IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 27, 2020, accepted September 11, 2020, date of publication September 15, 2020,
date of current version September 25, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3024193

High-Capacity Image Steganography Based on Improved FC-DenseNet

XINTAO DUAN[1], LIU NAO[1], GOU MENGXIAO[1], DONGLI YUE[1], ZIMEI XIE[1], YUANYUAN MA[1], AND CHUAN QIN[2], (Member, IEEE)

[1]College of Computer and Information Engineering, Henan Normal University, Xinxiang 453007, China
[2]School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

Corresponding author: Xintao Duan (duanxintao@htu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61672354, in part by the Key Scientific Research Project of the Henan Provincial Higher Education under Grant 19B510005 and Grant 20B413004, and in part by the Science and Technology Project of the Henan Provincial Department of Science and Technology under Grant 202102210165.

ABSTRACT Aiming at the problem that the traditional steganography based on carrier modification has the low steganographic capacity, a steganographic scheme based on Fully Convolutional Dense Connection Network (FC-DenseNet) is proposed. Since FC-DenseNet can effectively overcome the problems of gradient dissipation and gradient explosion, and a large number of features are multiplexed, the cascaded secret image and carrier image can reconstruct good image quality after entering the network. Effectively improve steganographic capacity. First, we reset the number of input channels of the first convolution block of FC-DenseNet and the number of output channels of the last convolution block and deleted the LogSoftmax() function. On the sender side, after the concatenated secret image and carrier image pass through the hidden network FC-DenseNet, the secret image is embedded in the carrier image to obtain a stego-image. At the receiving side, the extraction network reconstructs the secret image from the stego-image. Experimental results show that our proposed steganography scheme not only has a high Peak Signal-to-Noise Ratio(PSNR) and Structural Similarity(SSIM) but also can realize large-capacity image steganography, with an average image payload capacity of 23.96 bit per pixel.

INDEX TERMS Convolutional neural network, image steganography, deep learning, FC-DenseNet.

I. INTRODUCTION

Cyberspace security has always attracted people's attention. While the number of global netizens continues to grow, network security is in a cycle of decreasing and increasing. How to effectively ensure that information is not intercepted by attackers in the process of Internet communication is crucial [1]. In addition, personal privacy leakage is also an important issue today. For example, the photographed electronic ID cards and bank cards were stolen by criminals and used for illegal purposes. It not only endangers people's property safety, but also violates personal privacy.

The most common steganography is to embed secret information into the Least Significant Bits(LSB) of natural images, LSB digit of the pixel is easily replaced by secret information, which is easy to produce visual artifacts. In the case of known carrier images and stego-image, it is very easy for an attacker to find secret information in stego-image. In [2], Cheonshik Kim *et al.* proposed a lossless data hiding method for Absolute Moment Block Truncation (AMBTC) images. In order to increase the steganographic capacity, the coefficients of the bit planes in each 4×4 block are calculated, and then the histogram modification is effectively expanded, which improves the steganography effect to a certain extent. In order to further optimize the results of [2], Dongkyoo Shin *et al.* in [3] added Hamming codes and lookup tables to optimize codewords, which are mainly used to reduce the distortion of image pixels and improve image quality. In [4], Vinodhini RE *et al.* used DNA algorithms to insert secret information into DNA sequences to complete image steganography. In [5], Lin Weixiong *et al.* proposed a blind watermarking algorithm based on the significant difference between wavelet coefficient quantization and copyright protection. Each of the seven non-overlapping wavelet synergies of the carrier image is grouped into a block, which mainly quantizes the local maximum coefficient to determine the

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

position of the watermark insertion. In [6], Tomáš *et al.* proposed an adaptive steganographic lattice code quantization method that minimizes the embedding impact. STC is usually used in combination with the cost. STC coding and cost calculation are divided into two processes. The cost is calculated first (Assuming the effect of distortion after embedding information), STC coding is performed. Currently, based on the method of modifying the carrier image [4]–[6], it mainly includes modifying the spatial domain of the image and modifying the frequency domain of the image. The change of the image spatial domain is to directly modify the pixel value of a certain area or a certain area to achieve expected goals, and the image frequency domain changes by modifying the transform domain coefficients to achieve information embedding. We are different from traditional image steganography. Traditional image steganography is a cost function designed by hand and embedded with STC class. The cost function is designed by ourselves based on our own experience. This is a difficulty in STC coding. Due to the great potential of deep learning [7], it has achieved great success in pattern recognition, natural language processing, data mining, etc., and it has also penetrated and developed in the field of information security and forensics. Therefore, introducing deep learning into the field of steganography and using artificial intelligence to design the cost function is a novel solution to the difficulty of manually designing the cost function.

Text information is hidden into the image in [8]–[10]. In the generative adversarial network of [9] and [10], the stego-image image generated by the generator makes the discriminator unable to distinguish between the carrier image and the stego-image. At the same time, it is impossible to accurately distinguish whether the image is a carrier image or a stego-image in our human vision, so it is difficult for us to find that the image contains secret information. In [10], the secret information is segmented and converted into random noise. The noise enters the generator and the output result is a secret image, and the secret image is extracted by the extractor and the output result is a noise vector. Finally, according to the mapping of secret information and noise vector relationship, reconstruct the secret information. To increase security, in [11], the attention model and the generative adversarial network are combined to realize the embedding and extraction of secret images. Nowadays, there are relatively few attempts to hide images using the deep neural network itself [12]–[14]. In [13], the convolutional network has an encoder and a decoder. The encoder is mainly used to hide the color secret image, and the decoder is mainly used to extract the secret image. The convolutional network implements the hiding and extraction of secret images, but after hiding the stego-image, we can easily see the distortion of setgo-image. Most of them are steganographic text [4]–[6], [8]–[10]. The number of bits per pixel occupied by the text is relatively small in the image, so these methods have the advantages of steganalysis good resistance, but low payload capacity. A good information hiding challenge arises because the appearance and underlying statistical changes of the carrier image are easily caused by embedding the message. There are two main reasons for the change in the appearance of the carrier image and the statistics of the basic data. First, the amount of information that will be hidden. The most common is to hide relatively few bits of text information hidden, such as in [15]. Second, the extent of the visible change in the image of the carrier depends on the image of the carrier itself. Hiding secret information in the high-frequency region of the image is better than hiding the information in artificially detectable low-frequency regions. The capacity for estimating information hiding can be found in [16].

In March 2020, China's Internet penetration rate was 64.5% and the number of Internet users was 904 million. The problem of personal privacy leakage continues to arise. For example, for some private photos stored by the user, the user can embed the private photos into a natural picture through the scheme we designed. In [17]–[19], the text is converted to binary data, and then the position of the LSB is automatically selected by the neural network for embedding. In contrast, in our work, the difference between us and [20] is that the preparation network is removed, and the hidden network uses Full Convolution Dense Connection Network (FC-DenseNet) [21]. In our work, first, the secret image and the carrier image are encoded as a stego-image via a hidden network, and secondly, the stego-image is decoded into a secret image via a decoding network. Ultimately, our proposed solution not only improves steganographic capacity but also has a high Peak Signal to Noise Ratio (PSNR) [22] and the Structural Similarity Index (SSIM) [23].

Perhaps the best neural network is used at the same time as the work presented here [24]–[26]. In addition, in standard steganography studies, these methods encode small amounts of information but are visual of good quality. As a summary, our work consists of the following three parts:

–In [21] the image is segmented by semantics, and the segmentation effect is very good. Unlike [21], we applied FC-DenseNet for information hiding for the first time. The number of input channels of the first convolution block of FC-DenseNet and the number of output channels of the last convolution block is reset. In addition, we did not use the object tag category and deleted LogSoftmax().

– Hidden secret information does not need to be perfectly coded and can accept small errors. It can clearly balance the reconstruction quality of the carrier image and the secret image, as shown in Figure 1.

– Unlike an encrypted noise image or an image that is obviously visually impacted after adding secret information. Instead of the images we transmit are meaningful images, and large orders of secret information are hidden, the ratio of the magnitude of the secret image to be hidden to the image of the carrier is 1:1, payload capacity reaches 23.96bpp.

II. RELATED KNOWLEDGE

Deep networks have achieved unprecedented success in many image segmentation tasks such as [21], [27]. In this article, we mainly apply FC-DefnsNet to our hidden network to

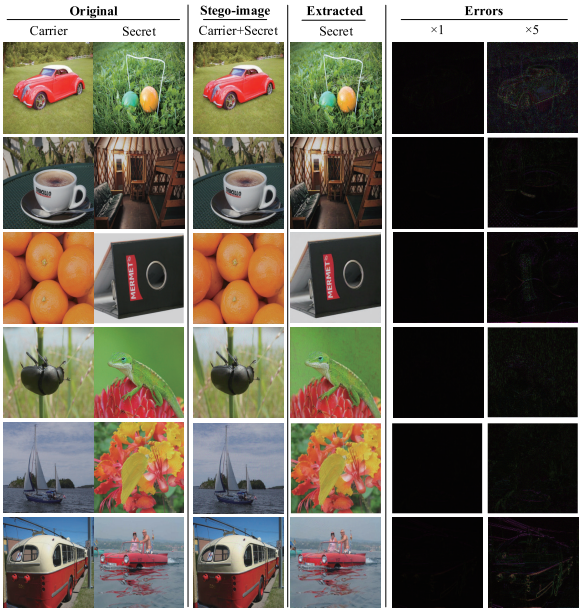


FIGURE 1. Random samples from the Encoder-Decoder (hidden and extracted) system. Starting from the left: carrier image, secret image, the stego-image (It not only looks like a carrier image, but also hides the secret image in it), and the recovered secret image——this is extracted from only the stego-image. Columns 5 and 6 are the errors for the stego-image vs. carrier (enhanced ×1, ×5).

hide information. FC-DenseNet is extended from Densely Connected Convolutional Networks (DenseNet) [28]. To utilize deep networks in image steganography, we will briefly introduce the application of DenseNet and FC-DenseNet to lay the foundation for the next section.

A. DENSE CONVOLUTIONAL NETWORK (DenseNet)

DenseNet is not only used for the classification of data but also for the super-resolution of images and image segmentation. In [29], Tong Tong al. proposed a method to apply DenseNet to image super-resolution. After the image passes DenseNet, it is followed by deconvolution and reconstruction, which have achieved good results. In view of the fact that the traditional convolutional network has an L layer, there is an L number of connections. However, there is a connection between each layer of the DenseNet network and each subsequent layer, and the total number of connections is $L(L+1)/2$.

The advantage of DenseNet is to eliminate the vanishing-gradient problem, enhance feature propagation, a large number of features are multiplexed, and can effectively reduce the parameters of the model. The structure of DenseNet is shown in Figure 2. Figure 2 briefly depicts the layout of DenseNet. Consequently, the l^{th} layers accept all layers in front of feature maps, $X_0, X_1, \ldots, X_{l-1}$ as input:

$$X_l = Y_l([X_0, X_1, \ldots, X_{l-1}]). \tag{1}$$

where $[X_0, X_1, \ldots, X_{l-1}]$ represents the concatenation of the feature maps generated in the 0 to $l-1$ layers. Combine multiple tensors into one tensor for better application. is a composite function with three operations: namely

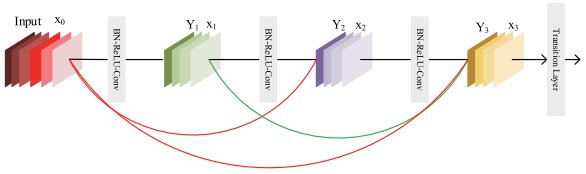


FIGURE 2. A 4-layer dense block with hyperparameter $k = 4$. The hyperparameter determines how much information each layer contributes globally.

Batch Normalization (BN) and Rectified Liner Unit (ReLU), Convolution (Conv).

B. APPLICATION AND DEVELOPMENT OF FC-DenseNet

FC-DenseNet, as the name suggests, is a dense convolutional neural network. It is derived from DenseNet, FC-DenseNet is mainly used for semantic segmentation, according to the label, it can segment the sky, tree, vehicle, pedestrian, road, etc. in the picture, and also applies to the segmentation in the video. The main process is to add a downsampling and upsampling function behind the Dense Block. The downsampling is a 2×2 convolution block used to extract the maximum value of the feature map to reduce the image and simplify the calculation. The upsampling is composed of a 3×3 transposed convolution, and the convolution kernel has a step size of 2 to compensate for the pooling operation. Finally, the classification is done by 1×1 convolution and the number of categories. More straightforward, the number of output feature-maps is the number of different categories, and the effect of the segmentation is attractive.

Based on FC-DenseNet, it is not necessary to split different objects in the image but to hide information in our work. We did not mark the different objects in the image, deleted the label category and the 1×1 convolution to achieve perfect hiding of the information.

In the next section, we will describe how to train both the encoder and the decoder to hide secret pictures and recover secret pictures, as well as the structure of the encoder and decoder. Quantitative assessment and visual assessment will be carried out in Section IV. Conclusions are present in Section V.

III. PROPOSED IMAGE STEGANOGRAPHY SCHEME

This part mainly describes the detailed structure of hidden network and extraction network, and how to hide secret information and reconstruct secret information. In addition, the loss function we designed is introduced in detail.

A. NETWORK ARCHITECTURE

As depicting Figure 3, our proposed architecture mainly contains concatenation and an encoder (Hiding Network) and a decoder (Reveal Network). Concatenation can be described as let $c \in R^{W \times H \times D}$ and $s \in R^{W \times H \times D'}$ be two tensors of the same width and height and depth, D and D', where c and s are the carrier image and the secret image, respectively;

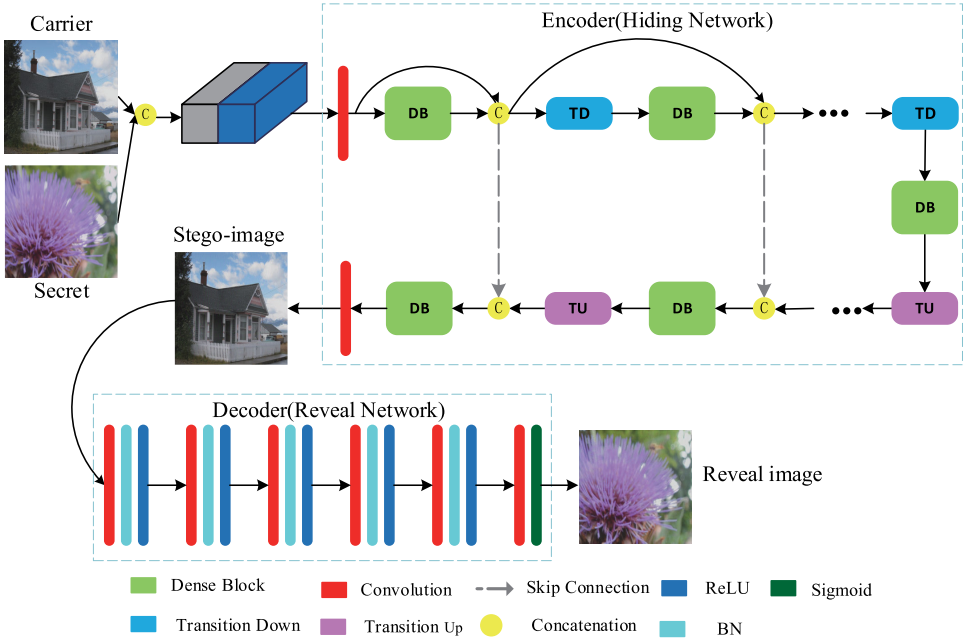


FIGURE 3. Framework for steganography model diagram.

then, let concatenation: $(c, s) \rightarrow \varphi \in R^{W\times H\times(D+D')}$ be the concatenation of the two tensors along the depth axis. The image size requirements input into the framework: $W \geq 256$, $H \geq 256$, $D = 3$, uniformly adjusted to 256×256 through the image size function in python. The encoder is mainly used to hide secret information and the decoder is mainly used to extract secret information. Our goal is to train encoders and decoders at the same time to achieve perfect information hiding and extraction.

B. HIDE ENCODER OF THE SECRET IMAGE

The encoder was previously a concatenation operation. This function spliced the carrier image and the secret image and output a 6-channel tensor to prepare for the next full-size image hiding. The structure of the encoder is given in Table 1. In Table 1, the first column indicates the architecture of the hidden network, The second column indicates the size of the output feature map, D is the number of channels. The input to the encoder is a 6-channel tensor and the output is a 3-channel tensor. The foremost goal of the encoder is to encode a tensor with a channel number of 6 and the output stego-image should be as close as possible to the carrier image. Encoder components include two 3×3 convolution, 11 Dense Block (DB), 10 Concatenation, 5 Transition Down (TD) and 5 Transition Up (TU). The input of the first convolution block is a 6-channel tensor and the output is a 48 channel tensor. The input of the last convolution block is a 192 channels tensor, and its output is a 3-channel tensor. TD is composed of BN, ReLU, 1×1Conv, 2×2maximum pooling. TU is a 3×3 Transposed Convolution *strides* $= 2$, *padding* $= 1$. TD is mainly used to extract and integrate the features of secret images and carrier images. The role of TU is to gradually restore the characteristics of the input. The combination of error τ and TU can restore the carrier image with high accuracy. In Figure 4, each DB consists of 4 layers, each layer including batch normalization, activation, 3×3Conv, *Droupt* $= 0$. Our goal is to minimize the loss between stego-image and carrier image:

$$\tau = \|c - c'\|. \tag{2}$$

where c and c' represent the carrier image and stego-image.

TABLE 1. Architecture of Hidden networks.

Step	*Feature-map(W×H×D)*
3×3Conv	256×256×48
4(BN+ReLU+Conv)+TD	128×128×96
4(BN+ReLU+Conv)+TD	64×64×144
4(BN+ReLU+Conv)+TD	32×32×192
4(BN+ReLU+Conv)+TD	16×16×240
4(BN+ReLU+Conv)+TD	8×8×288
4(BN+ReLU+Conv)	8×8×336
TU+4(BN+ReLU+Conv)	18×18×372
TU+4(BN+ReLU+Conv)	32×32×336
TU+4(BN+ReLU+Conv)	64×64×288
TU+4(BN+ReLU+Conv)	128×128×240
TU+4(BN+ReLU+Conv)	256×256×192
1×1Conv	256×256×3

C. REVEAL DECODER OF THE STEGO-IMAGE

In Table 2, the input to the decoder is a 3-channel tensor and the output is also a 3-channel tensor. The Decoder components include 6 Conv with a convolution core of 3×3, 5 BN, 5 ReLU, and 1 Sigmoid. The main goal of the Decoder is to decode the secret image s from the stego-image and the

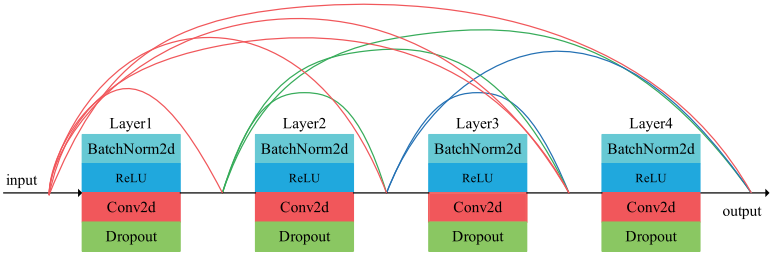


FIGURE 4. A Dense connection Block has four layers. Layers and layers are connected to each other.

TABLE 2. Architecture of Reveal Networks.

Step	*(padding, stride)*	*Feature-map(W×H×D)*
3×3Conv+BN+ReLU	(1,1)	256×256×64
3×3Conv+BN+ReLU	(1,1)	256×256×128
3×3Conv+BN+ReLU	(1,1)	256×256×256
3×3Conv+BN+ReLU	(1,1)	256×256×128
3×3Conv+BN+ReLU	(1,1)	256×256×64
3×3Conv+Sigmoid	(1,1)	256×256×3

decoded s' should be as similar to the original secret image as possible. Our goal of the decoder is to minimize the loss of the decoded secret image and the original secret image:

$$\gamma = \left\| s - s' \right\| \tag{3}$$

where s and s' represent the original secret image and the reveal secret image, respectively.

D. LOSS FUNCTION

In our work, the target of the encoder is to hide the secret image into the carrier image to get the stego-image, the decoder extracts secret images from stego-image. The training of the entire network is done by reducing the following errors: (β is their weight):

$$\zeta = \tau + \beta\gamma \tag{4}$$

Note that as the error τ changes, it has no impact on the weight of the decoder, only the weight of the encoder. Because the decoder does not need to reconstruct the cover image; it only needs to recover secret information from stego-image. Both the encoder and the decoder are trained from the error signal $\beta\gamma$ since the hidden network and the extracted network are responsible for saving and forwarding information about secret images. Better hiding and extracting information by propagating error signals.

IV. EXPERIMENTS

In this article, the dataset is taken from the public ImageNet image database, and we randomly selected 23,500 images, of which 20,000 were used only for training, and 3,500 were used for testing. The hyperparameter $\beta = 0.75$ of the model, the number of input channels of the first convolution block of the hidden network is 6, and the number of output channels of the last convolution block is 3. The initial learning rate of the network is 0.001. The batch into which the image enters the model is set to 16, and the number of iterations of the training is set to 200. Experimental hardware configuration: 2 GPUs, GPU model is GeForce GTX1080, 1 GPU memory is 8G. Software environment required for the experiment: deep learning framework Pytorch1.1.0, programming experiment with python3.6. For performance evaluation, we used two assessment methods: visual evaluation and quantitative evaluation. In addition, steganalysis tools are introduced to detect the anti-steganography security of our proposed scheme. Image distortion is evaluated by the SSIM and the PSNR. In addition, the space occupied by our model parameters is 8.24M.

A. VISUAL EVALUATION

In check to see the performance of the model, we performed a human visual assessment of the carrier image, the secret image, and the stego-image. In Figure 5, after comparison, our eyes do not see the difference between the carrier image and the stego-image. The stego-image image is subtracted from the carrier image to obtain an error image. In the 1, 2, 3, 5, and 7 lines of Figure 5, after the error image is magnified 10 times and 20 times, we can only see artifacts of the carrier image without seeing the secret image artifacts.

To further verify the generalization ability of the model, we selected several images from the CelebA dataset and the COCO dataset for testing. The results of the test are shown in Figure 6. The high-frequency region of the carrier image has a strong anti-interference ability, so the artifact of the carrier image appears in the error image. In the 4th line and the 6th line of Figure 5, in the 7th line of Figure 6, after the error image is magnified 10 times and 20 times, we only see the artifacts of the secret image blurred because the area of the carrier image is relatively flat.

B. QUANTITATIVE EVALUATION

In addition to visual assessment, we also measure the statistical mathematical distribution quality of stego-image. One widely-used metric for measuring image quality is the PSNR. PSNR is mainly used to measure the distortion rate of an image and display it as a score. Given two images X and Y of size (W, H), the PSNR is defined as a function of the Mean

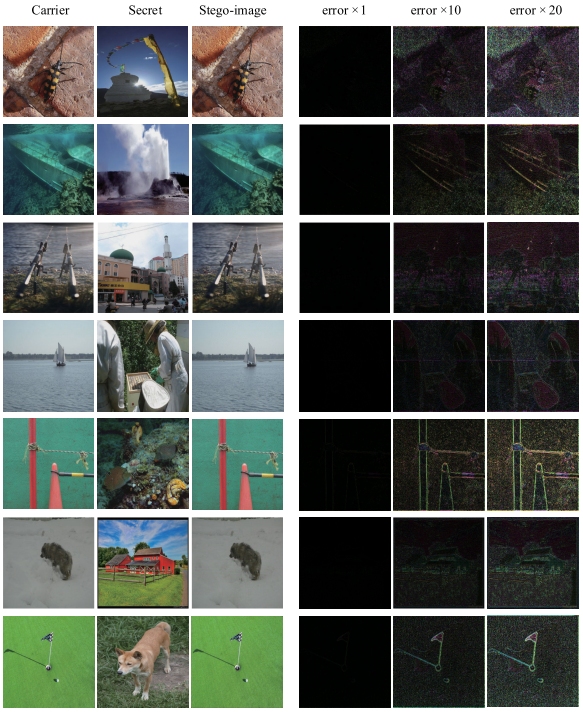


FIGURE 5. Starting from the left: carrier image, secret image, the stego-image. The last three columns are the errors for the setgo-image vs. carrier (enhanced ×1, ×10, ×20).

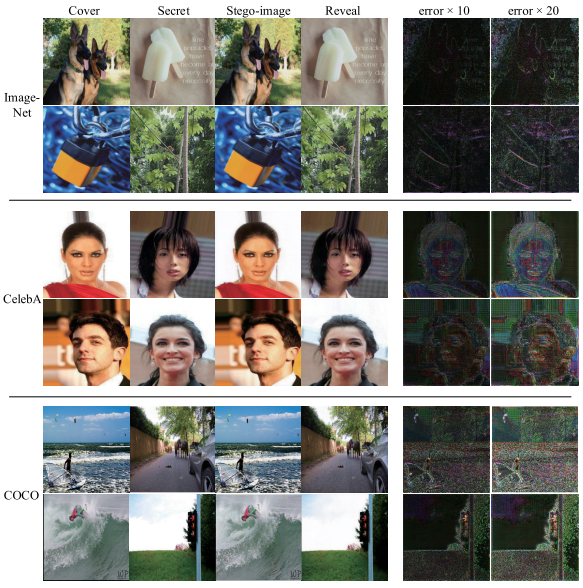


FIGURE 6. Sarting from the left: carrier image, secret image, the stego-image (stego-image not only contains secret images, but also looks like a carrier image), and the recovered secret image this is to say extracted from only the stego-image. 1 row and 2 rows, 3 rows and 4 rows, 5 rows and 6 rows respectively of ImageNet dataset, CelebA dataset, COCO dataset.

Squared Error (MSE):

$$MSE = \frac{1}{WH}\sum_{i=1}^{W}\sum_{j=1}^{H}(X_{i,j} - Y_{i,j})^2 \tag{5}$$

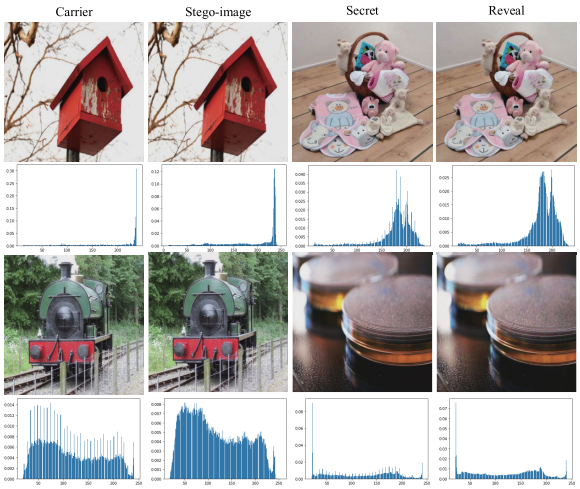


FIGURE 7. From the left: carrier image, stego-image, secret image, and the recovered secret image. The 2th and 4th line histograms correspond to lines 1 and 3, respectively.

$$PSNR = 10\log_{10}\frac{(2^n - 1)^2}{MSE} \tag{6}$$

where 2^{n-1} is the maximum value of the image point color and n is the number of bits of the sample point.

For a more complete evaluation of stego-image, we declare the SSIM between the carrier image and the stego-image. Given two images X and Y, the SSIM could be computed using the means, μ_X and μ_Y, Variances, σ_X^2 and σ_Y^2, and covariance σ_{XY}^2 of the images as shown below:

$$SSIM = \frac{(2\mu_X\mu_Y + k_1R)(2\sigma_{XY} + k_2R)}{(\mu_X^2 + \mu_Y^2 + k_1R)(\sigma_X^2 + \sigma_Y^2 + k_2R)} \tag{7}$$

hereby default $k_1 = 0.01$, $k_2 = 0.03$, and the return value is between [-1, 1], where 1.0 means the two images are identical.

To further illustrate the effectiveness of steganography, in Table 3, we introduce the carrier image and the stego-image, the extracted secret image, and the PSNR and SSIM of the original secret image. In the third row of Table 3, we can see that the average of PSNR and SSIM for the carrier image and the stego-images reached (39.556, 0.985), and the average of PSNR and SSIM of the reconstructed secret image and the original secret image reached (37.092, 0.975). Table 4 reports a comparison of our approach to other solutions, and our PSNR and SSIM are the highest. Besides, in Figure 7, we plot the carrier image, stego-image, secret image, and the histogram of the reconstructed secret image, we can observe the carrier image and the stego-image, the secret original image and the reconstructed secret image, and the histogram difference between them is not very large. It is worth noting that because these evaluation indicators are relatively good, the solution we designed will not destroy the visual integrity.

C. STEGANOGRAPHY CAPACITY ANALYSIS

The traditional LSB steganography has a relatively small steganographic capacity. Since our steganography scheme is

TABLE 3. Compare Carrier Image and Stego-Image,Secret Image and Reconstructed Secret Image PSNR and SSIM.

Figure	*Carrier vs stego-image (PSNR, SSIM)*	*Reconstructed secret vs. secret (PSNR, SSIM)*
Fig.7 row1	40.196, 0.995	36.185, 0.984
Fig.7 row3	40.799, 0.998	37.331, 0.982
ImageNet (Average)	39.556, 0.985	37.092, 0.975

TABLE 4. We compare the PSNR and SSIM of the scheme [10], [11],and we can see that our indicators are higher than theirs.

Schemes	*Carrier vs stego-image (PSNR, SSIM)*	*Reconstructed secret vs. secret (PSNR, SSIM)*
[12]	32.9, 0.96	36.6, 0.96
[13]	34.89, 0.9681	33.42, 0.9474
ours	39.556, 0.985	37.092, 0.975

TABLE 5. The Steganographic Capacity Comparison Result.

Schemes	*Absolute capacity (bytes/image)*	*Stego-image size*	*Relative capacity (bytes/pixel)*
[10]	≥ 37.5	64×64	9.16e-3
[30]	1.125	512×512	4.29e-6
[31]	3.72	$\geq 512 \times 512$	1.42e-5
[32]	1533 4300	1024×1024	1.46e-3 4.10e-3
[33]	0.375	32×32	3.7e-4
[34]	64×64	800×800	6.04e-3
Ours	256×256	256×256	1

relatively new, it is more intuitive to compare with other schemes, there is a hiding scheme including the steganography based on carrier selection and the steganography based on carrier synthesis. The results of the comparison are shown in Table 5, we can clearly see that the payload capacity of our proposed scheme is better than other schemes. Here, column 1, column 2, column 3, column 4 are steganography, steganography capacity per image, stego-image size, relative steganography capacity (steganography capacity per pixel):

$$Relative\ capacity = \frac{Absolute\ capacity}{Image\ size} \tag{8}$$

In addition, our proposed model is suitable for most images. As shown in Figure 8, we introduce CT images, remote sensing images, and images captured by aerial vehicles. The load capacity and the change rate of the carrier image in Figure 8 are shown in Table 6. The extraction rate (r) of the extraction network is calculated by the following formula:

$$r = 1 - \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} \left| s_{i,j} - s'_{i,j} \right| \tag{9}$$

The change rate of the carrier image (cgr) is calculated by the following formula:

$$cgr = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} \left| c_{i,j} - stego_{i,j} \right| \tag{10}$$

TABLE 6. Carrier image change rate and load capacity results.

FIGURE	*cgr*	*payload capacity(Bpp)*
FIGURE 8 Column 1	0.12%	23.97
FIGURE 8 Column 2	0.12%	23.97
FIGURE 8 Column 3	0.11%	23.95
FIGURE 8 Column 4	0.13%	23.96
FIGURE 8 Column 5	0.11%	23.96
FIGURE 8 Column 6	0.10%	23.97
Average	0.115%	23.96

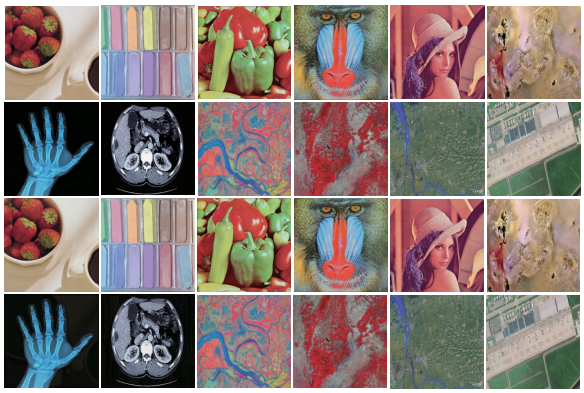


FIGURE 8. Experimental images of CT images, remote sensing images, and images captured by aerial cameras. Lines 1, 2, 3 and 4 are carrier image, secret image, stego-image, and extracted secret image respectively.

FIGURE 9. The first line is the carrier image, the second line is the secret image, the third line is stego-image, and the fourth line is the secret image extracted. The carrier image and the secret image in the first column are the original images. The carrier image and the secret image in the second, third, fourth, and fifth columns are the images processed by the JPEG high-quality compression factor, and the quality factors are 70, 75, 80, and 85 respectively.

The calculation method of payload capacity is:

$$payload\_capacity = r \times 8 \times 3 \tag{11}$$

where 8 indicates that the space occupied by a single pixel is 8, and 3 indicates that each image has 3 channels.We randomly selected 500 pairs of images as the test set, and the calculated average load rate was 23.96bpp.

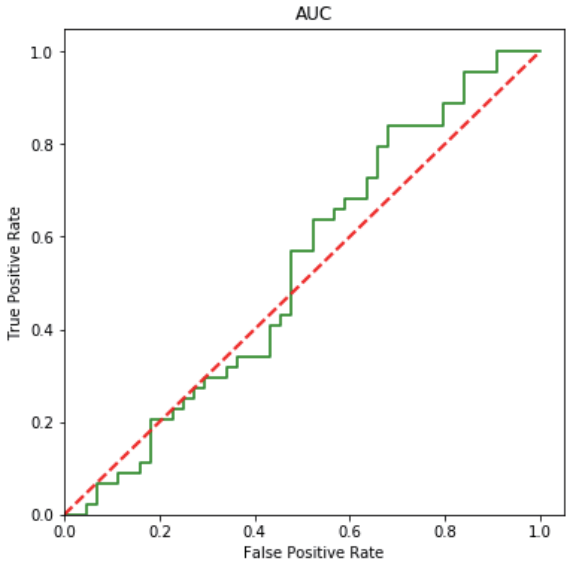


FIGURE 10. Get ROC curve with StegExpose analysis tool.

To verify the influence of image compression on our model, before the image enters our model, the carrier image and the secret image are compressed separately with JPEG high-quality factors. The final experimental results are shown in Figure 9. We can hardly see the difference between them. When the quality factor is 70, we cannot see the difference between them and the original image experimental results.

Finally, we use the steganalysis tool StegExpose [35] to analyze our data, using a standard threshold of 0.2, and the analysis results are shown in Figure 9. The horizontal axis indicates that a file that is not organized is judged as a steganographic file, and the vertical axis indicates that an organized file is judged as a steganographic file. The red dashed line represents random guessing, and the green solid line represents the ROC drawn. It can be seen that the performance of the steganalysis tool is only slightly better than random guessing.

V. CONCLUSION

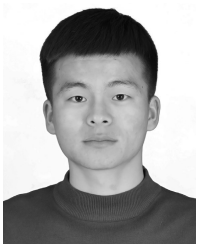
This article is different from the traditional STC coding steganography scheme. We use artificial intelligence to design the cost function instead of the cost function manually designed by STC coding. Through the end-to-end training model, the full-size image is hidden and the image distortion is small. Experimental results show that the method has advanced visual effects and high steganography capacity, and the model has strong generalization ability, which can achieve steganography and extraction of different data sets. In the next step of this article, we will compress the secret image and prepare to try to hide two secret images in the encoder, at the same time, the decoder extracts two secret images from stego-image. Further realize fast transfer and high-capacity steganography.

REFERENCES

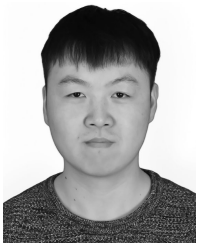
[1] A. Kokaj, "Cyber war and terrorism in kosovo," *Academic J. Bus., Admin., Law Social Sci.*, vol. 5, no. 1, pp. 124–128, Mar. 2019.
[2] C. Kim, D. Shin, L. Leng, and C.-N. Yang, "Lossless data hiding for absolute moment block truncation coding using histogram modification," *J. Real-Time Image Process.*, vol. 14, no. 1, pp. 101–114, Jan. 2018.
[3] C. Kim, D.-K. Shin, C.-N. Yang, and L. Leng, "Hybrid data hiding based on AMBTC using enhanced Hamming code," *Appl. Sci.*, vol. 10, no. 15, p. 5336, Aug. 2020.
[4] R. Vinodhini and P. Malathi, "DNA based image steganography," in *Computational Vision and Bio Inspired Computing*, vol. 28. Cham, Switzerland: Springer, 2018, pp. 819–829.
[5] W.-H. Lin, S.-J. Horng, T.-W. Kao, P. Fan, C.-L. Lee, and Y. Pan, "An efficient watermarking method based on significant difference of wavelet coefficient quantization," *IEEE Trans. Multimedia*, vol. 10, no. 5, pp. 746–757, Aug. 2008.
[6] T. Filler, J. Judas, and J. Fridrich, "Minimizing embedding impact in steganography using trellis-coded quantization," *Proc. SPIE*, vol. 7541, Jan. 2010, Art. no. 754105.
[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.
[8] J. Zhu, R. Kaplan, J. Johnson, and F. Li, "HiDDeN: Hiding data with deep networks," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 682–697.
[9] K. Alex Zhang, A. Cuesta-Infante, L. Xu, and K. Veeramachaneni, "SteganoGAN: High capacity image steganography with GANs," 2019, *arXiv:1901.03892*. [Online]. Available: http://arxiv.org/abs/1901.03892
[10] D. Hu, L. Wang, W. Jiang, S. Zheng, and B. Li, "A novel image steganography method via deep convolutional generative adversarial networks," *IEEE Access*, vol. 6, pp. 38303–38314, 2018.
[11] C. Yu, "Attention based data hiding with generative adversarial networks," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 1, 2020, pp. 1120–1128.
[12] A. ur Rehman, R. Rahim, M. S. Nadeem, and S. ul Hussain, "End-to-end trained CNN encode-decoder networks for image steganography," 2017, *arXiv:1711.07201*. [Online]. Available: http://arxiv.org/abs/1711.07201
[13] R. Zhang, S. Dong, and J. Liu, "Invisible steganography via generative adversarial networks," *Multimedia Tools Appl.*, vol. 78, no. 7, pp. 8559–8575, Apr. 2019.
[14] F. Chen, Q. Xing, and F. Liu, "Technology of hiding and protecting the secret image based on two-channel deep hiding network," *IEEE Access*, vol. 8, pp. 21966–21979, 2020.
[15] A. Cheddad, J. Condell, K. Curran, and P. Mc Kevitt, "Digital image steganography: Survey and analysis of current methods," *Signal Process.*, vol. 90, no. 3, pp. 727–752, Mar. 2010.
[16] F. Yaghmaee and M. Jamzad, "Estimating watermarking capacity in gray scale images based on image complexity," *EURASIP J. Adv. Signal Process.*, vol. 2010, no. 1, Dec. 2010, Art. no. 851920.
[17] V. Kavitha and K. S. Easwarakumar, "Neural based steganography," in *Proc. Pacific Rim Int. Conf. Artif. Intell.*, vol. 3157, 2004, pp. 429–435.
[18] Y. Sun, H. Zhang, T. Zhang, and R. Wang, "Deep neural networks for efficient steganographic payload location," *J. Real-Time Image Process.*, vol. 16, no. 3, pp. 635–647, Jun. 2019.
[19] R. Khare, R. Mishra, and I. Arya, "Video steganography using LSB technique by neural network," in *Proc. Int. Conf. Comput. Intell. Commun. Netw.*, Nov. 2014, pp. 898–902.
[20] S. Baluja, "Hiding images within images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 7, pp. 1685–1697, Jul. 2020.
[21] S. Jegou, M. Drozdzal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: Fully convolutional DenseNets for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jul. 2017, pp. 11–19.
[22] A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *Proc. Int. Conf. Pattern Recognit.*, Aug. 2010, pp. 2366–2369.
[23] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust registration of multimodal remote sensing images based on structural similarity," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2941–2958, May 2017.
[24] W. Tang, S. Tan, B. Li, and J. Huang, "Automatic steganographic distortion learning using a generative adversarial network," *IEEE Signal Process. Lett.*, vol. 24, no. 10, pp. 1547–1551, Oct. 2017.
[25] J. Hayes and G. Danezis, "Generating steganographic images via adversarial training," in *Proc. Neural Inf. Process. Syst.*, 2017, pp. 1951–1960.
[26] H. Shi, J. Dong, W. Wang, Y. Qian, and X. Zhang, "SSGAN: Secure steganography based on generative adversarial networks," in *Proc. Pacific Rim Conf. Multimedia.*, 2017, pp. 534–544.

[27] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2881–2890.
[28] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
[29] T. Tong, G. Li, X. Liu, and Q. Gao, "Image super-resolution using dense skip connections," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4799–4807.
[30] Z. Zhou, H. Sun, R. Harit, X. Chen, and X. Sun, "Coverless image steganography without embedding," in *Proc. Int. Conf. Cloud Comput.*, vol. 9483, 2015, pp. 123–132.
[31] Z. L. Zhou, Y. Cao, and X. M. Sun, "Coverless information hiding based on bag-of-words model of image," *J. Appl. Sci.*, vol. 34, no. 5, pp. 527–536, 2016.
[32] K.-C. Wu and W. Chung-Ming, "Steganography using reversible texture synthesis," *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 130–139, Jan. 2015.
[33] M.-m. Liu, M.-q. Zhang, J. Liu, Y.-n. Zhang, and Y. Ke, "Coverless information hiding based on generative adversarial networks," 2017, *arXiv:1712.06951*. [Online]. Available: http://arxiv.org/abs/1712.06951
[34] J. Xu, X. Mao, X. Jin, A. Jaffer, S. Lu, L. Li, and M. Toyoura, "Hidden message in a deformation-based texture," *Vis. Comput.*, vol. 31, no. 12, pp. 1653–1669, Dec. 2015.
[35] B. Boehm, "StegExpose—A tool for detecting LSB steganography," 2014, *arXiv:1410.6656*. [Online]. Available: http://arxiv.org/abs/1410.6656

XINTAO DUAN received the master's degree in computer application technology from Shanghai Normal University, in 2004, and the Ph.D. degree in communication and information systems from Shanghai University, in 2011. He has been teaching and researching with Henan Normal University, since July 2004. His research interests include image encryption, information hiding, image forensics, and deep learning.



LIU NAO received the B.S. degree from Henan Agricultural University, China, in 2018. He is currently pursuing the M.S. degree with the College of Computer and Information Engineering, Henan Normal University. His research interests include image processing, deep learning, and image steganography.



GOU MENGXIAO received the B.S. degree from Henan Agricultural University, China, in 2018. He is currently pursuing the M.S. degree with the College of Computer and Information Engineering, Henan Normal University. His research interests include image processing, deep learning, and image steganography.

DONGLI YUE received the M.S. degree from the Chinese Academy of Sciences, Hefei, China, in 2003. She is currently an Associate Professor with Henan Normal University, Xinxiang, China. Her current research interests include image processing and machine learning.

ZIMEI XIE received the M.S. degree from Shantou University, Shantou, China, in 2006. She is currently a Lecturer with Henan Normal University, Xinxiang, China. Her current research interests include image processing and machine learning.

YUANYUAN MA received the B.S. and M.S. degrees from Henan Normal University, Xinxiang, China, in 2004 and 2007, respectively, and the Ph.D. degree from the Zhengzhou Information Science and Technology Institute, Zhengzhou, China, in 2019. She is currently a Lecturer with Henan Normal University. Her research interest includes image steganalysis technique.

CHUAN QIN (Member, IEEE) received the B.S. degree in electronic engineering and the M.S. degree in signal and information processing from the Hefei University of Technology, Anhui, China, in 2002 and 2005, respectively, and the Ph.D. degree in signal and information processing from Shanghai University, Shanghai, China, in 2008. Since 2008, he has been with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, where he is currently a Professor. He was with Feng Chia University, Taiwan, as a Postdoctoral Researcher, from 2010 to 2012. His research interests include image processing and multimedia security. He has published more than 110 articles in these research areas.