# A Novel Steganography Algorithm Based on Instance Segmentation

**Ruohan Meng[1, 2], Qi Cui[1, 2], Zhili Zhou[1, 2], Chengsheng Yuan[1, 2, 3] and Xingming Sun[1, 2, *]**

**Abstract:** Information hiding tends to hide secret information in image area where is rich texture or high frequency, so as to transmit secret information to the recipient without affecting the visual quality of the image and arousing suspicion. We take advantage of the complexity of the object texture and consider that under certain circumstances, the object texture is more complex than the background of the image, so the foreground object is more suitable for steganography than the background. On the basis of instance segmentation, such as Mask R-CNN, the proposed method hides secret information into each object's region by using the masks of instance segmentation, thus realizing the information hiding of the foreground object without background. This method not only makes it more efficient for the receiver to extract information, but also proves to be more secure and robust by experiments.

**Keywords:** Steganography, object detection, mask R-CNN, irregular region.

## 1 Introduction

With the rapid development of science and technology, computer network, multi-media technology and personal mobile communication technology have been widely used. A large number of multimedia information in the network is more convenient and fast transmission, but multimedia information brings convenience to people's lives, personal privacy is snooped, dissemination and other information security issues have gradually attracted people's attention. In recent years, some information security problems have been solved by information hiding technology. This method hides the secret information into the cover media such as image, video, text, sound and so on, which contains certain meanings. By using the redundancy of the cover image, the secret information cannot be detected. The cover images those embed secret information are called stego images. After the sender transmit the stego images to the receiver, the latter extracts the secret information through certain methods. In this process, the stego images and cover images

[1] School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing, 210044, China.

[2] Jiangsu Engineering Centre of Network Monitoring, Nanjing, 210044, China.

[3] Department of Electrical and Computer Engineering, University of Windsor, Windsor, Canada.

[*] Corresponding Author: Xingming Sun. Email: sunnudt@163.com.

Received: 28 December 2018; Accepted: 09 January 2019.

are indistinguishable to the naked eyes. By reducing people's concern about stego images, the security of covert communication can be increased.

In the early days, the most common method for information hiding of images is the least significant bit (LSB) steganography algorithm. The content adaptive steganography algorithms proposed in recent years. Using HUGO [Pevný, Filler and Bas (2010)], S-UNIWARD [Holub, Fridrich and Denemark (2014)], WOW [Holub and Fridrich (2012)] and other steganographic algorithms to selectively hide secret information into texture-rich image regions. Compared with the LSB steganography algorithm, the content adaptive steganography methods have more complex statistical features. At the same time, coverless information hiding is also concerned. Zhou et al. [Zhou, Wu, Yang et al. (2017); Zhou, Yang, Chen et al. (2016); Zhou, Wang, Wu et al. (2017); Zhou, Mu and Wu (2018); Cao, Zhou, Yang et al. (2018)] propose image coverless information hiding methods based on image similarity features. These methods search for the image block corresponding to secret information by hash mapping between image block and secret information to transmit secret information.

In recent years, researchers have combined in deep learning with information hiding to achieve more secure and hidden information hiding. Hayes et al. [Hayes and Danezis (2017)] construct a steganography framework based on Generative Adversarial Networks (GAN). The architecture includes three party: steganography, steganalysis and extraction. Among them, the generator contains steganography method, and the discriminator contains steganalysis method. The stego images generated by generator are adjusted according to the feedback of discriminator. Extractors use the stego images output by generators to extract the secret information. Volkhonskiy et al. [Volkhonskiy, Nazarov, Borisenko et al. (2017)] proposed the Steganographic Generative Adversarial Networks (SGAN) based on Deep Convolutional Generative Adversarial Networks (DCGAN) [Radford, Metz and Chintala (2015)]. It is mainly based on GAN to add a discriminator network for steganalysis of the generated stego images. Shi et al. [Shi, Dong, Wang et al. (2017)] use WGAN to replace DCGAN in SGAN to make the image with higher visual quality and faster model training. At the same time, GNCNN [Qian, Dong and Wang (2015)] is used instead of the steganalysis in Volkhonskiy et al. The resulting image is proved more suitable for steganography. In addition, Liu et al. [Liu, Zhang, Liu et al. (2017)] propose that using the ACGAN generator directly for coverless information hiding. The method divides and expresses secret information into image category information by establishing a mapping dictionary between image categories and text information. Then, the image category information is input into the generator, and an image is generated as a stego, thereby realizing coverless information hiding. In order to ensure the security of the generator method, Ke et al. [Ke, Zhang, Jia et al. (2017)] propose a generator which meets the Kerckhoffs principle, and directly used the key and the cover image as the input of the generator in GAN to generate the stego image.

In addition to using GAN, B. Shumeet [Baluja (2017)] proposes to use convolutional neural network to find suitable steganography positions in images. Then the secret information is embedded by auto-encoder. Wu et al. [Wu, Wang, Shi et al. (2016)] use machine learning method to implement LSB steganography. Meng et al. [Meng, Rice, Wang et al. (2018)] use the method of object detection based on Faster R-CNN, combined with multiple steganography algorithms, to achieve information hiding of multiple steganography methods. Atee et al. [Atee, Ahmad, Noor et al. (2017)] propose the learning of the extreme learning machine (ELM) and the optimal embedding information position. This method can better guarantee the visual effect of the image and has better imperceptibility.

For the method proposed by Meng et al. [Meng, Rice, Wang et al. (2018)], the object texture is more complicated than the background texture under certain circumstances. The secret information is hidden in the foreground object, which improve the security and robustness of the steganography. The steganography method is based on Faster R-CNN and there is still room for improvement. From the experiment, the Faster R-CNN only has a rectangle box for the object, and uses the existing steganography algorithms to hide the secret information in the rectangle box of each object. We have found that through this steganographic algorithm, secret information cannot be completely hidden only in foreground objects. And there is still hidden information in some of the backgrounds in the box. In order to solve this problem, a steganographic algorithm based on irregular shape of fixed area is proposed in this paper. Our method is mainly to further change Mask R-CNN by adding a steganography algorithm in the process to hide only the secret information in the foreground objects. This method, called SMask R-CNN (Steganographic Mask R-CNN), extends Mask R-CNN by adding a part for hiding information on each region of mask, belong to the last step of Fully Convolutional Network (FCN).

## 2 Related works

### *2.1 Information hiding*

Information hiding refers to embedding secret information into the cover by a specific algorithm, without changing the meaning conveyed by the cover, and without being discovered the similarities and differences between the stego images and cover images by the detector. As shown in the Fig. 1, information hiding is mainly divided into three parts, namely the sender, the receiver and the detector. The sender hides the secret messages into the cover image through a steganography algorithm to get the stego image. The secret messages in stego image are extracted by the receiver with an extraction algorithm. For the detector, after the stego image is published, the detector judges whether the image contains secret message by steganalysis algorithm, and gets the result of whether the image is cover image or stego image.

LSB is the simplest and representative steganographic algorithm, whose idea is to replace the minimum bit of an image directly with secret information. When only the lowest bit is replaced, the human eye's perception of small color changes is very low, so it can realize the imperceptibility of embedded information. The least significant bit steganography method has the characteristics of high embedding efficiency and high capacity. The method has generality and can be applied to transform domain.

Generally, for gray images, each pixel is 8 bits. For color images, there are R (red), G (green) and B (blue) channels. Each pixel of each channel has 8 bits. Therefore, each pixel of the color image is 24 bits, and the value of each pixel is 0 or 1. We convert secret information into binary code. By comparing the bit of the LSB of each pixel of the image with the bit of the secret information. If the LSB is the same as the bit of the secret information, then LSB of the pixel will not be modified. Otherwise, the LSB of the pixel will be ±1 randomly.

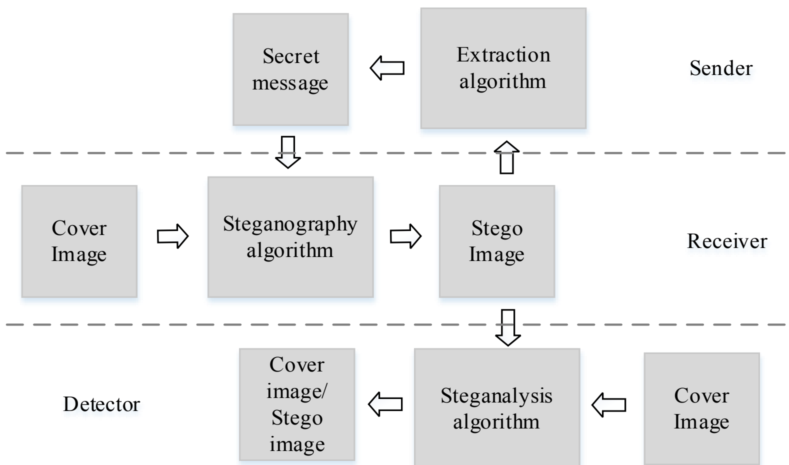



**Figure 1:** The structural diagram of information hiding

This paper combines the latest object detection method (Mask R-CNN) with the most typical information hiding method (LSB). The Mask R-CNN method is used to find a suitable hidden area, and then the secret information is converted into a binary bit. The least significant bit of the selected image portion is modified by the ±1 embedding method, and the secret information is hidden into the image of the fixed area.

Its minimum distortion function is:

$$D\left(X,\hat{X}\right)=\sum_{i=1}^{n1}\sum_{j=1}^{n2}\rho\left(X_{ij},\hat{X}_{ij}\right)\left|X_{ij}-\hat{X}_{ij}\right| \tag{1}$$

where X represents the cover image, $\hat{X}$ is the dense image, and $\rho\left(X_{ij},\hat{X}_{ij}\right)$ is the cost indicator of the X pixel variation for the steganography algorithm.

### *2.2 Object detection*

In the early days, object detection is mainly impleted by using traditional computer vision method. In recent years, with the enhancement of expressive and computational abilities of deep learning, it has been widely used in classification, target detection, medicine, natural language processing and so on.

An early object detection method based on deep learning is the OverFeat algorithm proposed by Sermanet et al. [Sermanet, Eigen, Zhang et al. (2013)], who uses CNNs to process multi-scale sliding windows. After that, the performance of object detection method based on deep learning has been greatly improved. This algorithm is the R-CNN (Regions with CNN features) algorithm proposed by Girshick et al. [Girshick, Donahue, Darrell et al. (2014)]. R-CNN is mainly divided into three parts: Selective Search algorithm is used to extract region proposals in images; CNN is used to extract features from each region proposal; Support Vector Machine (SVM) is used to classify each region proposal. Although the R-CNN algorithm achieves good recognition results, the algorithm is more cumbersome and time-consuming in training. To solve this problem, Kaiming He et al. [He, Zhang, Ren et al. (2015)] propose SPPNet to solve the problem of region proposal computational redundancy. After that, Fast R-CNN [Girshick (2015)] and Faster R-CNN [Ren, He, Girshick et al. (2015)] are proposed in turn to achieve the end-to-end goal. The biggest innovation of Faster R-CNN is to design a region proposal network and integrate multiple detection algorithms into the same network. At present, Faster R-CNN algorithm is a mainstream object detection algorithm. In 2018, the Mask R-CNN proposed by Kaiming He et al. [He, Gkioxari, Dollár et al. (2017)], which is an instance segmentation algorithm. It can complete many tasks, such as object classification, object detection, instance segmentation and so on. The advantages of Mask R-CNN are high speed, high accuracy, simple and intuitive. The algorithm mainly adds a FCN network to realize the accurate segmentation of the target in the image. At the same time, the ROI Align method is proposed to replace the ROI Pooling method, so that the pixels of the original image and the pixels in the feature map are completely corresponding, thus improving the accuracy of detection.

The object detection method mentioned above is based on region proposal's deep learning object detection method. In addition, there are deep learning object detection methods based on regression methods, such as YOLO [Redmon, Divvala, Girshick et al. (2016)], SSD [Liu, Anguelov, Erhan et al. (2016)] and other object detection methods. This type of method mainly uses the idea of regression, and directly obtains the object border and object category of this position through a forward propagation.

Since our proposed algorithm only wants to hide secret information into foreground objects, it needs more accurate instance segmentation. Faster R-CNN and other object detection methods can only achieve border regression and get the bounding box containing objects. The steganography algorithm proposed by Meng et al. [Meng, Rice, Wang et al. (2018)] is not very accurate, which hides secret information in the bounding box containing objects. So on the basis of this work, we use Mask R-CNN to realize accurate information hiding only on foreground objects.

## 3 Proposed method using LSB matching steganographic algorithms based on Mask R-CNN

This part mainly descripts SMask R-CNN in detail. Through SMask R-CNN, we can achieve information hiding in irregular fixed areas. The whole idea of SMask R-CNN algorithm is very simple, that is, on the basis of the original Mask R-CNN, LSB matching steganographic algorithm is added to FCN network to generate the corresponding mask of the stego image, namely Mask R-CNN+LSB maching, or Faster R-CNN+SFCN, more detailed is RPN+ROI Align+Fast Rcnn+SFCN. As shown in Fig. 2, firstly, the image is input into SMask R-CNN network, and the feature maps are extracted through VGG-16 and Feature Pyramid Networks (FPN); secondly, input feature maps into Region Proposal Network (RPN) and select the appropriate proposals; thirdly, proposals through the RoI Align layer, using a bilinear interpolation algorithm, so that the pixels in the original image and the pixels in the feature map are completely aligned, so there is no deviation; finally, in the process of bounding box and classification, Steganography Fully Convolutional Networks (SFCN) is added, and stego images is finally obtained.

The loss function of SMask R-CNN is:

$$L = L_{cls} + L_{box} + L_{mask} \tag{2}$$

Here, $L_{cls}$ represents the loss produced by classification; $L_{box}$ is the loss produced by the regression of bounding box; $L_{mask}$ is the loss of mask's segmentation.

Mask R-CNN mainly adds a mask branch to Faster R-CNN which uses the Fully Convolutional Networks (FCN). FCN is used for image semantic classification by extending image-level classification to pixel-level classification. CNN is composed of convolution, pooling, and fully connected layers, while FCN replaces the fully connected layer in the CNN with a convolutional layer. At the same time, the up-sampling layers are added after the full convolutional layer, that is, the reverse process of pooling. After the pooling layer, the amount of data is reduced. In order not to reduce the amount of data, the number of data is increased by up-sampling, so that each pixel can be predicted and classified. As shown in Fig. 3, it is a structural diagram of the FCN. First, an image with a resolution of H×W is input into the FCN network, and then the resolution of the image is sequentially reduced by 4, 8, 16, and 32 times by the convolutional layers and the pooling layers. The resolution of the feature map obtained in the last layer of the full convolutional layer is 1/32 of the original image. Next, by up-sampling the feature image, a feature map consistent with the resolution of the original image is obtained, and the resolution of the image is also H×W. Finally, a softmax function is used to estimate the probability of different categories, and the larger the corresponding probability value of the pixel, the greater the result of the pixel for this class.

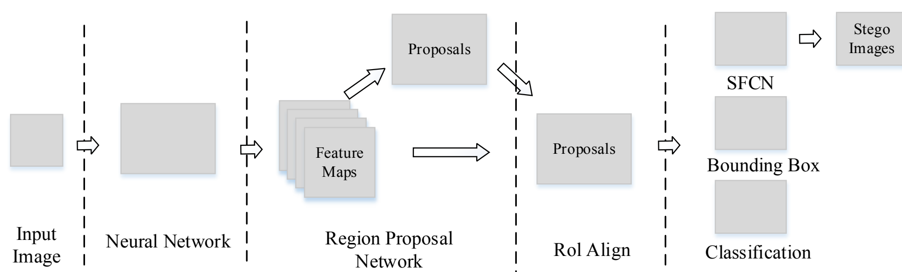



**Figure 2:** The architecture of the SMask R-CNN

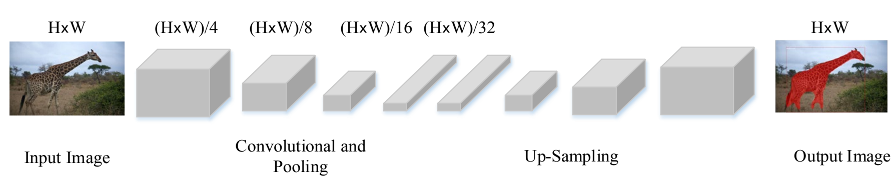



**Figure 3:** The structural diagram of the FCN. By convolution, pooling and up-sampling, the network gets feature maps with the same size as the input image, and then classifies each pixel to get the mask

Since we improve the FCN structure in Mask R-CNN to SFCN, SFCN will be described in detail. The structure of SFCN is shown in Fig. 4. Firstly, the feature map with the same number of pixels in cover image can be obtained by feature selection, proposal determination, and ROI Align. The feature map is fed into SFCN network, and the mask of foreground object is obtained by convolution, pooling, up-sampling and classification. In the process of classifying each pixel, a steganographic algorithm that is LSB matching is added. Finally, the stego image is obtained by embedding secret message.

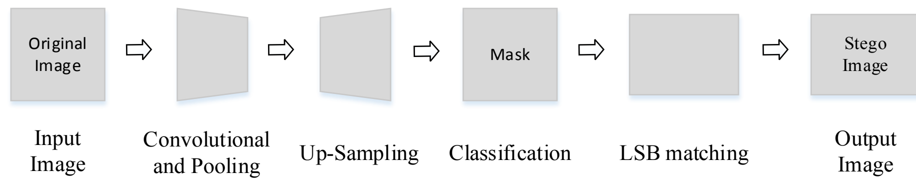



**Figure 4:** The structural diagram of the SFCN. It adds the LSB matching steganography part to the structure of FCN

## 4 Experiments

In this section, steganographic experiments are implement on proposed algorithm and LSB matching algorithm. Two sets of verification experiments are performed to demonstrate the effectiveness of the proposed method. The first group is steganalysis experiment, and the second group focuses on image quality analysis.

All the experiments are conducted on NVIDIA GTX1080 Ti GPU using tensorflow framework. The training data set in the experiments is COCO 2017 dataset. The network used by the image feature extraction section when training Mask R-CNN is VGG16.

### *4.1 Steganography process*

In the steganography process, we conduct two methods, SMask R-CNN steganography algorithm and the LSB matching steganography algorithm. The experimental results obtained from these two steganography algorithms are named SMask_R and LSB_R respectively. We train SMask R-CNN using training dataset of COCO 2017 on 4 GPUs, set batch size to 8. The learning rate is set to 0.02, weight decay and momentum is set to 0.0001 and 0.9. The LSB matching steganography algorithm shares the same image set with SMask R-CNN. We randomly select six images in the test image to show as experimental results. As shown in Fig. 5, the first row is the selected six cover images, the second row is the stego images obtained by the SMask R-CNN steganography algorithm, the third row is the residual image of the stego images and cover images obtained by the SMask R-CNN steganography algorithm, the fourth row is the stego images obtained by the LSB matching steganography method, and the last row is the residual maps of the stego images and cover images obtained by the LSB matching steganography method.

At the same time, we make the comparison of another two sets of residual maps by MSA_ROI steganography in Meng et al. [Meng, Rice, Wang et al. (2018)] and SMask R-CNN steganography algorithm. The experimental results as shown in Fig. 6, the first row is a random selection of six cover images, the second and the third rows are the residual maps between cover images and stego images by MSA_ROI steganography and proposed steganography.

Through experiments, it can be found that secret information can be accurately hidden into irregular objects by SMask R-CNN steganography algorithm, but with LSB matching steganography, secret information is distributed throughout the image, with MSA_ROI steganography, secret information is distributed in rectangular boxes of objects. This experiment can prove that we have achieved the goal of hiding secret information only in irregular objects.

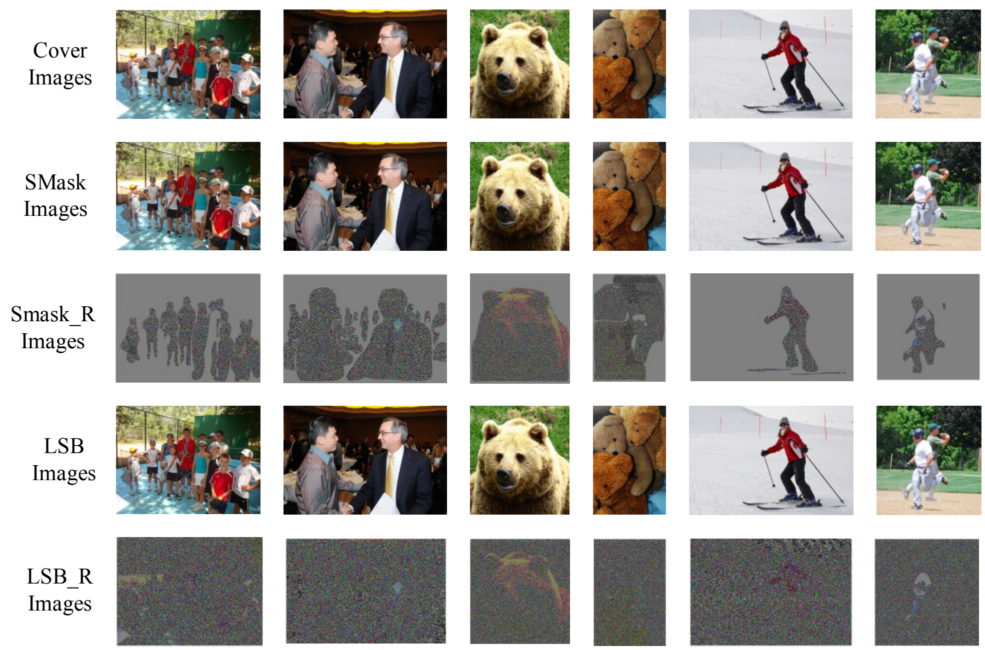



**Figure 5:** Two sets of stego images by SMask R-CNN steganography algorithm and LSB matching steganography. The residual graphs are enhanced

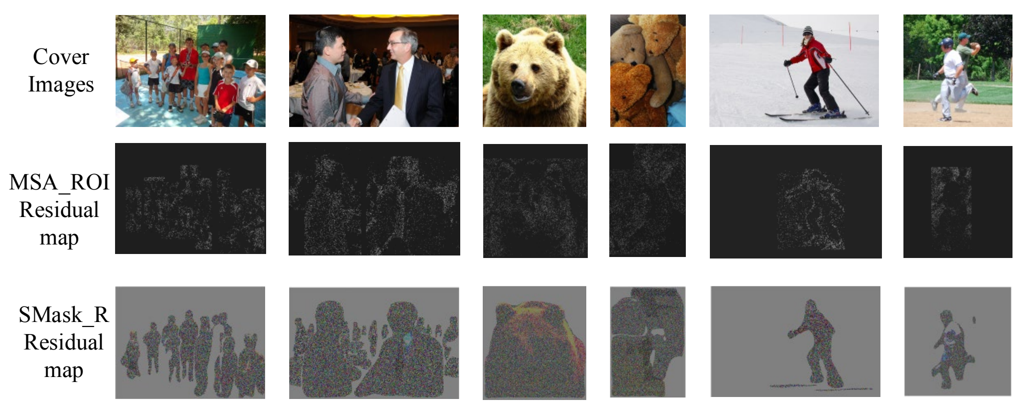



**Figure 6:** The comparison of two sets of residual maps by MSA_ROI steganography and SMask R-CNN steganography algorithm. The residual graphs are enhanced

### 4.2 Steganalysis process

In the steganalysis section, we mainly apply SPA (Sample Pair Analysis) steganalysis method proposed by Dumitrescu et al. [Dumitrescu, Wu and Wang (2002)]. SPA algorithm uses the spatial correlation of the signal for steganalysis to determine whether the image is a stego image. In the experiment, we analysis 5,000 stego images obtained by SMask R-CNN steganography method and 5,000 stego images obtained by LSB matching steganography method, respectively.

**Table 1:** Comparison of the proposed algorithm and LSB matching steganography algorithm in the number of stego images detected by SPA

| | Cover_ | Stego_ | Stego_P | Stego_R | Stego_R_P | Stego_G | Stego_G_P | Stego_B | Stego_B_P |
|---|---|---|---|---|---|---|---|---|---|
| SMask_R | 1642 | 3358 | 0.672 | 2644 | 0.529 | 2521 | 0.504 | 2959 | 0.592 |
| LSB_R | 97 | 4903 | 0.981 | 4371 | 0.874 | 4380 | 0.876 | 4390 | 0.878 |

Through SPA, the steganalysis results of the two sets of images contain a total of four parts: a. the number of images without secret information is named Cover_; b. the number of images containing secret information in the R channel is called Stego_R; c. the number of images containing secret information in the G channel (Stego_G); d. the number of images containing secret information in the B channel (Stego_B). To facilitate the comparison of the experimental results, we subtract the value of Cover_ from the total number of detected images, that is, the number of images containing secret information (Stego_). At the same time, we calculate the proportion of each result in the total number of steganographic images. The proportion of the detected image containing secret information to the total steganographic image is Stego_P. The ratio of the image containing secret information in R channel to the total steganography image is marked as Stego_R_P. Similarly, we can obtain Stego_G_P and Stego_B_P. The experimental results are shown in Tab. 1. In the 5,000 stego images obtained by SMask R-CNN steganography method, 3,358 stego images are detected by SPA algorithm. but the number of images which are detected by the LSB matching steganography method is up to 4,903. In the steganalysis of the three channel, the average value of images in the three channel detected by the SMask R-CNN steganography method is 2,708, but the average value of images in the three channel detected by the LSB matching steganography is about 4,380. Comparing the experimental results, the proposed method is more robust than the traditional LSB matching method.

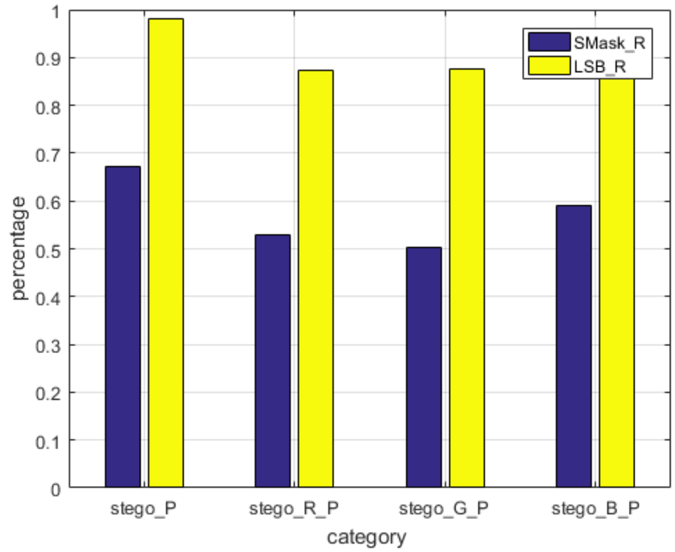



**Figure 7:** The histogram of ratio of the detected image to the total test image by SMask R-CNN steganography method and traditional LSB matching steganography method

**Table 2:** Comparison of the proposed algorithm and the LSB matching steganography algorithm in the ratio of the detected image to the total test image

| | Stego_P | Stego_R_P | Stego_G_P | Stego_B_P |
|---|---|---|---|---|
| SMask_R | 0.6716 | 0.5288 | 0.5042 | 0.5918 |
| LSB_R | 0.9806 | 0.8742 | 0.8760 | 0.8780 |

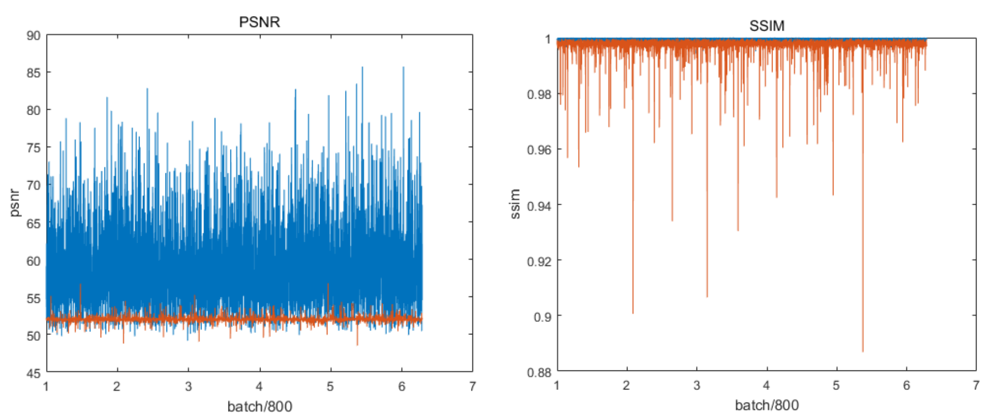
Tab. 2 and Fig. 7 show the results of the ratio of the detected stego images to the total stego images. As shown in Tab. 2, we can see that the detection rate of stego images obtained by the proposed method is 0.6716, while the detection rate of the stego image obtained by the traditional LSB matching steganography method is as high as 0.9806. Almost all stego images obtained by the traditional LSB matching steganography method can be detected. At the same time, in the steganalysis of three-channel, the detection rate of LSB matching steganography is generally about 30% higher than that of SMask R-CNN steganography, which shows that our method has a significant improvement in steganalysis resistance compared with LSB matching method. In Fig. 6 we can also clearly see that the value of SMask_R is higher than the value of LSB_R, that is, the detection rate of LSB matching is significantly higher than that of SMask R-CNN method. It has been proved by steganalysis experiments that our method's ability to resist steganalysis is greatly improved compared to traditional methods.

### *4.3 Stego image quality analysis*

In this section, we implement an image quality analysis experiment on stego images. Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) are applied as evaluating indicator methods. PSNR mainly calculates the error between corresponding pixels, that is, image quality assessment based on error sensitivity. SSIM mainly evaluates image similarity from three aspects: brightness, structure and contrast. The larger the PSNR value is, the smaller the distortion is. The range of SSIM is 0 to 1. The closer the value of SSIM is to 1, the smaller the image distortion is, the better the image quality performance.

In this experiment, 5,000 stego images obtained by SMask R-CNN steganography, 5,000 stego images obtained by LSB matching steganography and 5,000 cover images are used for PSNR and SSIM experiments respectively. The experimental results are shown in Fig. 8, as can be seen from (a), the orange dots are almost all distributed under the blue dots, and the average value of orange dots is about 52, the highest value is about 57, but the average value of blue dots is about 59, the highest value is even about 85. Experiments show that compared with the traditional LSB matching method, the proposed method has better image quality. As can be seen from (b), almost all the blue dots are distributed around 1.00, while the orange dots are distributed differently and generally lower than the blue dots. For the SSIM index, the closer the SSIM value is to 1, the greater the image similarity. Thus, the similarity between stego images and cover images obtained by our proposed method is great, and it is superior to the stego images obtained by LSB matching steganography method.

From the two experiments of PSNR and SSIM, we can conclude that the SMask R-CNN steganography method is more robust than traditional LSB matching steganography method.




(a) (b)

**Figure 8:** (a) and (b) are line charts of the PSNR values and SSIM values of the SMask R-CNN steganography method and traditional LSB matching steganography method. Blue dots represent the proposed method; orange dots represent LSB matching steganography

## 5 Conclusion and future work

This paper mainly proposes to find the foreground object as a selected region of steganography, and then use steganography algorithms to hide the secret information into the irregular foreground object accurately, avoiding the secret information hiding in the background completely. Comparison experiments show that the same steganographic algorithm is used to hide the secret information in the foreground object and all the images. The anti-detection and image quality of the stego image obtained by the former are better than the latter. This method improves the robustness and security of stego images. We will continue to consider the distinguishment of foreground object textures, determine which objects are suitable for hiding information, consider the edge information of objects, and design a more suitable steganography algorithm to hide secret information.

**Acknowledgement:** This work is supported by the National Key R&D Program of China under grant 2018YFB1003205; by the National Natural Science Foundation of China under grant U1836208, U1536206, U1836110, 61602253, 61672294; by the Jiangsu Basic Research Programs-Natural Science Foundation under grant numbers BK20181407; by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) fund; by the Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET) fund, China.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

**Atee, H. A.; Ahmad, R.; Noor, N. M.; Rahma, A. M. S.; Aljeroudi, Y.** (2017): Extreme learning machine based optimal embedding location finder for image steganography. *PLoS One*, vol. 12, no. 2, e0170329.

**Baluja, S.** (2017): Hiding images in plain sight: deep steganography. *Conference on Neural Information Processing Systems*, pp. 2069-2079.

**Cao, Y.; Zhou, Z.; Yang, C. N.; Sun, X.** (2018): Dynamic content selection framework applied to coverless information hiding. *Journal of Internet Technology*, vol. 19, no. 4, pp. 1179-1186.

**Dumitrescu, S.; Wu, X.; Wang, Z.** (2002): Detection of LSB steganography via sample pair analysis. *International Workshop on Information Hiding*, pp. 355-372.

**Girshick, R.** (2015): Fast R-CNN. *IEEE International Conference on Computer Vision*, pp. 1440-1448.

**Girshick, R.; Donahue, J.; Darrell, T.; Malik, J.** (2014): Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580-587.

**Hayes, J.; Danezis, G.** (2017): Generating steganographic images via adversarial training. *Advances in Neural Information Processing Systems 30*, pp. 1-10.

**He, K.; Gkioxari, G.; Dollár, P.; Girshick, R.** (2017): Mask R-CNN. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 99, pp. 1.

**He, K.; Zhang, X.; Ren, S.; Sun, J.** (2015): Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 37, no. 9, pp. 1904.

**Holub, V.; Fridrich, J.** (2012): Designing steganographic distortion using directional filters. *IEEE International Workshop on Information Forensics and Security*, vol. 2, pp. 234-239.

**Holub, V.; Fridrich, J.; Denemark, T.** (2014): Universal distortion function for steganography in an arbitrary domain. *Eurasip Journal on Information Security*, vol. 2014, no. 1, pp. 1.

**Ke, Y.; Zhang, M.; Jia, L.; Su, T.; Yang, X.** (2017): Generative steganography with Kerckhoffs' principle based on generative adversarial networks. arXiv:1711.04916.

**Liu, M.; Zhang, M.; Liu, J.; Zhang, Y.; Ke, Y.** (2017): Coverless information hiding based on generative adversarial networks. arXiv:1712.06951.

**Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S. et al.** (2016): SSD: single shot multibox detector. *European Conference on Computer Vision*, pp. 21-37.

**Meng, R.; Rice, S. G.; Wang, J.; Sun, X.** (2018): A fusion steganographic algorithm based on Faster R-CNN. *Computers, Materials & Continua*, vol. 55, no. 1, pp. 1-16.

**Pevný, T.; Filler, T.; Bas, P.** (2010): Using high-dimensional image models to perform highly undetectable steganography. *International Workshop on Information Hiding*, pp. 161-177.

**Qian, Y.; Dong, J.; Wang, W.** (2015): Deep learning for steganalysis via covolutional neural networks. *Proceedings of SPIE-The International Society for Optical Engineering*, vol. 9409.

**Radford, A.; Metz, L.; Chintala, S.** (2015): Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv:1511.06434.

**Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A.** (2016): You only look once: unified, real-time object detection. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779-788.

**Ren, S.; He, K.; Girshick, R.; Sun, J.** (2015): Faster R-CNN: towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems*, pp. 91-99.

**Sermanet, P.; Eigen, D.; Zhang, X.; Mathieu, M.; Fergus, R. et al.** (2013): Overfeat: integrated recognition, localization and detection using convolutional networks. arXiv:1312.6229.

**Shi, H.; Dong, J.; Wang, W.; Qian, Y.; Zhang, X.** (2017): SSGAN: secure stganography based on generative adversarial networks. *Pacific Rim Conference on Multimedia*, pp. 534-544.

**Volkhonskiy, D.; Nazarov, I.; Borisenko, B.; Burnaev, E.** (2017): Stganographic generative adversarial networks. arXiv:1703.05502.

**Wu, H. Z.; Wang, H. X.; Shi, Y. Q.** (2016): Can machine learn steganogrphy? Implementing LSB substitution and matrix coding steganography with feed-forward neural networks. arXiv:1606.05294.

**Zhou, Z.; Mu, Y.; Wu, Q.** (2018): Coverless image steganography using partial-duplicate image retrieval. *Soft Computing*, vol. 2, pp. 1-12.

**Zhou, Z.; Wang, Y.; Wu, Q.; Yang, C.; Sun, X.** (2017): Effective and efficient global context verification for image copy detection. *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 48-63.

**Zhou, Z.; Wu, Q.; Yang, C.; Sun, X.; Pan, Z.** (2017): Coverless image steganography based on histograms of oriented gradients-based hashing algorithm. *Journal of Internet Technology*, vol. 18, no. 5, pp. 1177-1184.

**Zhou, Z.; Yang, C.; Chen, B.; Sun, X.; Liu, Q. et al.** (2016): Effective and efficient image copy detection with resistance to arbitrary rotation. *IEICE Transactions on Information and Systems*, vol. E99-D, no. 6, pp. 1531-1540.