

Evaluation of Deep Learning and Conventional Approaches for Image Steganalysis

Guoliang Xie¹, Jinchang Ren^{1,*}, Huimin Zhao², Sophia Zhao¹, Stephen Marshall¹

1 Dept. of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK

2 School of Computer Science, Guangdong Polytechnic Normal University, Guangzhou, P.R. China

*Corresponding author: jinchang.ren@strath.ac.uk

Abstract. Steganography is the technique for embedding secret messages into digital media without changing their appearances. As a countermeasure to steganography, steganalysis detects the presence of hidden data in digital content. For the last decade, the majority of image steganalysis approaches can be formed by two stages. The first stage is to extract effective features from the image content and the second is to train a classifier in machine learning by using the features from stage one. Ultimately the image steganalysis becomes a binary classification problem. Since Deep Learning related architecture unifies these two stages and saves researchers lots of time designing hand-crafted features, the design of a CNN-based steganalyzer has therefore received increasing attention over the past few years. In this paper, we will examine the development in image steganalysis, both in the spatial domain and in the JPEG domain, and discuss the future directions.

Keywords: Steganalysis, Deep-learning, Feature extractor.

1 Introduction

Image steganography is the method for communicating secret messages under the cover images, which needs a cover source, a secret message and a secure steganographic scheme that hides the secret message under the cover source. Once a cover image is embedded with the secret message, it becomes a stego image. Usually, the change of embedding is extremely subtle that people would not differentiate between a cover image and a stego image by their looks.

During the last decade, many adaptive steganographic algorithms have been proposed to hide secret messages inside carriers, which are much difficult to detect. Such algorithms can be found in the spatial domain, for example, MiPOD [1], S-UNIWARD [2], HILL [3], WOW [4] and HUGO [5]. And in the frequency domain or JPEG domain, J-UNIWARD [6] and UED-JC [7] are two classic adaptive steganographic algorithms.

Image steganalysis, on the other hand, is the method used for analyzing whether an image is a cover or a stego image. Overall, image steganalysis can be divided into two different types, the first is called specific steganalysis and the second is called

universal steganalysis. Specific steganalysis means a steganalysis method is designed for a designated steganographic method. While universal steganalysis means it works for any steganographic method in a specific domain. Since specific steganalysis is no longer updated for the last decade, we will focus on universal steganalysis only in this paper, both in the spatial and JPEG domain.

We will divide the universal image steganalysis into conventional image steganalysis and deep-learning-based image steganalysis. To avoid confusion, we defined conventional image steganalysis as using hand-crafted feature extractors to produce features from images for training the classifiers in machine learning. In conventional image steganalysis, statistical features are usually used in training the classifiers while in deep-learning-based methods, features are learned automatically by convolutional neural networks (CNN).

Fig.1 shows the comparison of the structures between conventional methods and deep-learning-based methods, and the figure is extracted from Qian’s work [26]. For a better comparison between conventional steganalysis and deep-learning-based steganalysis, we further divide the two-stage into three parts in Fig.1. As Fig.1 suggests, a conventional steganalysis technique is consist of the pre-processing part, the feature representation part and the classification part. Researchers need to find out how to combine these three parts effectively to make a good steganalyzer. Although a deep-learning-based technique is also composed of three parts, they are in fact inside the same architecture.

Starting from 2015 [26], researchers began to focus more on deep-learning-based image steganalysis methods as effective conventional hand-crafted feature extractors become more and more difficult to design, and conventional methods tend to have large dimensionalities in their features, which makes the computation difficult. Deep-learning based methods, however, merge the pre-processing part (image processing layer), feature extraction part (convolutional layer) and classification part (fully connected layer) into a complete convolutional neural network. These networks have helped the researchers save lots of time finding effective feature extraction techniques as the networks themselves would do so automatically. Besides, deep-learning-based methods are usually performed on GPUs (Graphics Processing Unit, GPU), which makes the computation much faster than using CPUs [8].

2 Conventional Universal Image Steganalysis

2.1 Conventional universal image steganalysis in the spatial domain

Due to space limitations, our survey will start with the SPAM feature [9], which was a powerful feature extractor in the spatial domain. In this paper, the authors believed the secret messages can be regarded as additive noise to the cover images, and they modelled the “local dependences” in the cover image as a Markov chain. The empirical probability transition matrixes in the Markov chain are used as features to train the Support Vector Machine (SVM).

Later, Fridrich et al. [10] proposed a model called the Spatial Rich Model (SRM), which used linear and nonlinear high-pass filters to acquire stego signal and resulting

in a large feature vector compared to SPAM. Besides, they proposed to use ensemble classifiers (EC) [11] to solve the problem of the “curse of dimensionality” of the features, which also provide slightly better performance than the SVMs. From then on, researchers mainly use EC to verify their proposed methods, whose featural dimensionalities are relatively large, such as [12-15].

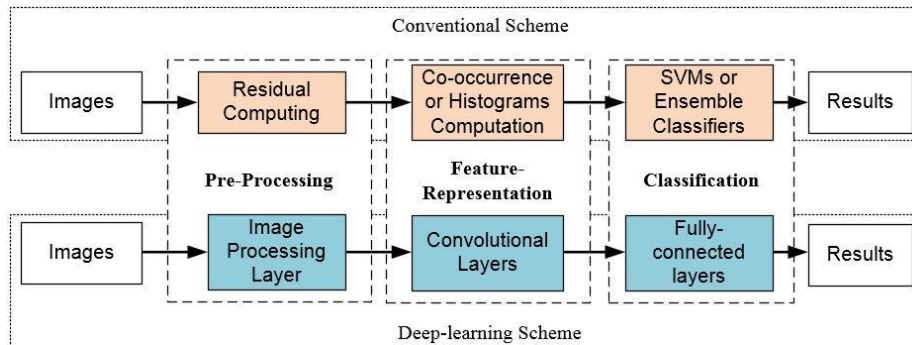


Fig. 1. Framework comparison between conventional methods (up) and deep-learning-based methods (down).

Instead of using the co-occurrence matrix as in [10], Holub et al. projected residuals of the images onto a set of random vectors and hence it is called PSRM (Projection Spatial Rich model) [13]. The method surpassed SRM in performance and it also reduced featural dimensionality. Besides, this method can be used in the JPEG domain, just as SRM.

In papers [13-14], the authors had found that by using the so-called selection-channel in steganalysis, the detection accuracy can be further improved, and Boroumand [37] et al. also proved it useful by using this technique in deep-learning-based methods later. The selection-channel is the probability of the elements in cover images that are being modified. By far, the maxSRM proposed in [14] is the most effective conventional feature extractor for image steganalysis.

Later, Li et al. [15] proposed an effective approach based on Threshold LBP (Local Binary Pattern) Operation in 2017. In this paper, although the authors still use the co-occurrence matrix to characterize pixel relationships, they found that TLBP operations are complementary to it in capturing local features. The results had proved that the TLBP is competitive to SRM, although it had not exceeded maxSRM in performance.

2.2 Conventional universal image steganalysis in JPEG domain

In the frequency domain, the images are usually compressed and coded according to some predefined transformation rules, and when performing image steganalysis in the frequency domain, the transformation rules have to be considered. In real-world applications, the JPEG format is the most common image format, which adopts discrete cosine transform to compress the images.

Since the modern adaptive steganographic schemes, such as J-UNIWARD [6], Uniform Embedding Distortion (UED) [7], would not introduce easily detected artefacts in DCT coefficients, these techniques can be better detected using Machine Learning instead of designing the specific steganalysis scheme.

We start from Chen’s feature extractor in the JPEG domain [16], which was classic and effective. The authors proposed to use both the “intra-block and inter-block correlations” of the coefficients in JPEG images as features to train classifiers, where “inter-block” means the JPEG coefficients in the same place in an 8×8 block. And we can capture the “intra-block correlation” by transition probability matrices of difference JPEG 2D arrays using Markov processes.

“Calibration” is the technique that helps to improve the accuracy both in JPEG and spatial image steganalysis by “providing the steganalyst with a reference image”. Although Calibration was first proposed in 2002 by Fridrich [17] et al., Kodovsky provides more detail in [18] to reveal the benefit of this technique and shows that it may have no benefit at all. By using the calibration technique, Liu’s method [19] managed to surpass the CC-PEV (Cartesian calibration, CC) technique proposed in [18] in terms of accuracy.

As a high-dimensional representation of images can provide better capability in capturing the complex features, Kodovsky et al. proposed their first high-dimensional rich model in the JPEG domain named CC-C300 [20]. Besides, they proposed their first ensemble classifier to deal with “the curse of dimensionality” reported in [20] as CC-C300 produces a 48600D feature for every jpeg image.

Kodovsky et al. proposed their JPEG rich model (JRM) with 11255D and its cartesian-calibrated version is abbreviated as CC-JRM with 22510D in paper [21], they also confirmed that “steganalysis can benefit from multiple-domain approaches” through the results tested by combining the Spatial Rich Model (SRM) and CC-JRM.

To solve the dimensionality-cursed problem, Holub et al. provided a relatively low computational complexity scheme named DCTR (Discrete Cosine Transform Residual) [22], which can be implemented quickly while still being competitive to the high-dimensional ones. The DCTR feature set is constructed from “the first-order statistics of quantized noise residuals” using 64 kernels of DCT.

Holub et al. found that smaller kernels are better in capturing the artefacts introduced by the adaptive JPEG steganographic schemes, therefore they proposed their PHARM (PHase Aware of pRojection Model) scheme [23], which used a bunch of “small-support kernels” to obtain residuals, and project them randomly as in PSRM. Also, they further confirmed that by decompressing the JPEG images to the spatial domain, adaptive steganographic schemes can be better detected. Later, Xia et al. [42] improve this method both in efficiency and effectiveness. In paper [24], Song et al. proposed to use 2D Gabor filters to acquire residuals from the decompressed JPEG images, and then they extracted the histogram features and merged them as the so-called GFR (Gabor Filter Residuals) feature set.

To provide a JPEG-version’s selection channel aware feature set, Denmark et al. utilized the selection channel technique to improve the performance of DCTR, PHARM and GFR in [25]. They proved that a selected channel can help to increase the performance significantly, especially when applied to the low-embedding rate situations.

In paper [43], Feng et al. proposed to use “diverse base filters” to extract stego noise. The MD-CFR (Maximum Diversity Cascade Filter Residual) they proposed jointly with SCA-GFR (Selection Channel Aware, SCA) managed to be an effective feature extractor for JPEG image steganalysis. In this scheme, they designed various base filters and then cascaded these base filters to get the high order filters, which are finally optimized through their filter-selection method.

3 Deep-learning based Image Steganalysis

A Convolutional-Neural-Network-based steganalyzer is usually composed of the following parts: a pre-processing part, a convolution part and a classification part. The pre-processing part is a set of high-pass filters for capturing the stego-like high-frequency signal. The convolution part is usually made of a convolution, an activation function, a pooling step and a normalization step. The classification part is usually equipped with a Softmax function which normalizes the output value between $[0,1]$.

3.1 The CNNs for spatial image steganalysis

Qian’s work [26] is in literature the first to use CNN in image steganalysis successfully, inspired by Krizhevsky’s CNN [27] in 2012. However, the achieved performance was 3-4% worse than the EC with the rich model. This paper sent a message that CNN could be very powerful in steganalysis if tuned and trained carefully.

Pibre et al. [28] followed up and showed the steganalysis results of a special case where the steganographer uses the same embedding key for different images all the time. In their report, they showed that when well parameterized, the CNN can provide a result better than that of the combination of a Rich Model (RM) with an Ensemble Classifier (EC).

Xu et al. [29] reported a CNN architecture that considers the advantages taken from conventional steganalysis. We refer to this network as Xu-Net in this paper. Although Xu-Net used only one high-pass filter, the proposed CNN was competitive compared to the SRM with ensemble classifiers on the BOSSbase [45]. The CNN they proposed did defeat the SRM+EC in detecting the S-UNIWARD and HILL algorithms.

In paper [30], Xu et al. improved the results by using an ensemble scheme. They formed a group of CNNs and trained them independently on a “random subsample of the training dataset”, and then their “intermediate feature representations” were extracted and pooled, which were then fed into the next classification.

Hoping to use transfer learning in image steganalysis, Qian et al. [31] proposed a transfer learning framework to help the training of CNN. In this work, they reported that the performance of a CNN on a low payload steganographic algorithm can be improved by transferring a pre-trained CNN on a higher payload.

Ye et al. [32] proposed the first CNN architecture that outperformed the best conventional spatial steganalysis scheme, maxSRM, and we refer to this work as Ye-Net. In this paper, they finally used 30 high-pass filters instead of one in previous works to extract the stego noise, and they proposed their activation function in their CNN model.

To further boost the performance, they used the selection-channel information from maxSRM.

Yedrouj et al. proposed an efficient Yedrouj-Net in [33] and claimed that their CNN provides better performance than Ye-Net even without the knowledge of the selection-channel. That will raise a question: Is the selection channel helping in the deep learning technique? Since the networks themselves may learn the possible embedding positions when sufficiently trained.

Li et al. proposed a wide structure named ReST-Net in paper [34], and they proposed to use diverse activation modules (DAMs) in this CNN architecture. The different activation modules will “activate the convolution outputs differently” and by using a parallel structure of CNN subnets, the ReST-Net can capture more “preprocessed information”.

Zhang et al. proposed their Zhu-Net in [35]. They proposed to use smaller kernels in the preprocessing layer, and used “separable convolution” for extracting “spatial and channel correlation of residuals”. Furthermore, spatial pyramid pooling is used to process the images with an arbitrary size, which also helps to aggregate local features.

Wu et al. [36] also used one kernel in preprocessing layer, which is the same as in [26], but they proposed to use residual-learning for preserving the stego noise. Experiments showed promising results when compared with the maxSRMd2+EC scheme.

Boroumand et al. proposed their SRNet in [37], and SRNet is the first CNN in the literature that can be used both for spatial images and JPEG images. Besides, according to the paper, the network is also the first to use selection-channel for JPEG domain steganalysis. The most important thing about this network is that they managed to free the designer from introducing the elements needed in training a CNN for image steganalysis, for example, the fixed kernels.

Note that Xu-Net, Ye-Net, Yedrouj-Net and Zhu-Net are using a fixed high-pass filter(s), some of these filters are updated during the CNN training process. However, SRNet randomly initialized these filters and they are updated during the training process. This scheme is proved to be one more step closer to automatic image steganalysis.

3.2 The CNNs for JPEG image steganalysis

Xu et al. proposed a convolutional neural network to detect J-UNIWARD in the paper [38]. In this report, different from the one in [37], where the authors disabled pooling in their front end, the authors confirmed that the pooling method is critical for performance. The 20-layer CNN proposed in the paper had been confirmed to outperform the best feature-based method, SCA-GFR. To better verify the performance, Xu et al. used a dataset called “CLS-LOC” from ImageNet [47], which contains over 1.2 million images.

Chen et al. proposed two CNNs for JPEG images called VNet and PNet in [39], they varied both in volume and inaccuracy. Besides, these networks were the first to consider JPEG phase-awareness. To further boost the performance, they proposed the “catalyst kernel” to capture the stego noise, which works much better than the KV kernel proposed in Qian’s work [26]. That is because the high-frequency noise in the

spatial domain was changed when being transformed in the JPEG domain, and the kernel that worked well in the spatial domain was no longer effective.

Yang et al. [40] proposed a relatively deeper CNN, which is composed of 32 layers. The network was designed so deep to “reuse the features by concatenating all features from the previous layers”, and in this way, model parameters are greatly reduced. Besides, this is the first paper that combines CNN and SCA-GFR methods by using the ensemble technique.

Zeng et al. proposed their “Hybrid Deep-Learning” framework [41], which was composed of two parts. The first part includes convolution, quantization and truncation as in the classic method [15], while the second part is composed of “multiple deep subnets”. That was the first work reported to use more than five million cover images in image steganalysis, which provides a benchmark for large-scale JPEG image steganalysis.

Hu et al. proposed their scheme based on “visual attention and reinforce learning” [44]. In their scheme, they used “attention-focused regions” to mark the area of interest, and they replace the images that are wrongly classified with the SoAFRs (the Summary of attention-focused regions), which is “merged by continuous decision-making” through reinforcement learning.

4 Results Comparisons

In this section, we would like to provide direct comparisons between some classic methods, both in the spatial domain and in the JPEG domain. In the spatial domain, we selected S-UNIWARD, WOW, HUGO and HILL as representative steganographic schemes, while in the JPEG domain, we chose J-UNIWARD and UED-JC. We put the results for 0.1bpp (bit per pixel) and 0.4bpp in spatial schemes, and the results for 0.1bpac (bit per AC-coefficient) and 0.4bpac in JPEG schemes since these payloads are the most frequently used.

For conventional methods, two databases BOSSbase [45] and BOWS-2 [46] are frequently used, they both contain 10,000 grayscale images. For deep-learning methods, besides the two datasets listed above, the CLS-LOC dataset [47] from “ImageNet” is used as well, in which 250,000 images were used.

In Table 1, the experimental results are shown in the second column (SPAM+EC) to the fifth column (TLBP+EC) are tested by us on the BOSSbase dataset only. The last two columns are extracted from paper [37], and the experiments were performed on BOSSbase and BOWS-2.

Table 1. Comparison of error-rate results from different steganalyzers in Spatial domain.

Feature Extractor	SPAM [9] (686D)	SRM [10] (34671D)	maxSRM [14]	TLBP [15] (29040D)	SCA- YeNET [32]	SRNeT [37]
S-UNI/ 0.1bpp	0.4561	0.4013	0.3682	0.4079	0.3220	0.3104
S-UNI/ 0.4bpp	0.3270	0.2074	0.1990	0.1913	0.1281	0.1023

0.4bpp						
HUGO/ 0.1bpp	0.4641	0.3666	0.3066	0.3768	-	-
HUGO/ 0.4bpp	0.3065	0.1807	0.1654	0.2014	-	-
WOW/ 0.1bpp	0.4740	0.4009	0.3015	0.3957	0.2442	0.2587
WOW/ 0.4bpp	0.3785	0.2092	0.1634	0.2081	0.0959	0.0893
HILL/ 0.1bpp	0.4765	0.4332	0.3796	0.4120	0.3380	0.3134
HILL/ 0.4bpp	0.3848	0.2509	0.2251	0.2325	0.1708	0.1414

Table 2 shows the results from four conventional methods and three CNNs for JPEG images. The results for conventional methods were extracted from paper [25], while the results for CNNs were extracted from [37]. The leftmost column means the steganographic method is J-UNIWARD or UED-JC with a quality factor of 75 or 95, and the embedding rate is 0.1 bpac or 0.4 bpac.

As the results suggest, although the maxSRM method was proposed in 2014, it is still a powerful tool in image steganalysis. Note that the selection-channel technique was not used in TLBP. Also, we can see from the result provided by SPAM that it can detect adaptive steganographic schemes to some extent but fails in comparing with any rich models. For deep-learning methods, we showed only two state-of-art networks, and the SCA-YeNET means Ye-Net [32] with the selection-channel.

In table 2, we put only one of the selection-channel-aware versions of the rich model in the JPEG domain, i.e. SCA-GFR, due to space constraints. Also, it can provide the best performance in these rich models. Besides, SRNeT is capable of detecting adaptive steganographic methods in both domains, so we put the results in this table.

Although CNN related methods had surpassed conventional methods in both domains, we still recommend researching conventional and deep-learning methods simultaneously. The reason is easy, as Boroumand et al. suggested in [37], we are still far away from fully automatic steganalysis, thus we still need insights from conventional methods to help improve the detecting performance.

Table 2. Comparison of error-rate results from different steganalyzers in the JPEG domain.

Feature Extractor	DCTR [22] (8000D)	GFR [24] (17000D)	PHARM [23] (12600D)	SCA-GFR [25]	J-XuNeT [38]	SRNeT [37]
J-UNI/ 75/0.1bpac	0.4400	0.4095	0.4284	0.3589	0.4310	0.3201
J-UNI/ 0.4bpac	0.1553	0.1005	0.1259	0.0792	0.1207	0.0670

75/0.4bpac						
UED-JC/	0.3593	0.3153	0.3376	0.2352	0.2144	0.1311
75/0.1bpac						
UED-JC/	0.0593	0.0346	0.0418	0.0377	0.0287	0.0188
75/0.4bpac						
J-UNI/	0.4826	0.4756	0.4835	0.4634	0.4812	0.4277
95/0.1bpac						
J-UNI/	0.3081	0.2721	0.3079	0.2617	0.3232	0.1762
95/0.4bpac						
UED-JC/	0.4598	0.4323	0.4490	0.3931	0.3848	0.3044
95/0.1bpac						
UED-JC/	0.2180	0.1663	0.1947	0.1662	0.1292	0.0877
95/0.4bpac						

5 Conclusions

Throughout this paper, we can conclude that the research in conventional image steganalysis has slowed down, while deep-learning-based image steganalysis is attracting more and more researchers. However, although many deep-learning-based methods have shown remarkable performance when compared with conventional methods, most of these schemes still rely on classical conventional methods. The fully-automatic image steganalysis is still far from reaching, therefore we still need to research both classic methods and CNN-related ones.

Papers such as Zeng's [41], revealed that GAN (Generative Adversarial Network) related work will be one of the interesting points in the future. However, since training a GAN equals training two CNNs simultaneously, how to accelerate the training speed and how to use the limited GPU resources effectively will become our problems. Furthermore, as training any of these networks usually takes an expensive graphic card or more, keep finding effective and efficient CNNs will always be the task.

Acknowledgement. This work was partly supported by National Natural Science Foundation of China (61772144, 61672008), Innovation Team Project (Natural Science) of the Education Department of Guangdong Province (2017KCXTD021), Foundation for Youth Innovation Talents in Higher Education of Guangdong Province (2018KQNCX139), Innovation Research Project (Natural Science) of Education Department of Guangdong Province (2016KTSCX077), Project for Distinctive Innovation of Ordinary Universities of Guangdong Province (2018KTSCX120), and Foreign Science and Technology Cooperation Plan Project of Guangzhou Science Technology and Innovation Commission (201807010059).

References

1. Sedighi, V. et al. "Content-adaptive steganography by minimizing statistical detectability." *IEEE Transactions on Information Forensics and Security* 11.2: 221-234. (2015)
2. Holub, V. et al. "Universal distortion function for steganography in an arbitrary domain." *EURASIP Journal on Information Security* 2014.1. (2014)
3. Li, B. et al. "A new cost function for spatial image steganography." 2014 IEEE International Conference on Image Processing (ICIP). (2014)
4. Holub, V. et al. "Designing steganographic distortion using directional filters." 2012 IEEE International workshop on information forensics and security (WIFS). (2012)
5. Pevný, T. et al. "Using high-dimensional image models to perform highly undetectable steganography." *International Workshop on Information Hiding*. Springer, Berlin. (2010)
6. Holub, V. et al. "Universal distortion function for steganography in an arbitrary domain." *EURASIP Journal on Information Security*. (2014)
7. Guo, L. et al. "Uniform embedding for efficient JPEG steganography." *IEEE transactions on Information Forensics and Security* 9.5: 814-825. (2014)
8. Schlegel, D. "Deep Machine Learning on Gpu." *University of Heidelber-Ziti* 12. (2015)
9. Pevny, T. et al. "Steganalysis by subtractive pixel adjacency matrix." *IEEE Transactions on Information Forensics and Security* 5.2: 215-224. (2010)
10. Fridrich, J. et al. "Rich Models for Steganalysis of Digital Images." *IEEE Transactions on Information Forensics and Security* 7.3:868-882. (2012)
11. Kodovsky, J. et al, and Vojtěch Holub. "Ensemble classifiers for steganalysis of digital media." *IEEE Transactions on Information Forensics and Security* 7.2: 432-444. (2011)
12. Holub, V. et al, Random Projections of Residuals as an Alternative to Co-occurrences in Steganalysis, *Proc. SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics XV*, vol. 8665, San Francisco, CA. (2013)
13. Holub, V. et al. "Random projections of residuals for digital image steganalysis." *IEEE Transactions on Information Forensics and Security* 8.12: 1996-2006. (2013)
14. Denemark, T. et al. "Selection-channel-aware rich model for steganalysis of digital images." 2014 IEEE International Workshop on Information Forensics and Security (WIFS). (2014)
15. Li, B. et al. "New steganalytic features for spatial image steganography based on derivative filters and threshold LBP operator." *IEEE Transactions on Information Forensics and Security* 13.5: 1242-1257. (2017)
16. Chen, C. et al. "JPEG image steganalysis utilizing both intrablock and interblock correlations." 2008 IEEE International Symposium on Circuits and Systems. IEEE. (2008)
17. Fridrich, J. et al "Steganalysis of JPEG images: Breaking the F5 algorithm." *International Workshop on Information Hiding*. Springer, Berlin. (2002)
18. Kodovský, J. et al "Calibration revisited." *Proceedings of the 11th ACM workshop on Multimedia and security*. ACM. (2009)
19. Liu, Q. "Steganalysis of DCT-embedding based adaptive steganography and YASS." *Proceedings of the thirteenth ACM multimedia workshop on Multimedia and security*. ACM. (2011)
20. Kodovský, J. et al "Steganalysis in high dimensions: Fusing classifiers built on random subspaces." *Media Watermarking, Security, and Forensics III*. Vol. 7880. International Society for Optics and Photonics. (2011)
21. Kodovský, J. et al "Steganalysis of JPEG images using rich models." *Media Watermarking, Security, and Forensics 2012*. Vol. 8303. International Society for Optics and Photonics. (2012)

22. Holub, V. et al "Low-complexity features for JPEG steganalysis using undecimated DCT." *IEEE Transactions on Information Forensics and Security* 10.2: 219-228. (2014)
23. Holub, V. et al "Phase-aware projection model for steganalysis of JPEG images." *Media Watermarking, Security, and Forensics 2015*. Vol. 9409. International Society for Optics and Photonics. (2015)
24. Song, X, et al. "Steganalysis of adaptive JPEG steganography using 2D Gabor filters." *Proceedings of the 3rd ACM workshop on information hiding and multimedia security*. ACM. (2015)
25. Denmark, T. et al "Steganalysis features for content-adaptive JPEG steganography." *IEEE Transactions on Information Forensics and Security* 11.8 (2016): 1736-1746.
26. Qian, Y, et al. "Deep learning for steganalysis via convolutional neural networks." *Media Watermarking, Security, and Forensics 2015*. Vol. 9409. International Society for Optics and Photonics. (2015)
27. Krizhevsky, A. et al "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. (2012)
28. Pibre L. et al. Deep learning is a good steganalysis tool when embedding key is reused for different images, even if there is a cover source mismatch[J]. *Electronic Imaging*, 4(8):1-11. (2016)
29. Xu, G. et al "Structural design of convolutional neural networks for steganalysis." *IEEE Signal Processing Letters* 23.5: 708-712. (2016)
30. Xu, G. et al. "Ensemble of CNNs for steganalysis: An empirical study." *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*. ACM. (2016)
31. Qian, Y. et al. "Learning and transferring representations for image steganalysis using convolutional neural network." 2016 *IEEE international conference on image processing (ICIP)*. (2016)
32. Ye, J. et al "Deep learning hierarchical representations for image steganalysis." *IEEE Transactions on Information Forensics and Security* 12.11: 2545-2557. (2017)
33. Yedroudj, M. et al "Yedroudj-Net: An Efficient CNN for Spatial Steganalysis." 2018 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. (2018)
34. Li, B. et al. "ReST-Net: Diverse activation modules and parallel subnets-based CNN for spatial image steganalysis." *IEEE Signal Processing Letters* 25.5: 650-654. (2018)
35. Zhang, R. et al. "Efficient feature learning and multi-size image steganalysis based on CNN." *arXiv preprint arXiv:1807.11428*. (2018)
36. Wu, S. et al. "Deep residual learning for image steganalysis." *Multimedia tools and applications* 77.9: 10437-10453. (2018)
37. Boroumand, M. et al "Deep residual network for steganalysis of digital images." *IEEE Transactions on Information Forensics and Security* 14.5: 1181-1193. (2018)
38. Xu, G. "Deep convolutional neural network to detect J-UNIWARD." *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*. ACM. (2017)
39. Chen, M. et al. "JPEG-phase-aware convolutional neural network for steganalysis of JPEG images." *Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security*. ACM. (2017)
40. Yang, J. et al. "JPEG steganalysis based on densenet." *arXiv preprint arXiv:1711.09335*. (2017)
41. Zeng, J. et al. "Large-scale JPEG image steganalysis using hybrid deep-learning framework." *IEEE Transactions on Information Forensics and Security* 13.5: 1200-1214. (2017)
42. Xia, C. et al. "Improved PHARM for JPEG Steganalysis: Making PHARM More Efficient and Effective." *IEEE Access* 7: 50339-50346. (2019)

43. Feng, G. et al. "Diversity-based Cascade Filters for JPEG Steganalysis." *IEEE Transactions on Circuits and Systems for Video Technology*. (2019)
44. Hu, D. et al. "Digital image Steganalysis based on Visual Attention and Deep Reinforcement Learning." *IEEE Access*. (2019)
45. Bas, P. et al. "Break our steganographic system": the ins and outs of organizing BOSS." *International workshop on information hiding*. Springer, Berlin. (2011)
46. Bas, P. et al.. BOWS-2. [Online]. Available: <http://bows2.ec-lille.fr> (2007)
47. Russakovsky, O. et al. "Imagenet large scale visual recognition challenge." *International journal of computer vision*115.3: 211-252. (2015)