applied sciences

MDPI

Article

Identification of Content-Adaptive Image Steganography Using Convolutional Neural Network Guided by High-Pass Kernel

Saurabh Agarwal [1,2] and Ki-Hyun Jung [2,*]

[1] Department of Computer Science and Engineering, Amity School of Engineering & Technology, Amity University Uttar Pradesh, Noida 201313, India
[2] Department of Software Convergence, Andong National University, Gyeongbuk 36729, Republic of Korea
* Correspondence: khanny.jung@gmail.com or kingjung@anu.ac.kr; Tel.: +82-54-820-7968; Fax: +82-54-820-6257

Abstract: Digital images are very popular and commonly used for hiding crucial data. In a few instances, image steganography is misused for communicating with improper data. In this paper, a robust deep neural network is proposed for the identification of content-adaptive image steganography schemes. Multiple novel strategies are applied to improve detection performance. Two non-trainable convolutional layers is used to guide the proposed CNN with fixed kernels. Thirty-one kernels are used in both non-trainable layers, of which thirty are high-pass kernels and one is the neutral kernel. The layer-specific learning rate is applied for each layer. ReLU with customized thresholding is applied to achieve better performance. In the proposed method, image down-sampling is not performed; only the global average pooling layer is considered in the last part of the network. The experimental results are verified on BOWS2 and BOSSBase image sets. Content-adaptive steganography schemes, such as HILL, Mi-POD, S-UNIWARD, and WOW, are considered for generating the stego images with different payloads. In experimental analysis, the proposed scheme is compared with some of the latest schemes, where the proposed scheme outperforms other state-of-the-art techniques in the most cases.

Keywords: image steganalysis; image steganography; convolutional neural network; deep neural network; content-adaptive steganography

Citation: Agarwal, S.; Jung, K.-H. Identification of Content-Adaptive Image Steganography Using Convolutional Neural Network Guided by High-Pass Kernel. *Appl. Sci.* **2022**, *12*, 11869. https://doi.org/10.3390/app122211869

Academic Editor: Giacomo Fiumara

Received: 14 October 2022
Accepted: 19 November 2022
Published: 21 November 2022

1. Introduction

Image steganography schemes are applied to hide secret data/information that cannot be noticed by the naked eye. In most content-adaptive steganography schemes, syndrome-trellis code (STC) is considered to make the schemes less vulnerable. In this paper, four notable content-adaptive steganography schemes, such as HILL [1], Mi-POD [2], S-UNIWARD [3], and WOW [4], are analyzed. In steganography, there is only one unit value (addition or subtraction) in some image pixels, which is in accordance with the steganography scheme and message. STC ensures that there is no visible change in the image. In comparison to other classification problems, such as object recognition and texture classification, steganalysis is more challenging due to the low value of stego noise. Difference arrays of cover and stego images in the different steganography schemes are displayed in Figure 1, where the cover image [5] is shown in the first row. Difference arrays of cover and HILL, Mi-POD, S-UNIWARD, and WOW stego images are shown in the second and third rows for 0.2 bit per pixel (bpp) and 0.4 bpp payloads, respectively. Although it is difficult to interpret the difference between different steganography schemes, the changes are more evident with a higher value of the payload.

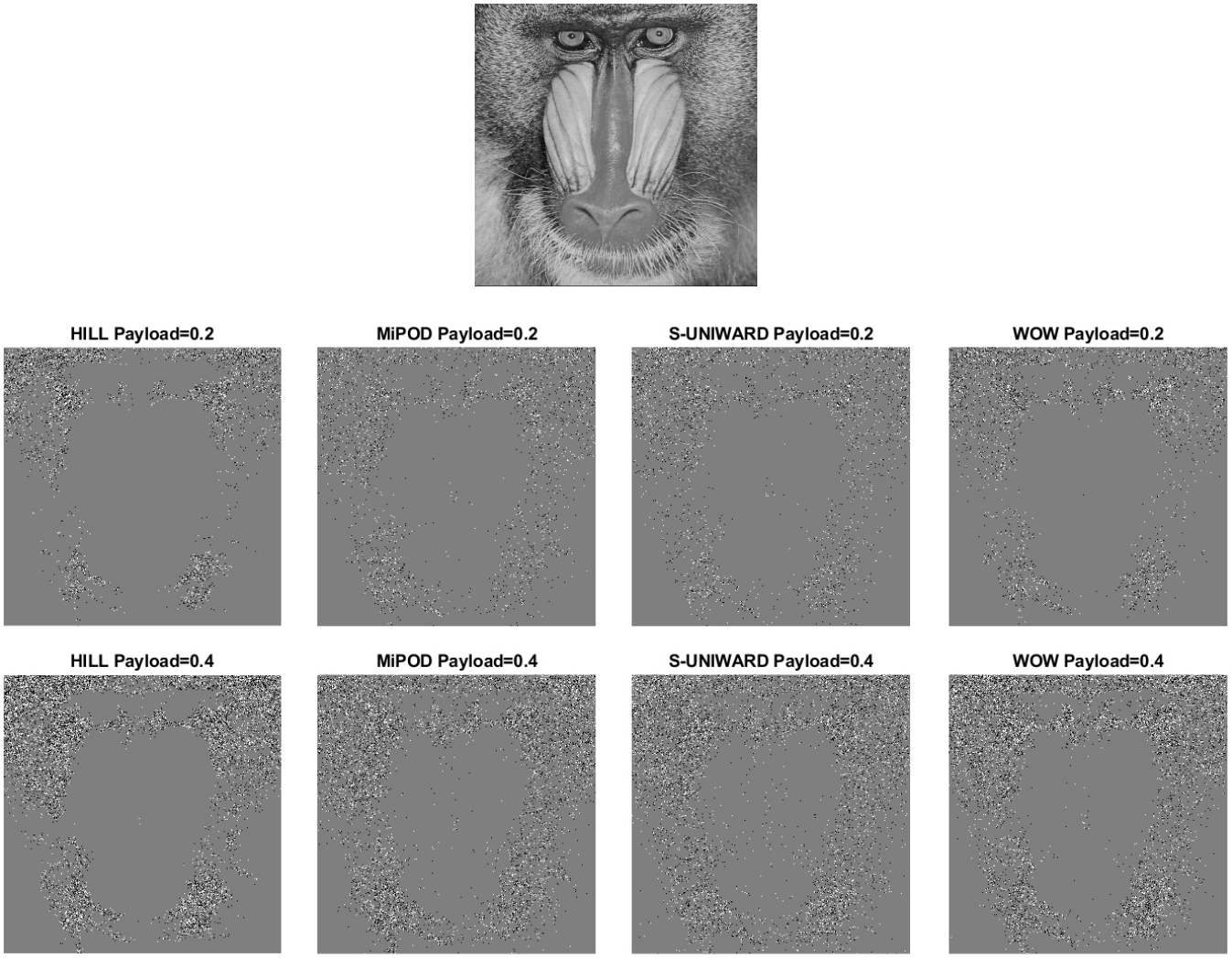


Figure 1. Cover image with residual arrays of cover and stego images.

The effect of steganography with different schemes cannot be understood effectively using natural images, as displayed in Figure 1. A computer-generated (CG) image with (0–255) is displayed in the first row of Figure 2 to understand a better effect of different steganography schemes. There is a difference of one unit value in adjacent pixels, with the exception of major diagonal pixel values, as well as a difference of 255 unit values in major diagonal pixel values. In different images of cover and stego images, it can be observed that each scheme has different behavior. In the HILL scheme, the changes in pixel values are more diversified than Mi-POD, followed by S-UNIWARD. However, WOW behavior is entirely different. The variations in the behavior of schemes are also reflected in experimental analysis.

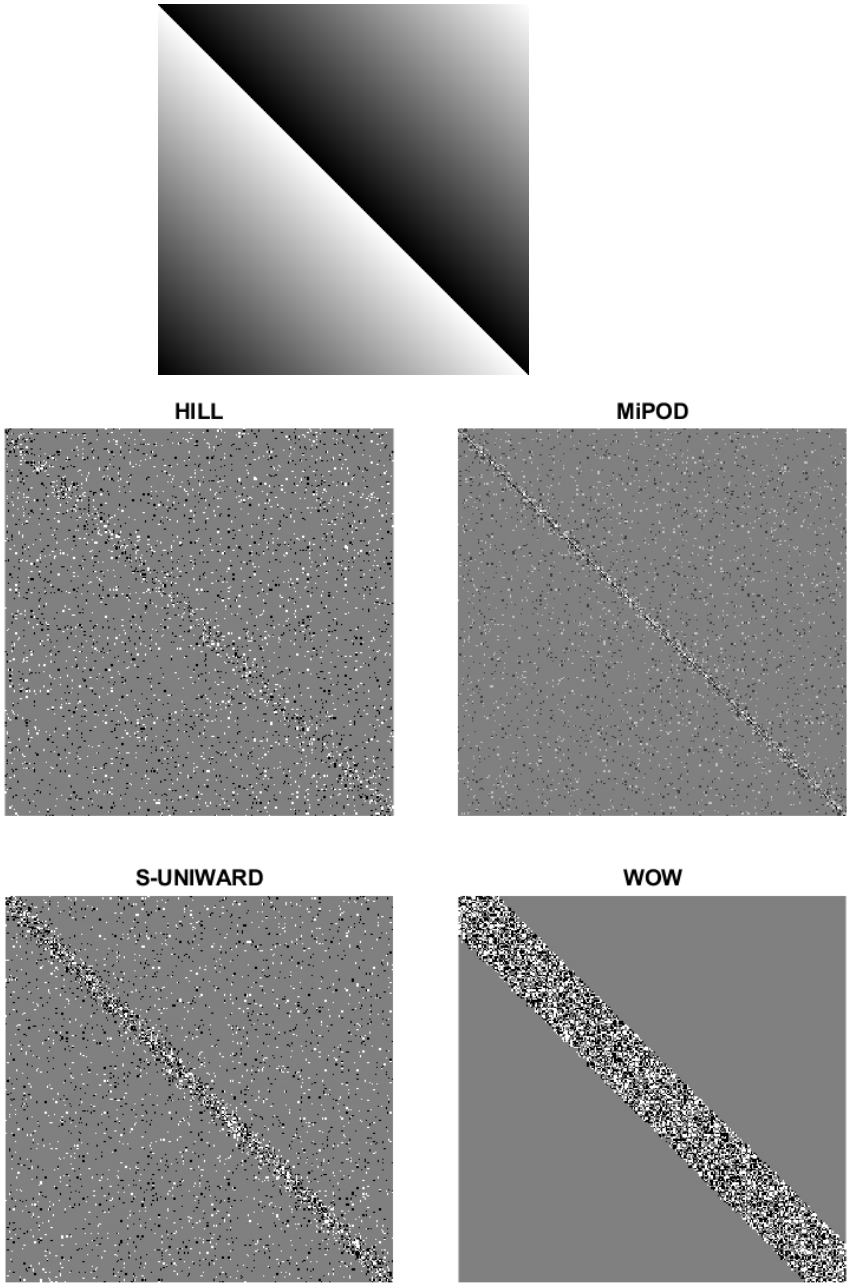


Figure 2. CG cover image and residual arrays of cover-stego images.

Several attempts have been made for the detection of image steganography. Initially, the thriving steganalysis is performed using Markov model-based features. Moreover, some researchers proposed steganalysis schemes based on texture descriptors. The first order and higher order difference arrays and residual 2D arrays are exploited for large-size feature vectors. As the large-size feature vector is more effective in steganalysis due to low stego noise, a feature vector is extracted from the spatial and/or frequency domain. Lyu and Farid [6] exploited the statistics of magnitude and phase information in the frequency domain for the detection of F5 [7] and Outguess [8] steganography schemes. Li et al. [9] used the texture and Markov features, with a dimension of 22,153, for the detection of HUGO [10] steganography scheme. Fridrich and Kodovsky [11] introduced thirty high-pass kernels that are frequently recognized as SRM kernels. These kernels are based on the first and higher order difference arrays and multiple quantization factors. Additionally, SRM kernels are applicable in most deep networks. The feature vector with a dimension of 34,671 is extracted using a Markov model, where the scheme is verified on HUGO and ± 1 variants

steganography schemes. Penvy et al. [12] extracted Markov features of the second order from the first order difference arrays for the detection of the ±1 steganography scheme. The variants of the scheme [11], abbreviated as maxSRM and maxSRMd2, are proposed by Denemark et al. [13]. The formation process of the co-occurrence matrix is modified to decrease the feature vector size in the maxSRM and maxSRMd2 schemes. The Denemark scheme is verified on WOW, S-UNIWARD, and S-UNIGARD steganography schemes. Furthermore, Tang et al. [14] utilized the Markov model, although the dense and edge region is exploited for better results. The local correlation pattern [15] is proposed for identifying HUGO, S-UNIWARD, and the LSB matching revisited LSBMR [16] steganography schemes. Li et al. [17] proposed the steganalysis scheme based on texture operator features and verified it on HUGO and WOW steganography schemes. In this scheme, images are preprocessed with high-pass kernels. Moreover, PCA is used to decrease the size of feature vector and increase the performance. In [18–20], popular texture operators, such as the variant of the local binary pattern (LBP) with some improvements, are proposed for steganalysis. Moreover, Markov features are combined for better performance. In [18,19], schemes are verified on HILL, Mi-POD, and S-UNIWARD steganography schemes. A steganalysis scheme [20] is applied for the detection of HILL, CMD-HILL [21], and Mi-POD steganography schemes. At present, many recent steganalysis schemes depend on deep networks, as the high computing power hardware is not an obstacle. Qian et al. [22,23] introduced the deep network for the detection of HUGO, S-UNIWARD, and WOW steganography schemes. Images are preprocessed by one 5×5 high-pass kernel. However, the results are inferior to some manual schemes. In the Xu et al. method [24], the tanh is applied rather than ReLU for the identification of HILL and S-UNIWARD steganography schemes. In [25,26], numerous residual 2D arrays and one 5×5 high-pass kernel are used to enhance the performance on many types of stego images. Ye et al. [27] used the residual array after processing with thirty high-pass kernels [11]. Truncated linear unit (TLU) is considered rather than ReLU to outperform previous schemes on WOW, HILL, and S-UNIWARD. A deep network known as steganalysis residual network (SRNet) [28] is introduced, where skip connections and pooling layers with other types of layers are used to create three categories of blocks. Moments are extracted from a trained deep network to identify HILL, S-UNIWARD, and WOW for uncompressed images and J-UNIWARD and UED-JC [29] for compressed images. Yedroudj et al. [30] utilized thirty high-pass kernels [11] in a non-trainable convolutional layer. In [31], a new concept of shared batch normalization is proposed to improve the performance. The images are preprocessed with twenty high-pass kernels [11], which is verified on HILL, HUGO, S-UNIWARD, and WOW. In Zhang et al. [32], thirty high-pass kernels [11] are utilized in a non-trainable layer. In particular, two concepts are incorporated with the suggested CNN, bottleneck approach and spatial pyramid pooling, for the detection of three popular types of stego images. Furthermore, Xiang et al. [33] utilized thirty high-pass kernels in a non-trainable convolutional layer with some improvements in the order of layers. Wang et al. [34] preprocessed the image with thirty high-pass kernels and included the transfer learning approach. The authors exploited both spatial and frequency domains. The network was initialized with trained CNN from low-capacity stego images and verified on S-UNIWARD and WOW stego images. As can be seen from previous research on image steganalysis, the majority of the previous articles can improve the performance by exploiting the high-pass kernels in preprocessing to increase the stego noise for better classification. In this paper, two measures are considered to reduce the detection error without increasing the computational cost. The key measurements of the proposed scheme are outlined as follows:

- In the proposed scheme, thirty-one kernels are used; thirty are high-pass kernels and one is the neutral kernel.
- Two non-trainable convolutional layers are considered using thirty-one kernels; one layer is used at the beginning of the network and the second before the middle of the network.
- To retain a complete statistical information, down-sampling is not performed.
- The layer-specific learning rate is considered for better results.

- The clipped ReLU layer is applied with a customized cut-off value for better control on the CNN.
- Softmax classifier is the popular choice in CNN. However, several classifiers are investigated and the SVM classifier is the most suitable.
- The outcomes of the proposed scheme are equated with the popular schemes Zhu-Net, SRNet, Yedroudj-Net, and Ye-Net.
- The comprehensive outcomes are discussed for HILL, Mi-POD, S-UNIWARD, and WOW steganography schemes with 0.2, 0.3, and 0.4 bpp payloads.
- The proposed scheme is discussed in detail in Section 2. The experimental evaluation is carried out in Section 3 and the conclusions are presented in Section 4.

2. The Proposed Scheme

Numerous image steganography schemes are used to hide some secret data in an image. At present, deep networks are found to be effective in state-of-the-art steganalysis schemes. In this paper, a robust deep neural network for the steganalysis of content-adaptive steganography schemes is proposed. The variation in different steganography schemes can be understood from Figure 2, and only one standard solution can be proposed for steganalysis. The effect of HILL, Mi-POD, S-UNIWARD, and WOW steganography schemes can be understood by covariance plots of image entropy in Figure 3. A total of twenty thousand images from BOWS2 [35] and BOSSBase [36] datasets is considered as a cover image. The same number of corresponding stego images with payload PL = {0.1, 0.2, 0.3, 0.4} is created using HILL, Mi-POD, S-UNIWARD, and WOW steganography schemes. In Figure 3, it is evident that a very small gap exists for cover and HILL PL = 0.1, which indicates that the detection of HILL may be challenging. However, the gap is more evident in Mi-POD followed by S-UNIWARD and WOW. Moreover, the behavior of covariance plot is reflected in the classification of cover and stego images. The highest detection error is found in HILL, followed by Mi-POD, S-UNIWARD, and WOW.

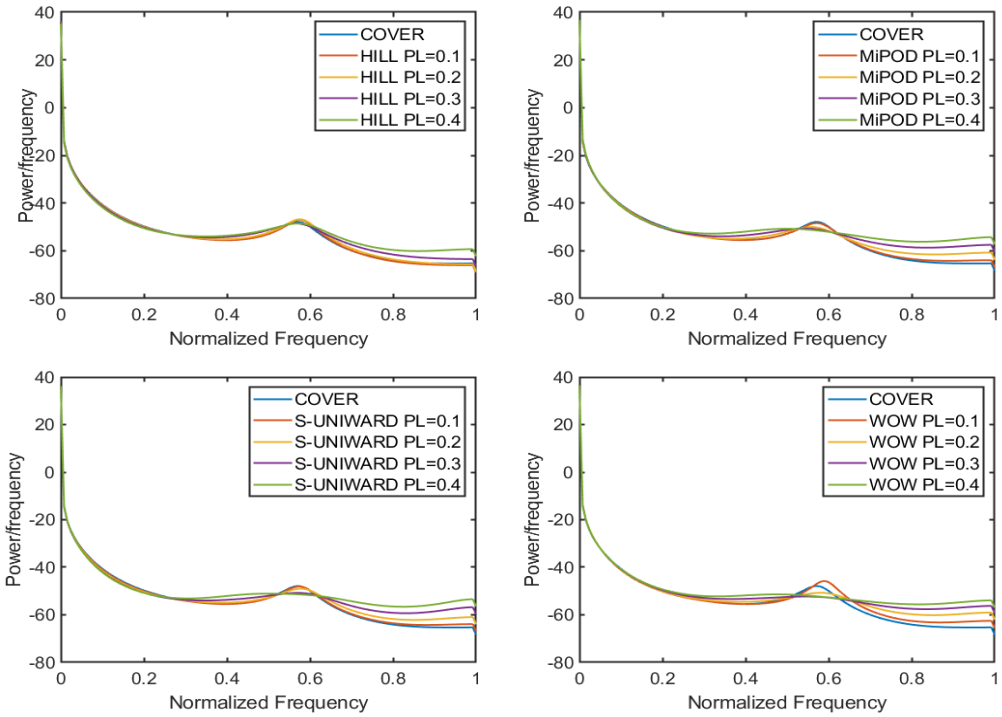


Figure 3. Covariance plots of cover and stego images entropy.

The block diagram of the proposed scheme is shown in Figure 4. The proposed convolutional neural network has two non-trainable convolutional layers. Thirty high-pass kernels and one neutral kernel of size 5×5 are used in non-trainable convolutional layers.

In Figure 5, one neutral kernel and three high-pass kernels are displayed to understand the effect of kernels. Each kernel provides a different type of information. High-pass kernels increase the stego noise. Conventionally, in [27,30–34], only one non-trainable convolutional layer was used with thirty high-pass kernels. The neutral kernel is applied to retain the unprocessed information from the preceding layer. Moreover, a neutral filter increases the statistical information as unprocessed information is added with high-pass kernel operated information. The effect is verified in experimental analysis. The proposed CNN is guided by two non-trainable convolutional layers. It has already been established that a single non-trainable convolutional layer with thirty high-pass kernels enhances the detection performance for cover and stego images. In experimental analysis, it is discovered that the performance can improve significantly using an additional non-trainable convolutional layer. However, there is an adverse effect when three non-trainable convolutional layers are tried. Therefore, two non-trainable layers are considered. The second non-trainable layer provides the best results when used not very far from the first non-trainable layer. If the distance between non-trainable layers becomes very far, then there is a drop in the performance of the network.

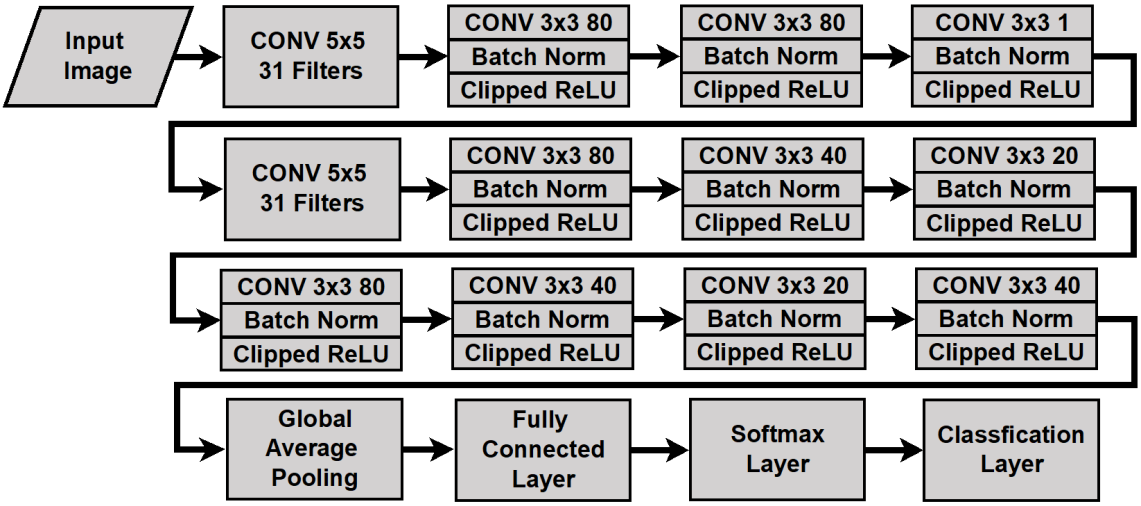


Figure 4. Block diagram of the proposed scheme.

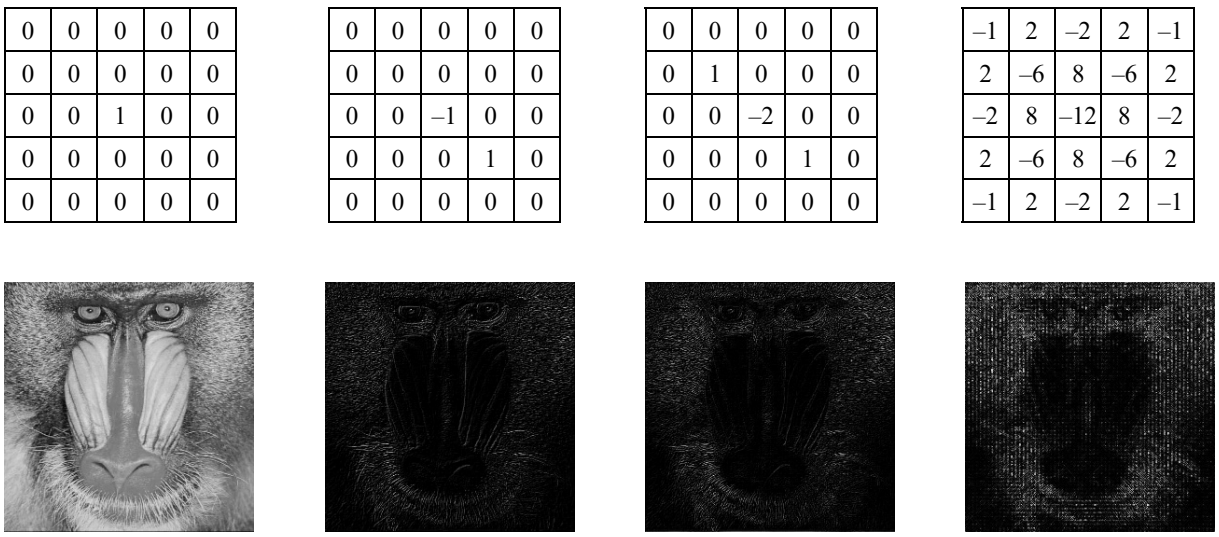


Figure 5. Neutral kernel and high-pass kernels.

The network considers ten trainable convolutional layers, and their weights are initialized using Glorot and Bengio [37] scheme. Each trainable convolutional layer is followed by batch normalization and a clipped ReLU layer [38]. The layer-specific learning rate is considered for each trainable convolutional layer. In previous research, a fixed learning rate, irrespective of the layer, was considered. Layer-specific learning rate [39] can improve the performance substantially which is found to be effective in many other applications [40,41]. The layer-specific learning is adapted by rigorous experimental analysis.

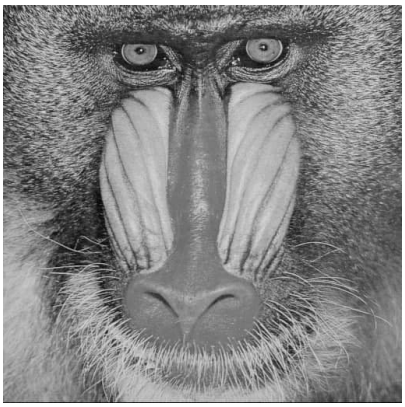
The ReLU layer is utilized in the majority of the earlier networks. ReLU replaces the negative elements with zero and remains intact in non-negative elements. The ReLU can be defined as:

$$E(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

In this paper, the clipped ReLU (CReLU) is used to give an improved control in the proposed CNN. Similar to ReLU, CReLU replaces the negative components with zero while replacing the non-negative items that are higher than a threshold value with threshold value (t). Positive elements smaller than the threshold value remain intact. The CReLU can be defined as:

$$C(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x < t \\ t, & x \geq t \end{cases}$$

In the proposed CNN, down-sampling is not performed in order to gain the utmost statistical information from the preceding layers. After the last clipped ReLU layer, the global average pooling (GAP) layer is applied [30,42]. The GAP layer is effective in decreasing the detection error. Stochastic gradient descent (SGD) optimizer is used to train the CNN. The initial learning rate of 1×10^{-2} and L2 regularization factor 1×10^{-4} are considered. One hundred epochs are considered to fine-tune the proposed CNN. Mini-batch of size 10 is considered. In the experimental analysis, the SVM classifier with the quadratic kernel gave better results than the softmax classifier. Therefore, the SVM classifier is utilized in the proposed CNN to classify cover and stego images. In Figure 6, the effect of four kernels in layers three (second row), thirteen (third row), and twenty-eight (fourth row) is displayed after training the CNN. Each layer has a different type of effect on the image. The depth of the network has a significant effect on the kernels of the layer. As the depth increases, the specific details become more prominent.



(**a**)

Figure 6. *Cont.*

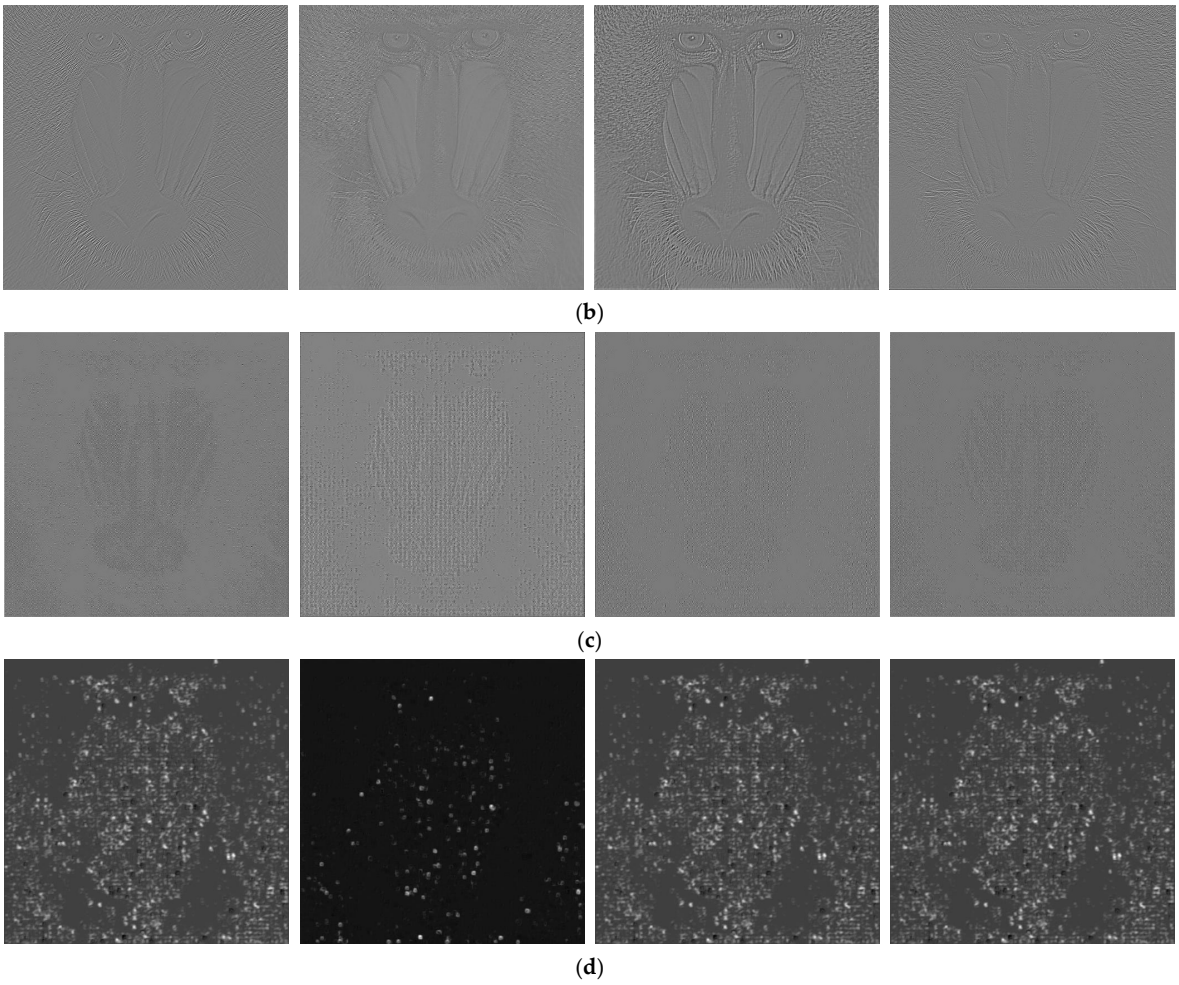


Figure 6. Effect of different layer kernels. (**a**) An image, (**b**) layer 3 kernels effect, (**c**) layer 13 kernels effect, (**d**) layer 28 kernels effect.

3. Experimental Analysis

Digital image steganography is very popular for covert communication. Steganalysis is challenging as there are little changes with a lower impact on the stego image. In this paper, four popular content-adaptive steganography schemes are evaluated by classifying cover and stego images. HILL, Mi-POD, S-UNIWARD, and WOW steganography schemes are tested using the proposed scheme on BOWS2 [35] and BOSSBase [36] image sets. Each set has ten thousand images with a dimension of 512 × 512 pixels of different varieties. Experimental analysis is performed after changing the size of the image to 256 × 256 pixels using interpolation. In most of the previous schemes, 256 × 256 pixel images are considered. The proposed scheme is compared with the latest schemes, SRNet [28], Ye-Net [27], Yedroudj-Net [30], and Zhu-Net [32]. BOSSBase image set is used to create cover and stego images for experimental analysis and experimental results are shown in Tables 1–6. Five thousand images of each class, stego and cover, are considered for training the CNN. One thousand and four thousand images are considered for validation and testing, respectively. In Tables 7 and 8, experimental results are shown using BOWS2 [35] and BOSSBase [36] image sets. The results are defined in terms of detection error. The small value of detection error represents a better performance.

In Table 1, the detection errors are compared using thirty and thirty-one kernels in non-trainable layers. As can be seen in Table 1, a neutral filter can improve the performance and reduce the detection error. In earlier articles, only thirty kernels are used. The additional kernel passes some unchanged information to subsequent layers, which can increase

the detection accuracy. The improvement in experimental results can be seen in four steganography schemes with three payloads.

Table 1. Detection errors using thirty and thirty-one kernels.

Steganography Scheme	**HILL**			**Mi-POD**			**S-UNIWARD**			**WOW**		
Payload (bpp)	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
Thirty Non-trainable kernels	0.3875	0.3242	0.2712	0.3475	0.2583	0.2378	0.3308	0.2166	0.1702	0.2523	0.1888	0.1512
Thirty-one Non-trainable kernels	0.3796	0.3198	0.2661	0.3414	0.2548	0.2343	0.3234	0.2132	0.1682	0.2491	0.1857	0.1484

In general, one preprocessing or non-trainable convolutional layer is considered in previous networks. In the proposed scheme, CNN is guided by two non-trainable convolutional layers that contain thirty-one kernels. The first non-trainable layer is considered after the image input layer, and the second non-trainable layer after three trainable convolutional layers. In experiments, three non-trainable layers are also considered. However, the best performance is achieved when two non-trainable layers are used. There is a noticeable reduction in detection error when compared with the single non-trainable layer, as depicted in Table 2.

Table 2. Detection errors using different non-trainable convolutional layers.

Steganography Scheme	**HILL**			**Mi-POD**			**S-UNIWARD**			**WOW**		
Payload (bpp)	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
Single Non-trainable Layer	0.3925	0.3287	0.2739	0.3489	0.2624	0.2420	0.3337	0.2217	0.1749	0.2545	0.1905	0.1521
Two Non-trainable Layers	0.3796	0.3198	0.2661	0.3414	0.2548	0.2343	0.3234	0.2132	0.1682	0.2491	0.1857	0.1484
Three Non-trainable Layers	0.3841	0.3246	0.2696	0.3479	0.2578	0.2390	0.3289	0.2157	0.1707	0.2525	0.1894	0.1503

The layer-specific learning rate is considered in the proposed scheme. Typically, a common learning rate is considered in each convolutional layer. Although the customized learning rate reduces the detection error moderately as shown in Table 3, customized learning rates are found after an exhaustive experimental analysis. Four steganography methods with three different payloads can be more effective in the proposed steganalysis scheme.

Table 3. Detection errors of fixed and layer-specific learning rate.

Steganography Scheme	**HILL**			**Mi-POD**			**S-UNIWARD**			**WOW**		
Payload (bpp)	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
Fixed learing rate	0.3898	0.3278	0.2715	0.3506	0.2606	0.2388	0.3289	0.2164	0.1731	0.2548	0.1898	0.1516
Layer-specific learning rate	0.3796	0.3198	0.2661	0.3414	0.2548	0.2343	0.3234	0.2132	0.1682	0.2491	0.1857	0.1484

In Table 4, experimental results are compared using ReLU and clipped ReLU. High-pass kernels in non-trainable layers increase the values, although CReLU is considered to control the increase. As discussed in the previous section, some specific values are controlled by the threshold in clipped ReLU. Similar to ReLU, CReLU replaces the negative components with zero while replacing the non-negative items that are higher than a threshold value with a particular threshold value. Positive elements below the threshold value are retained.

Table 4. Detection errors using ReLU and clipped ReLU.

Steganography Scheme	**HILL**			**Mi-POD**			**S-UNIWARD**			**WOW**		
Payload (bpp)	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
ReLU	0.3856	0.3236	0.2720	0.3462	0.2581	0.2397	0.3302	0.2170	0.1721	0.2540	0.1900	0.1519
Cipped ReLU	0.3796	0.3198	0.2661	0.3414	0.2548	0.2343	0.3234	0.2132	0.1682	0.2491	0.1857	0.1484

Support vector machine (SVM) is a very popular classifier. The proposed scheme replaces the softmax classifier with an SVM classifier with a quadratic kernel for a better detection of cover and stego images. There is a fair deduction in detection errors using the SVM classifier as shown in Table 5. A detailed analysis regarding the outperformance of SVM over softmax classifier is performed by Tang [43]. While SVM seeks to maximize the margin between data points belonging to different classes, the softmax classifier minimizes the cross-entropy. The proposed steganography scheme can successfully classify the cover and four steganography techniques with 0.2, 0.3, and 0.4 bpp payloads.

Table 5. Detection errors using softmax and SVM classifiers.

Steganography Scheme	**HILL**			**Mi-POD**			**S-UNIWARD**			**WOW**		
Payload (bpp)	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
Softmax Classifier	0.3854	0.3262	0.2735	0.3476	0.2620	0.2422	0.3298	0.2202	0.1757	0.2553	0.1933	0.1567
SVM Classifier	0.3796	0.3198	0.2661	0.3414	0.2548	0.2343	0.3234	0.2132	0.1682	0.2491	0.1857	0.1484

The proposed scheme is compared with the recent schemes SRNet [28], Ye-Net [27], Yedroudj-Net [30], and Zhu-Net [32] as shown in Table 6. The proposed scheme outperforms other schemes with the exception of two cases. As represented in Figure 2, the spread of stego noise is more uniform than other steganography schemes, making the detection of HILL very challenging. HILL has the maximum detection error than the other schemes. WOW has the least detection error as the stego noise is more concentrated in some specific areas. The proposed scheme outperforms other schemes with a payload of 0.2 bpp, with the exception of Mi-POD and S-UNIWARD. A detection error reduction from 2.37% to 13.77% is observed in most cases when the proposed scheme is used, with an average decline of 6.92% in classification error.

Table 6. Detection errors of the proposed and other schemes using BOSSBase image set.

Steganography Scheme	**HILL**			**Mi-POD**			**S-UNIWARD**			**WOW**		
Payload (bpp)	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
SRNet	0.4560	0.3789	0.3305	0.4221	0.3417	0.2806	0.3568	0.2686	0.2225	0.2850	0.2226	0.1783
Ye-Net	0.4672	0.4202	0.3736	0.4311	0.3733	0.3470	0.4058	0.3301	0.2706	0.3228	0.2922	0.2214
Yedroudj-Net	0.4710	0.4216	0.3372	0.4294	0.3798	0.2952	0.4122	0.3189	0.2757	0.3074	0.2652	0.2071
Zhu-Net	0.3888	0.3339	0.2878	0.3385	0.2828	0.2576	0.3167	0.2391	0.1951	0.2689	0.2339	0.1489
Proposed Scheme	0.3796	0.3198	0.2661	0.3414	0.2548	0.2343	0.3234	0.2132	0.1682	0.2491	0.1857	0.1484

In Table 7, both BOWS2 and BOSSBase image sets are considered for experimental analysis. Seven thousand images from each dataset are considered as the cover image for training the CNN. Corresponding stego images are created using a particular steganography scheme with specific payloads. The trained CNN is evaluated using six thousand remaining images of BOWS2 and BOSSBase image sets. There is a significant reduction in detection error after considering twenty-eight thousand images of both classes for training and twelve thousand images for testing the proposed CNN. Moreover, the experimental results show that the categorization error increases as the payload increases, which implies that a low payload makes the classification difficult.

Table 7. Detection errors of proposed and other schemes using BOWS2 and BOSSBase image sets.

Steganography Scheme	**HILL**			**Mi-POD**			**S-UNIWARD**			**WOW**		
Payload (bpp)	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
SRNet	0.4180	0.3468	0.3170	0.3967	0.3212	0.2638	0.3392	0.2495	0.2087	0.2505	0.1832	0.1401
Ye-Net	0.4601	0.3994	0.3527	0.4052	0.3509	0.3262	0.3864	0.3065	0.2511	0.3009	0.2402	0.1947
Yedroudj-Net	0.4426	0.3819	0.3257	0.4036	0.3570	0.2775	0.3835	0.2968	0.2543	0.2831	0.2130	0.1677
Zhu-Net	0.3718	0.3169	0.2414	0.3209	0.2658	0.2421	0.2875	0.2102	0.1695	0.2315	0.1702	0.1057
Proposed Scheme	0.3608	0.2867	0.2476	0.3109	0.2495	0.2203	0.2691	0.1922	0.1546	0.2066	0.1561	0.1011

In Table 8, experimental results are displayed after augmenting BOWS2 and BOSSBase image sets. One hundred thousand images are considered for training the CNN, and twenty-five thousand images for testing the trained CNN. There is a tremendous reduction in detection error using augmented images. Then, ten thousand images (Table 6) and twenty-eight thousand images (Table 7) are considered for training the CNN. The proposed scheme outperforms the other schemes and Zhu-Net by decreasing the detection error from 4% to 10%. Thirty high-pass kernels were applied by Ye et al. [27], Yedroudj et al. [30], and Zhang et al. [32]. The proposed scheme outcomes have been improved significantly with the use of thirty-one kernels and two non-trainable convolutional layers.

Table 8. Detection errors of proposed and other schemes using augmented images.

Steganography Scheme	HILL			Mi-POD			S-UNIWARD			WOW		
Payload (bpp)	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4	0.2	0.3	0.4
SRNet	0.3739	0.3135	0.2920	0.3095	0.2505	0.2058	0.2930	0.2136	0.1807	0.2226	0.1592	0.1185
Ye-Net	0.4324	0.3682	0.3220	0.3161	0.2737	0.2544	0.3599	0.2761	0.2135	0.2847	0.2175	0.1697
Yedroudj-Net	0.4093	0.3456	0.3140	0.3148	0.2785	0.2165	0.3479	0.2673	0.2292	0.2608	0.1774	0.1381
Zhu-Net	0.3469	0.2981	0.2291	0.2703	0.2073	0.1889	0.2312	0.1936	0.1400	0.1932	0.1413	0.0795
Proposed Scheme	0.3166	0.2655	0.2114	0.2482	0.1868	0.1718	0.2129	0.1859	0.1286	0.1765	0.1296	0.0718

4. Conclusions

In this paper, a robust steganalysis scheme has been proposed for the detection of content-adaptive steganography schemes. Two non-trainable convolutional layers with fixed thirty-one kernels were incorporated to guide the proposed CNN, and multiple novel strategies were considered to increase the performance of CNN. Moreover, customized learning rates were applied to convolutional layers. In the proposed scheme, the clipped ReLU was considered rather than ReLU for improvement and the SVM classifier was used as an alternative to softmax classifier. Furthermore, the proposed scheme was verified on popular content-adaptive steganography schemes, such as HILL, Mi-POD, S-UNIWARD, and WOW. A remarkable reduction in the detection errors from 2% to 10% was observed in the classification of cover and stego images. Furthermore, in the most cases, the performance of the proposed scheme was more outstanding than other state-of-the-art techniques.

Author Contributions: Both authors discussed the details of the manuscript. S.A. designed and wrote the manuscript. S.A. implemented the proposed technique and provided the experimental results. K.-H.J. drafted and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Brain Pool program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (2019H1D3A1A01101687) and Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1I1A3049788).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used in this paper are publicly available, and their links are provided in the reference section.

Acknowledgments: We thank the anonymous reviewers for their valuable suggestions that improved the quality of this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, B.; Wang, M.; Huang, J.; Li, X. A new cost function for spatial image steganography. In Proceedings of the 2014 IEEE International Conference on Image Processing, ICIP 2014, Paris, France, 27–30 October 2014; pp. 4206–4210. [CrossRef]
2. Sedighi, V.; Cogranne, R.; Fridrich, J. Content-Adaptive Steganography by Minimizing Statistical Detectability. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 221–234. [CrossRef]
3. Holub, V.; Fridrich, J.; Denemark, T. Universal distortion function for steganography in an arbitrary domain. *EURASIP J. Inf. Secur.* **2014**, *2014*, 1. [CrossRef]

4. Holub, V.; Fridrich, J. Designing steganographic distortion using directional filters. In Proceedings of the WIFS 2012—Proceedings of the 2012 IEEE International Workshop on Information Forensics and Security, Costa Adeje, Spain, 2–5 December 2012; pp. 234–239. [CrossRef]
5. Weber, A.G. The USC-SIPI Image Database: Version 5. *USC-SIPI Rep.* **2006**, *315*, 1–24.
6. Lyu, S.; Farid, H. Steganalysis Using Higher-Order Image Statistics. *IEEE Trans. Inf. Forensics Secur.* **2006**, *1*, 111–119. [CrossRef]
7. Westfeld, A. F5—A Steganographic Algorithm. In *International Workshop on Information Hiding*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 289–302.
8. Provos, N.; Honeyman, P. Detecting Steganographic Content on the Internet. 2001. Available online: http://niels.xtdnet.nl/papers/detecting.pdf/ (accessed on 21 January 2022).
9. Li, B.; Huang, J.; Shi, Y.Q. Textural features based universal steganalysis. In Proceedings of the Security, Forensics, Steganography, and Watermarking of Multimedia Contents X, San Jose, CA, USA, 27–31 January 2008; p. 681912. [CrossRef]
10. Pevný, T.; Filler, T.; Bas, P. Using High-Dimensional Image Models to Perform Highly Undetectable Steganography. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2010; Volume 6387, pp. 161–177.
11. Fridrich, J.; Kodovsky, J. Rich Models for Steganalysis of Digital Images. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 868–882. [CrossRef]
12. Pevny, T.; Bas, P.; Fridrich, J. Steganalysis by Subtractive Pixel Adjacency Matrix. *IEEE Trans. Inf. Forensics Secur.* **2010**, *5*, 215–224. [CrossRef]
13. Denemark, T.; Sedighi, V.; Holub, V.; Cogranne, R.; Fridrich, J. Selection-channel-aware rich model for Steganalysis of digital images. In Proceedings of the 2014 IEEE International Workshop on Information Forensics and Security, WIFS 2014, Atlanta, GA, USA, 3–5 December 2015; pp. 48–53. [CrossRef]
14. Tang, W.; Li, H.; Luo, W.; Huang, J. Adaptive steganalysis against WOW embedding algorithm. In Proceedings of the 2nd ACM workshop on Information hiding and multimedia security—IH&MMSec '14, Salzburg, Austria, 11–13 June 2014; pp. 91–96. [CrossRef]
15. Xu, X.; Dong, J.; Wang, W.; Tan, T. Local correlation pattern for image steganalysis. In Proceedings of the 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), Chengdu, China, 12–15 July 2015; pp. 468–472. [CrossRef]
16. Mielikainen, J. LSB matching revisited. *IEEE Signal Process. Lett.* **2006**, *13*, 285–287. [CrossRef]
17. Li, F.; Zhang, X.; Cheng, H.; Yu, J. Digital image steganalysis based on local textural features and double dimensionality reduction. *Secur. Commun. Netw.* **2016**, *9*, 729–736. [CrossRef]
18. Li, B.; Li, Z.; Zhou, S.; Tan, S.; Zhang, X. New Steganalytic Features for Spatial Image Steganography Based on Derivative Filters and Threshold LBP Operator. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 1242–1257. [CrossRef]
19. Wang, P.; Liu, F.; Yang, C. Towards feature representation for steganalysis of spatial steganography. *Signal Process.* **2020**, *169*, 107422. [CrossRef]
20. Ge, H.; Hu, D.; Xu, H.; Li, M.; Zheng, S. New Steganalytic Features for Spatial Image Steganography Based on Non-negative Matrix Factorization. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2020; Volume 12022, pp. 337–351.
21. Li, B.; Wang, M.; Li, X.; Tan, S.; Huang, J. A Strategy of Clustering Modification Directions in Spatial Image Steganography. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 1905–1917. [CrossRef]
22. Qian, Y.; Dong, J.; Wang, W.; Tan, T. Deep learning for steganalysis via convolutional neural networks. *Media Watermarking Secur. Forensics* **2015**, *2015*, 94090J. [CrossRef]
23. Qian, Y.; Dong, J.; Wang, W.; Tan, T. Learning and transferring representations for image steganalysis using convolutional neural network. In Proceedings of the International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016. [CrossRef]
24. Xu, G.; Wu, H.-Z.; Shi, Y.-Q. Structural Design of Convolutional Neural Networks for Steganalysis. *IEEE Signal Process. Lett.* **2016**, *23*, 708–712. [CrossRef]
25. Wu, S.; Zhong, S.H.; Liu, Y. Steganalysis via deep residual network. In Proceedings of the International Conference on Parallel and Distributed Systems—ICPADS, Wuhan, China, 13–16 December 2016; pp. 1233–1236. [CrossRef]
26. Wu, S.; Zhong, S.; Liu, Y. Deep residual learning for image steganalysis. *Multimed. Tools Appl.* **2017**, *77*, 10437–10453. [CrossRef]
27. Ye, J.; Ni, J.; Yi, Y. Deep Learning Hierarchical Representations for Image Steganalysis. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 2545–2557. [CrossRef]
28. Boroumand, M.; Chen, M.; Fridrich, J. Deep Residual Network for Steganalysis of Digital Images. *IEEE Trans. Inf. Secur.* **2019**, *14*, 1181–1193. [CrossRef]
29. Guo, L.; Ni, J.; Shi, Y.Q. Uniform Embedding for Efficient JPEG Steganography. *IEEE Trans. Inf. Forensics Secur.* **2014**, *9*, 814–825. [CrossRef]
30. Yedroudj, M.; Comby, F.; Chaumont, M. Yedroudj-Net: An Efficient CNN for Spatial Steganalysis. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 2092–2096. [CrossRef]

31. Wu, S.; Zhong, S.; Liu, Y. A Novel Convolutional Neural Network for Image Steganalysis with Shared Normalization. *IEEE Trans. Multimed.* **2020**, *22*, 256–270. [CrossRef]
32. Zhang, R.; Zhu, F.; Liu, J.; Liu, G. Depth-Wise Separable Convolutions and Multi-Level Pooling for an Efficient Spatial CNN-Based Steganalysis. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 1138–1150. [CrossRef]
33. Xiang, Z.; Sang, J.; Zhang, Q.; Cai, B.; Xia, X.; Wu, W. A New Convolutional Neural Network-Based Steganalysis Method for Content-Adaptive Image Steganography in the Spatial Domain. *IEEE Access* **2020**, *8*, 47013–47020. [CrossRef]
34. Wang, Z.; Chen, M.; Yang, Y.; Lei, M.; Dong, Z. Joint multi-domain feature learning for image steganalysis based on CNN. *EURASIP J. Image Video Process.* **2020**, *2020*, 28. [CrossRef]
35. Bas, P.; Furon, T. Break Our Watermarking System. 2008. Available online: http://bows2.ec-lille.fr/ (accessed on 21 January 2022).
36. Bas, P.; Filler, T.; Pevný, T. "Break Our Steganographic System": The Ins and Outs of Organizing BOSS. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 6958, pp. 59–70.
37. Glorot, Y.B.X. Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.
38. Hannun, A.; Case, C.; Casper, J.; Catanzaro, B.; Diamos, G.; Elsen, E.; Prenger, R.; Satheesh, S.; Sengupta, S.; Coates, A.; et al. Deep Speech: Scaling up end-to-end speech recognition. *arXiv* **2014**, arXiv:1412.5567.
39. Singh, B.; De, S.; Zhang, Y.; Goldstein, T.; Taylor, G. Layer-Specific Adaptive Learning Rates for Deep Networks. In Proceedings of the 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), Miami, FL, USA, 9–11 December 2015; pp. 364–368. [CrossRef]
40. Hermessi, H.; Mourali, O.; Zagrouba, E. Deep feature learning for soft tissue sarcoma classification in MR images via transfer learning. *Expert Syst. Appl.* **2019**, *120*, 116–127. [CrossRef]
41. Alarifi, J.S.; Goyal, M.; Davison, A.K.; Dancey, D.; Khan, R.; Yap, M.H. Facial Skin Classification Using Convolutional Neural Networks. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Berlin/Heidelberg, Germany, 2017; Volume 10317, pp. 479–485.
42. Xu, G. Deep convolutional neural network to detect J-UNIWARD. In Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security, Philadelphia, PA, USA, 20–22 June 2017. [CrossRef]
43. Tang, Y. Deep Learning Using Linear Support Vector Machines. *arXiv* **2013**, arXiv:1306.0239.