganography based on LSB matching revisited," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 2, pp. 201–214, Jun. 2010.
[110] J. Mielikainen, "LSB matching revisited," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 285–287, May 2006.
[111] I. Lubenko and A. Ker, "Steganalysis using logistic regression," *Proc. SPIE*, vol. 7880, pp. 193–203, Feb. 2011.
[112] A. D. Ker, "Steganalysis of embedding in two least-significant bits," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 1, pp. 46–54, Mar. 2007.
[113] S. Hetzl and P. Mutzel, "A graph–theoretic approach to steganography," in *Proc. IFIP Int. Conf. Commun. Multimedia Secur.* Berlin, Germany: Springer-Verlag, 2005, pp. 119–128.
[114] P. Sallee, "Model-based steganography," in *Proc. Int. Workshop Digit. Watermarking*, vol. 2939, 2003, pp. 154–167.
[115] Y. Yang and I. Ivrissimtzis, "Mesh discriminative features for 3D steganalysis," *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 10, no. 3, pp. 1–13, Apr. 2014.
[116] Z. Li and A. G. Bors, "3D mesh steganalysis using local shape features," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 2144–2148.
[117] Z. Li and A. G. Bors, "Selection of robust features for the cover source mismatch problem in 3D steganalysis," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 4256–4261.
[118] D. Kim, H.-U. Jang, H.-Y. Choi, J. Son, I.-J. Yu, and H.-K. Lee, "Improved 3D mesh steganalysis using homogeneous kernel map," in *Proc. Int. Conf. Inf. Sci. Appl.* 2017, pp. 358–365.
[119] Z. Li and A. G. Bors, "Steganalysis of 3D objects using statistics of local feature sets," *Inf. Sci.*, vols. 415–416, pp. 85–99, Nov. 2017.
[120] Z. Li, D. Gong, F. Liu, and A. G. Bors, "3D steganalysis using the extended local feature set," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 1683–1687.
[121] H. Zhou, K. Chen, W. Zhang, C. Qin, and N. Yu, "Feature-preserving tensor voting model for mesh steganalysis," *IEEE Trans. Vis. Comput. Graph.*, vol. 27, no. 1, pp. 57–67, Jan. 2021.
[122] Z. Li and A. G. Bors, "Steganalysis of meshes based on 3D wavelet multiresolution analysis," *Inf. Sci.*, vol. 522, pp. 164–179, Jun. 2020.
[123] M. Lounsbery, T. D. DeRose, and J. Warren, "Multiresolution analysis for surfaces of arbitrary topological type," *ACM Trans. Graph.*, vol. 16, no. 1, pp. 34–73, 1997.
[124] Y. Yang, N. Peyerimhoff, and I. Ivrissimtzis, "Linear correlations between spatial and normal noise in triangle meshes," *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 1, pp. 45–55, Jan. 2013.
[125] B. R. Ghosh, S. Banerjee, A. Chakraborty, S. Saha, and J. K. Mandal, "A deep learning based image steganalysis using gray level co-occurrence matrix," in *Proc. 2nd Int. Conf. Adv. Electr., Comput., Commun. Sustain. Technol. (ICAECT)*, Apr. 2022, pp. 1–8.
[126] Y. Zang, C. Ding, W. Hu, and C. Fu, "HRANet: Histogram-residual-attention network used to measure neatness of toy placement," *Signal, Image Video Process.*, vol. 2022, pp. 1–9, Apr. 2022.
[127] T. Abazar, P. Masjedi, and M. Taheri, "An efficient ensemble of convolutional deep steganalysis based on clustering," in *Proc. 6th Int. Conf. Web Res. (ICWR)*, Apr. 2020, pp. 260–264.
[128] S. Tan and B. Li, "Stacked convolutional auto-encoders for steganalysis of digital images," in *Proc. Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA), Asia–Pacific*, Dec. 2014, pp. 1–4.
[129] J. Masci, U. Meier, D. C. Ciresan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," in *Proc. ICANN*, 2011, pp. 52–59.
[130] S. Wu, S. Zhong, and Y. Liu, "Deep residual learning for image steganalysis," *Multimedia Tools Appl.*, vol. 77, pp. 10437–10453, Feb. 2017.
[131] G. Xu, H.-Z. Wu, and Y. Q. Shi, "Ensemble of CNNs for steganalysis: An empirical study," in *Proc. 4th ACM Workshop Inf. Hiding Multimedia Secur.*, Jun. 2016, pp. 103–107.
[132] 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.
[133] M. Yedroudj, F. Comby, and M. Chaumont, "Yedroudj-Net: An efficient CNN for spatial steganalysis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 2092–2096.
[134] G. Xu, "Deep convolutional neural network to detect J-UNIWARD," in *Proc. 5th ACM Workshop Inf. Hiding Multimedia Secur.*, Jun. 2017, pp. 67–73.
[135] 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.
[136] J. Zeng, S. Tan, B. Li, and J. Huang, "Large-scale JPEG image steganalysis using hybrid deep-learning framework," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 5, pp. 1200–1214, May 2018.
[137] Y. Yousfi, J. Butora, J. Fridrich, and Q. Giboulot, "Breaking ALASKA: Color separation for steganalysis in JPEG domain," in *Proc. ACM Workshop Inf. Hiding Multimedia Secur.* New York, NY, USA: Assoc. Comput. Machinery, 2019, pp. 138–149.
[138] M. Boroumand, M. Chen, and J. Fridrich, "Deep residual network for steganalysis of digital images," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 5, pp. 1181–1193, May 2019.
[139] Y. Yousfi, J. Butora, E. Khvedchenya, and J. Fridrich, "ImageNet pre-trained CNNs for JPEG steganalysis," in *Proc. IEEE Int. Workshop Inf. Forensics Secur. (WIFS)*, Dec. 2020, pp. 1–6.
[140] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.
[141] M. Tan and Q. V. Le, "MixConv: Mixed depthwise convolutional kernels," 2019, *arXiv:1907.09595*.
[142] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[143] X. Wu, Z. Shao, P. Ou, and S. Tan, "Application of quantisation-based deep-learning model compression in JPEG image steganalysis," *J. Eng.*, vol. 2018, no. 16, pp. 1402–1406, Nov. 2018.
[144] C. Wang and J. Ni, "An efficient JPEG steganographic scheme based on the block entropy of DCT coefficients," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 1785–1788.