2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

A High-Imperceptibility Image Steganography Scheme via Makeup Transfer Network and Multiple Feature Fusion

Meihong Yang[1,2][0009-0006-9197-299X], Ziyi Feng[1,2][0009-0009-6783-2278],
Bin Ma[1,2(✉)][0000-0002-9030-7393], Qi Li[1,2][0000-0001-7729-1422], and Linna Zhou[3][0009-0004-0727-2436]

[1] Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), Qilu University of Technology (Shandong Academy of Sciences), Jinan, China
[2] Shandong Provincial Key Laboratory of Industrial Network and Information System Security, Shandong Fundamental Research Center for Computer Science, Jinan, China
[3] Beijing University of Posts and Telecommunications
sddxmb@126.com

Abstract. The existing image steganography will inevitably cause modification traces to the cover image, resulting in the risk of secret information leakage. Therefore, this paper proposes a color image steganography algorithm based on Makeup Transfer Network and multi-scale feature fusion. This paper aims to achieve the embedding of secret image in the process of makeup transfer. Specifically, the secret image is initially mapped into its latent representation, then, it performs multi-scale feature fusion with makeup features to generate a makeup-ed stego image, resulting in the excellent quality of steganographic image and the high imperceptibility of secret information. Moreover, the Information Compensation Network (ICN) was constructed for deep fine-grained feature fusion, by using the differences between the original and rebuilt secret information as network loss, the information of secret image is comprehensively compensated and its quality is further improved. Experimental results show that the proposed scheme exhibits superior image quality on both the target image and the recovered secret image, thus providing good security.

Keywords: Steganography, Makeup transfer, Information compensation, Multi-scale feature fusion.

1 Introduction

Data hiding technology is an algorithm that can embed data into media and extract data effectively, which is mainly used for ownership protection, identity verification, secret communication and other purposes [1] [2] [3]. However, traditional data hiding technologies often face problems such as low embedding capacity and poor robustness. Therefore, image steganography has been proposed, which can not only realize the protection of secret information, but also hide the embedding behavior of secret information [4].

Early image steganography is mostly based on heuristic image steganography methods, such as HUGO [5], WOW [6], S-UNIWARD [7], Hill [8], etc. These schemes involve adaptively modifying the pixel values or frequency coefficients of the cover image according to a predesigned secret data embedding constraint function, and dynamically adjusting the data embedding cost to reduce the detection probability of the steganalyzer. By decomposing steganography into secret data encoding and embedding distortion minimization, these methods achieve a high degree of indistinability in appearance, but face problems such as limited capacity and poor steganalysis resistance. Therefore, steganography based on deep learning has been widely studied in recent years. Baluja et al. [9] constructed multiple deep neural networks for image preprocessing and information hiding, which effectively enhanced the quality of stego images. In addition to the image quality of the stego image, the security of the stego image is also a key metric to consider. Zheng et al. [10] developed a component-aware image steganography to achieve visual security and resistance to deep steganalysis through self-generated supervision. They also improved the naturalness of stego images and ensured high steganographic capability by fusing rule-based synthesis methods and generative adversarial networks. Recently, in order to reduce the image distortion caused90 by the steganography algorithm in the process of large-capacity information embedding and enhance the anti-steganalysis ability of the generated stego image, Ma et al. [4] proposed a large-capacity differential steganography algorithm for color images based on multiple adversarial networks. The information hiding was completed by embedding the secret information into the differential plane generated by the two most similar channels of the cover image. Thus, the distortion of the stego image is minimized.

However, there is still a significant security risk with the above approach: if an attacker can obtain the original cover image, they can distinguish the stego image from the cover image by analyzing the differences in the histogram distribution. Therefore, embedding secret information into any region of the cover image will inevitably change the cover image itself. If the original cover image is corrupted, steganography will be easily detected. Therefore, steganography based on style transfer came into being. They embedded secret information into the cover image, and then stylized it to eliminate the disturbance of the secret information to the cover image [11] [12]. Recently, Shi et al. [13] proposed a steganography strategy based on style transfer. Steganography tools masquerade as deep neural networks (DNNS) that perform style transfer tasks. However, the style transfer steganography produced by their strategy has poor visual effect, and the embedded secret information is a binary watermark, which can not meet the requirements of large-capacity embedding scenes.

In this paper, an image steganography method based on makeup transfer and multi-scale feature fusion is proposed. In the process of beautifying the face without makeup, the generation features of the secret image and the makeup features are multi-scale fused, so as to avoid direct modification of the cover image. Compared with the existing methods of modifying image content to achieve data embedding, the proposed method is more secure. Compared with the most advanced style transfer steganography strategy, our proposed strategy has a strong embedding ability, and the resulting stego images have a better visual effect, realizing reliable visual optimization of the original images. The main contributions proposed in this paper are as follows:

- We propose a steganographic algorithm for same-size color images based on makeup transfer and multi-scale feature fusion. By multi-scale feature fusion of secret image generation features and makeup features, the secret image was embedded during makeup transfer to realize data hiding, and the makeup-ed stego image is generated. The direct modification of the carrier image pixels is avoided, which significantly reduces the image distortion caused by large-capacity data embedding, and improves the visual effect of makeup-ed stego image.
- The Multi-scale Feature Fusion Transfer Module (MFFTM) was designed to embed secret information under the premise of ensuring the makeup effect. The multi-scale feature fusion strategy is used to map the secret image into potential features. During the upsampling process, the potential features and makeup features of different scales of secret images are fully fused to realize the synchronization of makeup transfer and secret image embedding.
- In order to improve the image quality of the reconstructed secret image, Information Compensation Network (ICN) is introduced. Through ICN, part of the secret information lost in the makeup transfer process is embedded back into the makeup-ed stego image, which ensures the integrity of the secret image embedding and further improves the image quality of the reconstructed secret image.

2 Proposed Scheme

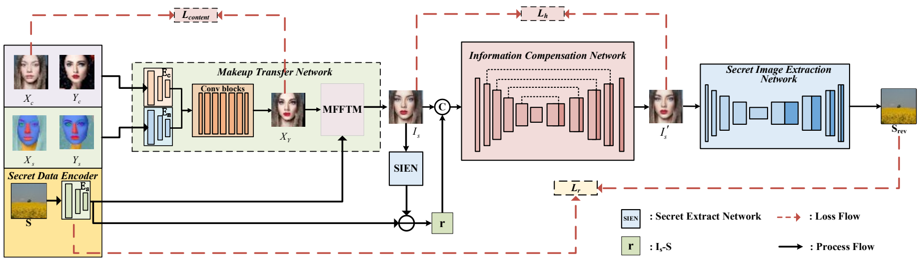

Fig. 1. Overview of the proposed model.

The aim of this study is to design a steganographic network that can realize both image style transfer and data embedding. Our proposed steganography strategy, as shown in Figure 1, systematically consists of three subnetworks: Makeup Transfer Network (MTN), Information Compensation Network (ICN), and Secret Image Extraction Network (SIEN). In MTN, the non-makeup image X_c, makeup image Y_c and their corresponding masks X_s and Y_s are taken as inputs, and the features of the images and their corresponding masks are fused by a series of convolutional blocks, and then the image features are transformed and aligned by calculating the symmetry relationship between the feature maps. Make the facial features of the face without makeup and the face with makeup correspond to each other, and generate the makeup fusion feature figure X_Y. Finally, the MFFTM realizes the dual functions of makeup transformation and secret information embedding at the same time, and obtains the initial makeup-ed stego image I_S. However, considering that the makeup transfer process may interfere with the embedded secret information, and thus lead to information loss, we designed ICN to embed the residual information of the secret image into the initial makeup-ed stego image I_S, perform information compensation, and generate the final makeup-ed stego image $I_S{}'$. SIEN is a deep learning network based on U-Net architecture. By introducing attention mechanism in the process of down-sampling and up-sampling, the expression ability of images is significantly enhanced. Steganographic images are fed into SIEN as inputs to reconstruct secret images from the makeup-ed stego image $I_S{}'$.

2.1 Multi-scale Feature Fusion Transer Module (MFFTM)

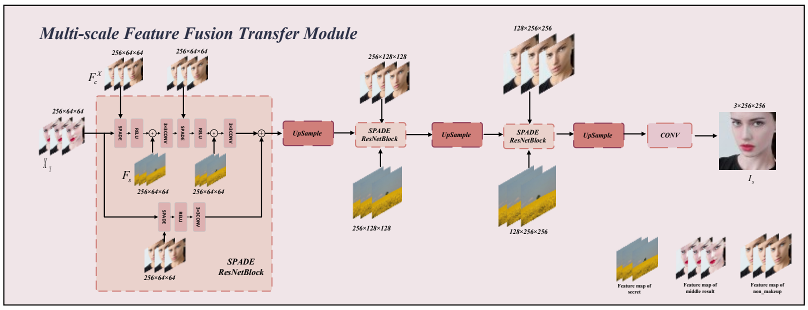

Fig. 2. The workflow of the multi-scale feature fusion transfer module.

The migration architecture plays a crucial role in generating the quality of makeup transfer images embedded with secret information. Specifically, a precise balance needs to be achieved between the effect of makeup transfer and the quality of the resulting image after embedding the secret information. Inspired by [14], this paper designed a multi-scale feature fusion transfer module (MFFTM) to achieve the dual goals of makeup transfer and secret information embedding. The overall framework of the proposed MFFTM is shown in Fig.2.

In MFFTM, by using multiple spatially adaptive denormalization (SPADE) blocks [15], the style features of the reference makeup image and the content features of the non-makeup image are fused at multi-scales to gradually generate natural makeup effects. Each SPADE block not only injects makeup style features, but also incorporates secret image features of different scales to achieve effective embedding of secret information.

For better feature fusion, feature extraction is first done by an independent encoder before the input data enters the MFFTM. Specifically, we design face encoder E_c and secret encoder E_s to extract the content features of non-makeup images and the embedded features of secret images, respectively:

$$F_c^X = E_c(X_c), F_s = E_s(S) \tag{1}$$

In each SPADE block, the normalized non-makeup feature map F_c^X and makeup fusion feature map X_Y are fused, and high-resolution feature maps are generated by layer by layer sampling, and a makeup-ed stego image with a resolution of 256×256 is finally output. The specific structure of the SPADE block is shown in Fig.3.

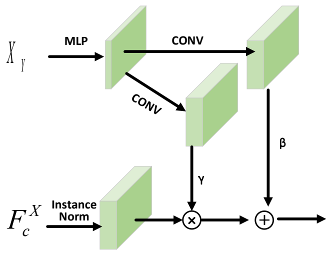

Fig. 3. The SPADE block structure.

Specifically, we use a multi-layer perceptron MLP (including a 3×3 convolution and a RELU activation function) to extract the underlying information of makeup fusion feature map X_Y, which mainly includes features of various areas of the face, and then pass the output of the MLP through two convolution layers to generate a scale factor γ and a bias parameter β, respectively. Finally, the normalized non-makeup feature map F_c^X is fused with γ and β, and the specific formula is as follows:

$$F_{out} = Norm(F_c^X)(1+\gamma)+\beta \tag{2}$$

where, γ represents the zoom ratio required in the makeup transformation, which controls the adjustment of brightness, contrast and other features of the final output image, and β corresponds to the tone change in the makeup adjustment process, which carries out translation operation on the color value of the output image.

At the same time, the secret image feature map F_s of different scales is further fused with the makeuped results obtained by each SPADE module to ensure that the secret information can be embedded in the final makeup-ed stego image. After multi-scale feature fusion processing of multiple SPADE ResNet blocks, the initial makeup-ed stego image I_s with secret information is finally generated. This multi-scale feature fusion strategy can not only generate realistic makeup effects, but also ensure that the embedding process of secret information and the makeup transfer operation are co-optimized, which improves the quality and steganography performance of the makeup-ed stego images.

2.2 Construction of Information Compensation Network

In MFFTM, the core task of makeup transfer is to embed secret information while applying makeup. However, under the premise of ensuring the makeup effect, it will inevitably lead to the loss of deep features of secret information, which will affect the quality of extracted information. In order to solve this problem, we construct the Information Compensation Network (ICN) based on the steganography strategy in [4], which aims to effectively compensate the secret information lost in the makeup transfer process. ICN consists of an encoder part and a decoder part, and the encoder part includes four downsampling modules, each of which consists of a 4 × 4 convolution, a Bach Normaliozation normalization layer, and a LeakyReLU operation. The decoder part consists of four upsampling modules, each of which includes a 4×4 transposed convolution, a Bach Normaliozation normalization layer, and a ReLU activation function. The encoder component is used to extract basic features from the input data, while the decoder component is tasked with reconstructing these features into the spatial structure of the original input. The workflow of ICN can be stated as follows:

Firstly, in order to obtain the lost secret information, the secret image is extracted from I_s through the secret image extraction network (SIEN, See the next section for details), and the residual calculation is performed with the original secret image S, which is defined as follows:

$$r = S - SIEN(I_s) \tag{3}$$

where, r represents the lost secret information.

Then, the secret information residual r and I_s are concatenated in the channel dimension and sent to ICN for feature compensation, and the makeup-ed stego image containing the compensated secret information $I_s{}'$ is generated, which is defined as follows:

$$I_s{}' = ICN(r, I_s) \tag{4}$$

2.3 Secret Image Extraction Network (SIEN)

The recovery ability of secret information is an important index to evaluate the performance of steganography schemes. In order to improve the accuracy of secret information extraction, the secret image extraction network (SIEN) needs to ensure that the extracted secret image has a higher similarity with the original secret image. Therefore, this paper adopts a similar architecture as ICN as the secret image extraction network. It is worth noting that in the up-down sampling process, in order to better retain the details of the image, a feature fusion module is added. In addition, in the jump connection of the upsampling stage, the channel attention mechanism [17] is introduced to extract the key information more effectively, so as to improve the quality and recovery ability of the extracted image. The process of reconstructing the secret information is as follows:

$$S_{rev} = SIEN(I_s{}') \tag{5}$$

where, S_{rev} represents the reconstructed secret image.

2.4 Loss Function

The loss function is an important part of the image steganography scheme based on makeup transfer and multi-scale feature fusion, which is mainly divided into the total loss of makeup transfer module, and the total loss of information compensation network and secret image extraction network.

2.4.1. Loss of makeup transer network.

In order to optimize the quality of the generated makeup-ed stego image, this paper defines the total loss of the makeup transfer module, which can be expressed as L_m and is mainly composed of the weighted sum of three parts, as shown in (6).

$$L_m = \lambda_1 L_{style} + \lambda_2 L_{content} + \lambda_3 L_{Hide} \tag{6}$$

The makeup loss L_m, is used to ensure that the generated initial makeup-ed stego image I_s is consistent with the reference makeup image Y_c in color and style. To this end, we use the paired dataset provided by [14], in which each pair of face X_c with non-makeup and face Y_c with makeup will have a corresponding distorted face X_y, as a criterion to improve the image quality of the generated initial makeup-ed setgo image I_s. To constrain the color similarity between X_Y and X_y we introduce Spatial Profile Loss (SPL) [16] to measure the color loss during makeup transfer:

$$L_{style} = SPL(X_Y, X_y) \tag{7}$$

SPL optimizes the color fidelity and detail consistency of the generated image in different Spaces and scales by calculating the difference between X_Y and X_y in RGB color space, YUV color space and the gradient in YUV space.

The content loss $L_{content}$ is used to ensure that I_s is consistent with the non-makeup image X_c in content information, SPL loss is also used to calculate the content loss between the intermediate result semantic feature maps X_Y and X_c:

$$L_{content} = SPL(X_Y, X_c) \tag{8}$$

Steganographic loss L_{Hide} aims to ensure that the makeup effect of the generated image is not destroyed in the process of embedding the secret image. In order to achieve this goal, a separate makeup module is added, which generates a pure makeup-ed image I without embedding the secret image through MTN, and uses I to constrain the embedding effect of I_s. The steganographic loss is mainly composed of MSE loss and VGG loss.

$$L_{Hide} = L_{MSE} + L_{VGG} \tag{9}$$

$$L_{MSE} = MSE(I, I_s) = \frac{1}{N}\sum_{i=1}^{N}(I(i) - I_s(i))^2 \tag{10}$$

$$L_{VGG} = VGG(I, I_s) = \frac{1}{N_l}\sum_{l=1}^{L}\|\phi_l(I) - \phi_l(I_s)\|_F^2 \tag{11}$$

where, N represents the total number of pixels in the image and $\phi_l(\cdot)$represents the features extracted by the lTH layer of the VGG16 network. N_l represents the product of the total number of channels, width and height of the feature at layer l, and $\|\cdot\|_F^2$ represents the Frobenius norm, which is the sum of element squares.

2.4.2. Loss of Information Compensation Network and Secret Image Extraction Network.

In order to ensure the quality of the stego image and the extracted image, we jointly train the information compensation network and the secret extraction network, and the total loss of these two modules is defined as L_{hr}, as shown in (12).

$$L_{hr} = \lambda_4 L_h + \lambda_5 L_r \tag{12}$$

Embedding loss L_h is used to ensure that the stego image will not be visually detected as containing secret information, requiring that the stego image and the cover image cannot be distinguished, defined as follows:

$$L_h = MSE(I_s, I_s') + VGG(I_s, I_s') + L_{TV} + L_{ST} \tag{13}$$

In addition to the MSE loss and VGG loss, we also introduce the total variation loss (TV loss) to minimize the absolute difference between adjacent pixels in the horizontal and vertical directions of the image, ensuring that the generated image is more natural and avoiding excessive artifacts or noise due to secret information embedding. The total variation loss L_{TV} is defined as follows:

$$L_{TV} = TV(I_s') = \sum_{i=1}^{C}\left(\sum_{h=1}^{H-1}\sum_{w=1}^{W-1}\left|(I_s')_{i,h,w+1} - (I_s')_{i,h,w}\right| + \sum_{h=1}^{H-1}\sum_{w=1}^{W-1}\left|(I_s')_{i,h+1,w} - (I_s')_{i,h,w}\right|\right) \tag{14}$$

where, I_s' size is C×H×W, C is the number of channels, H is the height, and W is the width.

In addition, in order to ensure that the makeup effect is not affected and that I_s' and I_s are as consistent as possible in style, style loss is also introduced, which measures the difference between I_s' and I_s Gram matrix by MSE. Style loss L_{ST} is defined as follows:

$$L_{ST} = \sum_l \left\| G_{I_s'}^{(l)} - G_{I_s}^{(l)} \right\|_2^2 \tag{15}$$

where, $G_x^{(l)}$ represents the Gram matrix of image x at the l-th layer.

The secret image recovered from the stego image must have high visual quality and be as identical as possible to the original secret image, and the extraction loss L_r is described as follows:

$$L_r = MSE(S, S_{rev}) + VGG(S, S_{rev}) \tag{16}$$

3 Experiments

3.1 Experimental setting

For the dataset, we choose SSAT [14] as the dataset of makeup transfer. SSAT includes 300 non-makeup images and 300 makeup images, which means there are a total of 90,000 sets of makeup transfer data, and the widely used dataset ImageNet is used as the secret image dataset for network training. For testing, we systematically constructed 3,000 sets of makeup transfer data by exhaustively pairing each of the 300 non-makeup images with 10 randomly selected makeup images from the SSAT dataset. Similarly, 3,000 color images were randomly selected from ImageNet as secret images for performance evaluation. During the training process, all trainable parameters are initialized normally and trained using the Adam optimizer with $\beta_1 = 0.5$ and $\beta_2 = 0.999$. For the total loss L_m of the makeup transfer module, we set $\lambda_1 = 1$, $\lambda_2 = 1$, $\lambda_3 = 1$ to balance the different loss functions; For the total loss L_{hr} of the information compensation network and the secret extraction network, we set $\lambda_4 = 1$ and $\lambda_5 = 0.75$. The whole network framework is implemented by pytorch and is iterated 90,000 times on a single RTX 3090 GPU card with 24GB VRAM, with the learning rate fixed at 0.0002 and batchsize set to 1.

3.2 Performance Analysis of the proposed scheme

The purpose of this paper is to realize the embedding operation of secret information while makeup transfer, while ensuring the image quality of reconstructed secret image. In order to verify the effectiveness of the designed information compensation network, we remove the information compensation network to analyze the impact of the compensation operation. In this case, we directly skip the ICN and input the initial makeup-ed stego image I_s into SIEN to extract the secret information. The comparison results are shown in Fig.4. (a) "Non-makeup" represents the image without makeup. (b) "Makeup" represents the reference makeup image. (c) "Secret images" represents the original secret image. (d) "I_s" represents the initial makeup-ed stego image generated without information compensation. (e) is the residual between the extracted secret image and the original secret image without ICN, and it can be seen that the residual is large and the color is bright. (f) "$I_s{}'$" represents the final makeup-ed stego image generated after information compensation. (g) represents the residual between the extracted secret image and the original secret image after ICN is added to compensate for the lost information. It can be seen that the residual is smaller and the color is relatively darker.

Fig. 4. Regarding the results of the ICN ablation experiments. (a) Non-makeup images. (b) Makeup images. (c) Secret images. (d)I_s. (e) The residual between $S_{rev}{}'$ and Secret images magnified by 3. (f)$I_s{}'$. (g) The residual between S_{rev} and Secret images magnified by 3.

In addition, to further analyze the impact of the information compensation network, we show the results of peak signal-to-noise ratio and structural similarity between the recovered secret image and the original secret image by our method with and without ICN, as shown in Table I.

Table 1. Ablation study of ICN

ICN	w/o	w
PSNR↑	21.02	35.73
SSIM↑	0.904	0.988

3.3 Compared With Other Advanced Schemes

In order to fully verify the effectiveness of the image steganography scheme based on makeup transfer and mutli-scale feature fusion proposed in this paper, we conduct experimental comparisons with several state-of-the-art methods in the current field on the same dataset. Considering that the proposed method involves two core tasks of makeup transfer and secret information steganography at the same time, the comparative experiments are carried out from the following two aspects.

Makeup performance: In order to verify the superiority of our makeup transfer strategy, we select three state-of-the-art makeup transfer methods, BeautyGAN [18], PSGAN [19] and EleGANt [20], as the comparison benchmarks, which have high visual effects and style expression ability in the field of makeup transfer. Fig.5 shows the comparison between our method and the above three methods in the makeup transfer effect. It is obvious that the makeup-ed images generated by our method are more visually pleasing, and the makeup style is highly similar to the reference makeup image. In the presence of spatial misalignment issues, our method shows the best transfer results and is able to achieve complete makeup transfer, while other methods suffer from local artifacts, uneven facial tones, and inadequate transfer of eyeshadow and lipstick. When there is no spatial misalignment problem, BeautyGAN performs well, but cannot effectively transfer eyeshadow; PSGAN can generate visually satisfactory results, but the effect of eye shadow transfer is still not ideal, and the local color uneven phenomenon of the lip is easy to occur when the lip is occlusive (as shown in the sixth row). The overall effect of EleGANt is more ideal, but there is still a problem of color loss in eye shadow transfer, which is obviously different from the reference makeup image. At the same time, in the case of lip occlusion, Elegant is also similar to PSGAN, which is characterized by uneven local tone of the lips. It is worth noting that the makeup-ed image generated by our method can still ensure excellent makeup effect when the secret image is embedded. Compared with other methods, our method is not only not inferior in the visual effect and style expression of makeup transfer, but also superior in color richness and spatial dislocation transfer.

Steganography performance: In this experiment, the same data set was used for testing, 300 makeup-ed images I_s generated by MTN of this scheme were used as the cover image, and 300 color images were randomly selected from ImageNet as secret images. Moreover, the proposed scheme is compared with three advanced steganography methods, namely, Baluja's scheme [9], Ma's scheme [4] and Zheng's scheme [10]. For objective evaluation, we evaluate the image steganography and extraction capabilities by peak signal-to-Noise ratio (PSNR) and structural similarity (SSIM). The experimental results are shown in Table II. It can be seen that the PSNR and SSIM of the beauty stego images generated by the proposed scheme are better than other advanced schemes. At the same time, the fidelity of the extracted secret image is also evaluated in the experiment. In Table II, it can be seen that the visual quality of the secret image extracted by the proposed scheme is also higher than that of other schemes. This is because the steganography strategy in this paper realizes data hiding by multi-scale fusion of the generation features and makeup features of the secret image and combining with makeup transfer, avoiding the direct modification of the cover image pixels, thus significantly reducing the distortion of the stego image caused by data embedding.

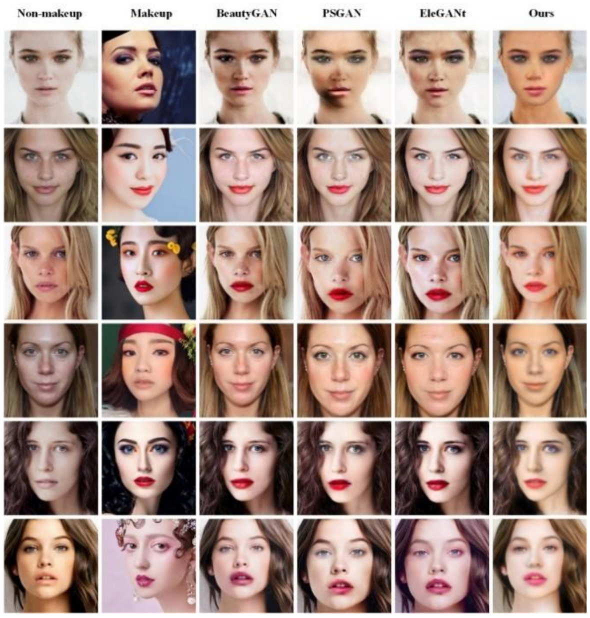

Fig. 5. Comparison with state-of-the-art methods. There is spatial misalignment in the first row and no spatial misalignment in the second to sixth rows.

Table 2. PSNR and SSIM of stego image and extracted image

	Stego (PSNR)	Stego (SSIM)	Extracted (PSNR)	Extracted (SSIM)
Baluja [9]	29.10dB	0.948	28.59dB	0.910
Zheng [10]	32.37dB	0.953	31.89dB	0.942
Ma [4]	37.97dB	0.974	33.85dB	0.944
ours	40.92dB	0.998	35.73dB	0.988

In order to fully evaluate the security of the proposed scheme, XuNet [21], SRNet[22] and ZhuNet[23] were used to verify the anti-steganalysis ability of the scheme. These three networks represent the mainstream methods in the current steganalysis field and can effectively detect steganographic information in images. By performing steganalysis on the stego images generated by different steganographic schemes, the average steganalysis detection rate of each steganographic scheme is calculated and shown in Table III. Experimental results show that our scheme exhibits significantly better steganalysis resistance than other existing steganography schemes. Specifically, the steganalysis detection rates of the stego images generated by the proposed scheme are 50.5%, 56.8%, and 57.2% with respect to the XuNet, SRNet, and ZhuNet steganalyzers, respectively, which maintain a low detection rate compared to other steganalysis schemes. This shows that the proposed steganography scheme has a clear advantage in resisting steganalysis and can effectively reduce the risk of detection while maintaining high quality image embedding.

Table 3. Detection accuracy of XuNet, SRNet, and ZhuNet steganalyzers on three steganographic schemes

	Li [11]	Li [12]	Ma [4]	Ours
XuNet	58.7%	56.9%	53.8%	50.5%
SRNet	67.9%	62.5%	57.3%	56.8%
ZhuNet	71.4%	71.6%	60.3%	57.2%

4 Conclusion

In this paper, an image steganography scheme based on makeup transfer and mutli-scale feature fusion is proposed to achieve large-capacity secret message embedding. By fusing the makeup features of different scales and secret image features, this method not only realizes the embedding of secret information, but also realizes the high-quality transfer of makeup. The proposed method not only has similar visual effects to the traditional makeup transfer method, but also avoids directly modifying the cover image pixels, which significantly reduces the distortion caused by data embedding and improves the security of the stego image. Compared with the existing steganography algorithms that embed data by modifying the cover image, the proposed method has larger embedding capacity without destroying the naturalness of the image, and can support the secret information embedding of the same size color image. In addition, in order to solve the problem that secret information may be lost in the process of makeup transfer, an information compensation network is designed to further improve the quality and fidelity of the reconstructed secret image by compensating the lost secret information in the makeup transfer. Experimental results show that the proposed scheme is superior to several existing advanced methods in terms of makeup transfer effect and steganography performance. While maintaining the visual quality of makeup images, the method achieves large-capacity information embedding, and largely compensates for the damage caused by makeup transfer to secret information. The information compensation network effectively enhances the quality of the extracted secret image, which proves the key role of the module in improving the steganography performance. In the future, we plan to further optimize the makeup transfer embedding network to improve the visual quality of beauty steganography images, and explore more robust embedding mechanisms to improve the embedding ability and anti-attack performance of the steganography system.

Acknowledgments. This work was supported by National Natural Science Foundation of China (62272255, 62302248); Shandong Taishan Scholars Program (tstp20240829); Key Project of Nature Foundation of Shandong Province (ZR2022LZH011); Key RD Projects of Shandong Province (2023CXGC010102, 2023TZXD049); Natural Science Foundation of Shandong Province (ZR2023QF018); Jinan ”New 20 universities” innovation team introduction project (202228016); Major Innovation Project of Science, Education and Industry Integration Pilot Project (2024ZDZX08).

References

1. B. Ma, Z. Tao, R. Ma, C. Wang, J. Li and X. Li, "A High-Performance Robust Reversible Data Hiding Algorithm Based on Polar Harmonic Fourier Moments," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 34, no. 4, pp. 2763-2774, April 2024, doi: 10.1109/TCSVT.2023.3311483.
2. B. Ma, H. Duan, R. Ma, Y. Xian and X. Li, "High-Performance Optimization Framework For Reversible Data Hiding Predictor," in IEEE Signal Processing Letters, doi: 10.1109/LSP.2024.3512357.
3. Bin Ma, Kaixin Du, Jian Xu, Chunpeng Wang, Jian Li, Linna Zhou,A screen-shooting resilient data-hiding algorithm based on two-level singular value decomposition,Journal of Information Security and Applications,Volume 78,2023,103605,ISSN 2214-2126.
4. B. Ma, H. Wang, J. Xu, X. Wang, X. Li and J. Li, "Color Image High-Capacity Differential Steganography Algorithm Based on Multiple Adversarial Networks," in IEEE Transactions on Circuits and Systems for Video Technology, doi: 10.1109/TCSVT.2024.3508849.
5. T. Pevn ́y, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," in Proc. 12th Int. Conf. Inf. Hiding, Jun. 2010, pp. 161–177.
6. V. Holub and J. Fridrich, "Designing steganographic distortion using directional filters," 2012 IEEE International Workshop on Information Forensics and Security (WIFS), Costa Adeje, Spain, 2012, pp. 234-239
7. V. Holub, J. Fridrich, and T. Denemark, "Universal distortion function for steganography in an arbitrary domain," EURASIP J. Inf. Secur., vol. 2014, no. 1, p. 1, Dec. 2014.
8. B. Li, M. Wang, J. Huang, and X. Li, "A new cost function for spatial image steganography," in Proc. IEEE ICIP, Paris, France, Oct. 2014,pp. 4206–4210.
9. Baluja S. "Hiding images in plain sight: Deep steganography". Advances in neural information processing systems, 2017, 30.
10. Z. Zheng, Y. Hu, Y. Bin, X. Xu, Y. Yang and H. T. Shen, "CompositionAware Image Steganography Through Adversarial Self-Generated Supervision," in IEEE Trans. Neural Netw. Learn. Syst., vol. 34, no. 11, pp. 9451-9465, Nov. 2023, doi: 10.1109/TNNLS.2022.3175627.
11. Q. Li, B. Ma, X. Fu, X. Wang, C. Wang and X. Li, "Robust Image Steganography via Color Conversion," in IEEE Transactions on Circuits and Systems for Video Technology, doi: 10.1109/TCSVT.2024.3466961.
12. Q. Li, B. Ma, X. Wang, C. Wang and S. Gao, "Image Steganography in Color Conversion," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 71, no. 1, pp. 106-110, Jan. 2024, doi: 10.1109/TCSII.2023.3300330.
13. R. Shi, Z. Wang, Y. Hao and X. Zhang, "Steganography in Style Transfer," in IEEE Transactions on Artificial Intelligence, vol. 5, no. 12, pp. 6054-6065, Dec. 2024, doi: 10.1109/TAI.2024.3379946.
14. Sun, Z., Chen, Y., & Xiong, S. (2021). SSAT: A Symmetric Semantic-Aware Transformer Network for Makeup Transfer and Removal. *ArXiv, abs/2112.03631*..
15. Park, T.; Liu, M.-Y.; Wang, T.-C.; and Zhu, J.-Y. 2019. Semantic Image Synthesis with Spatially-Adaptive Normalization. In Computer Vision and Pattern Recognition, 23372346.
16. Sarfraz, M. S.; Seibold, C.; Khalid, H.; and Stiefelhagen, R. 2019. Content and Colour Distillation for Learning Image Translations with the Spatial Profile Loss. In British Machine Vision Conference, 287.
17. Woo, S., Park, J., Lee, JY., Kweon, I.S. (2018). CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds) Computer Vision –

ECCV 2018. ECCV 2018. Lecture Notes in Computer Science(), vol 11211. Springer, Cham. https://doi.org/10.1007/978-3-030-01234-2_1.

18. Li, T.; Qian, R.; Dong, C.; Liu, S.; Yan, Q.; Zhu, W.; and Lin, L. 2018. BeautyGAN: Instance-level Facial Makeup Transfer with Deep Generative Adversarial Network. In ACM Multimedia, 645–653.
19. Jiang, W.; Liu, S.; Gao, C.; Cao, J.; He, R.; Feng, J.; and Yan, S. 2020. PSGAN: Pose and Expression Robust SpatialAware GAN for Customizable Makeup Transfer. In Computer Vision and Pattern Recognition, 5194–5202.
20. Yang, C., He, W., Xu, Y., Gao, Y. (2022). EleGANt: Exquisite and Locally Editable GAN for Makeup Transfer. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds) Computer Vision – ECCV 2022. ECCV 2022. Lecture Notes in Computer Science, vol 13676. Springer, Cham. https://doi.org/10.1007/978-3-031-19787-1_42.
21. G. Xu, H.-Z. Wu, and Y.-Q. Shi, "Structural Design of Convolutional Neural Networks for Steganalysis," IEEE Signal Process. Lett., vol. 23, no. 5, pp. 708–712, May 2016
22. M. Boroumand, M. Chen, and J. Fridrich, "Deep Residual Network for Steganalysis of Digital Images," IEEE Trans. Inf. Forensics Secur., vol. 14, no. 5, pp. 1181–1193, May 2019
23. R. Zhang, F. Zhu, J. Liu, and G. Liu, "Depth-Wise Separable Convolutions and Multi-Level Pooling for an Efficient Spatial CNN-Based Steganalysis," IEEE Trans. Inf. Forensics Secur., vol. 15, pp. 1138–1150, 2020