RELABEL, SCRAMBLE, SYNTHESIZE:A NOVEL COVERLESS STEGANOGRAPHY APPROACH VIA COLLAGE IMAGE

Koi Yee Ng*, Simying Ong*, Yuen Peng Loh†, and Chee Seng Chan*
* Faculty of Computer Science and IT, Universiti Malaya, Malaysia
E-mail: {wva190003@siswa., simying.ong@, cs.chan@}um.edu.my
† Faculty of Computing and Informatics, Multimedia University, Malaysia
E-mail: yploh@mmu.edu.my

***Abstract*—In this age of digital information, most correspondences are now transmitted through online channels, including private, confidential, and secret messages. Hence, a significant concern in such correspondence is the threat of eavesdropping. This work proposes a novel framework of coverless steganography by secret representations relabelling and positions scrambling to increase the complexity of a secret message encoding. This is followed by an ACGAN-based model implementation, namely StArtGAN, that is used in tandem with duotone effects to construct collage images from the encoded message, so as to produce a medium of inconspicuous communication. The decoding process then leverages on the discriminator of the StArtGAN model as the unique decoder to ensure the integrity and accurate recovery of the secret message. Investigations in this paper show that the proposed framework is scalable and holds potential to reduce the risk of eavesdropping in practical applications.**

I. Introduction

Steganography conceals external data into a cover such as image, video, text, audio, etc., for covert communications. The ability to resist the detection of communication and/or concealed data from eavesdroppers has been a great concern in steganography. Besides, there are other important considerations in steganography, including (a) maximizing the size of hidden data, and (b) resistance to the brute force attack, while (c) maintaining inconspicuous communication.

There are two major categories of steganography methods based on the cover utilization, namely cover-based methods and coverless-based methods. Cover-based image steganography method is often heavily dependent on the modification of the cover image for secret embedding. For instance, Least Significant Bits (LSB) insertion modifies the right-most bits of the image pixels to carry binary secret bit [1], [2], [3]. Histogram Shifting (HS) utilizes the peak bin (i.e., histogram bin with the highest frequency) and the adjacent emptied bin as the paired-channels to represent secrets [4], [5]. Difference Expansion (DE) exploits the difference between two adjacent pixels and appends them according to the secret data [6], [7]. All these methods modify the image pixels directly and purposely to carry secret data, which can inevitably leave modification traces which are prone to detection by steganalysis tools.

Recent years have seen the proposal of coverless image steganography, including the non-constructive- [8], [9] and constructive-based methods [10], [11], [12]. The key ideas in the aforementioned coverless methods are, the use of the cover image's contents for secret representation, or by synthesizing encoded images from the secret, respectively, both of which do not modify the cover image. However, the former method has low embedding capacity due to the limitation of image contents, i.e., image pixel values, image size, etc. Whilst for the latter method, the generated embedded images are not always natural enough to avoid suspicions and attacks.

More recently, research works [13], [14] had attempted the use deep learning methods, especially Generative Adversarial Networks (GANs) [15] for coverless steganography by performing encoded images synthesis. Inspired by [13], we theorize that a GAN can synthesize ambiguous and inconspicuous stego-images, while ensuring a promising secret recovery rate with resistance to detection from eavesdroppers.

In this paper, we propose a secret encoding framework, which comprises class-codeword relabelling, position scrambling, secret image synthesizing via ACGAN (Auxiliary Classifier Generative Adversarial Network) [16], and information hiding using synthesized images in duotone collage image. Relabelling and position scrambling increase the complexity of the secret encoding, whereas image synthesis induces an element of ambiguity into the encoding that would be difficult for unintentional recipients to decipher. Thus, the chances of successful unintended analysis of the secret can be greatly reduced. On top of that, the encoded images are presented in a duotone collage image, therefore reduce the suspicion of the generated image, meanwhile ensuring close to 100% secret recovery.

The main contributions of this work are (1) image synthesis to represent secret message and specialized classification using an ACGAN framework as a unique secret encoder-decoder model named StArtGAN, (2) information hiding via collage of synthesised images as a medium of transmission, and (3) the secret message relabelling and secret image position scrambling to improve the secrecy and strengthen the resistance to brute force attack.

II. Proposed Method

Our model consists of two main operations: secret encoding and secret decoding processes as illustrated in Fig. 1. In

TABLE I: Secret message relabelling using K and its corresponding class.

Secret Message Relabelling								
Orginal	000	001	010	011	100	101	110	111
Relabel	101	111	110	010	100	011	000	001
Class	5	7	6	2	4	3	0	1

this section, we present our detailed proposed framework and describe the architecture implemented.

A. Overall Framework

In the encoding phase, the secret message S_μ is transformed into its binary form S_ϵ based on ASCII Encoding. This binary secret message is then further transformed by utilizing a user defined digit representation table R (e.g., Table I) generated using a key K. We term this operation as relabeling to produce S_η which will act as part of an input to our image synthesis model, the Generator G of the ACGAN together with a random vector z to induce variations into the synthesized images. After that, the generated secret images I_η will be concatenated into I_ω and converted into a duotone effect collage image I_Σ. This collage image shall be the medium delivered to the intended recipient.

Following the reception of the collage image is the decoding process back into the original message. In this process, I_Σ will be recovered to obtain I'_ω and then I'_η, which are fed into a decoder, specifically the Discriminator D' of the ACGAN that transforms I'_η back into a code S'_η, where $S'_\eta \sim S_\eta$. This code will then undergo relabelling using K to transform it back into the original coding format, S'_ϵ where $S'_\epsilon \sim S_\epsilon$. Finally, the secret message S'_μ can be decoded back to its original form.

The relabelling step increases the level of complexity that is harder to be breached by eavesdroppers of the secret transmission. Moreover, incorporating synthesized image into the pipeline introduces three additional advantages, they are: (1) decoding that is only possible with a specifically trained model, (2) well synthesized encoding images, which in this context, images that are sufficiently ambiguous and not easily distinguish in any shape or form, and (3) incorporating photos effect elements that imparts a natural appearance that do not arouse the suspicion of eavesdroppers. We further detail the specifications of each stage of the encoding and decoding in the subsequent subsections.

B. The Encoding Phase

The secret encoding phase consists of four phases: (a) secret-to-binary transformation; (b) secret message relabelling; (c) encoded images synthesis; and (d) duotone collage image information hiding.

1) secret-to-binary transformation: Secret message S_μ is the data that the sender wants to send to the recipient covertly via a public communication channel, and it can be in any form including text, image, etc. In our proposed method, S_μ will be first transformed using the following equation:

$$f_1^e(S_\mu) \rightarrow S_\epsilon \tag{1}$$

where f_1^e is the secret-to-binary translation function and S_ϵ is the secret message in binary form (viz., 0 and 1 only).

2) Secret Message Relabelling: The transformed S_ϵ is then divided into $\log_2 \Lambda$ bits per group, where Λ is the number of class in our model. To ease the illustration, $\Lambda = 8$ will be utilized in the following discussion, hence $\|S_\eta\| = log_2 8 = 3$.

In this phase, random sequences will be generated using secret key K to relabel the secret message. The first row of the table indicates the original secret message, while the second row shows the relabelled message (translated from the random sequences) in representing the original secret message. Meanwhile, the third row of the table shows the corresponding class of each group, which is named as S_η.

$$f_2^e(S_\epsilon) \xrightarrow{\text{Table R}} S_\eta \tag{2}$$

3) Encoded Images Synthesis: For this use case, we designed an image synthesizer (encoding model) and classifier (decoding model) based on the ACGAN [16] architecture as the base design enables the inclusion of our secret code into the image generation process as well as produce a unique decoder.

ACGAN: Unlike the general GAN architecture where only a single noise vector z is used to generate synthetic images, every sample generated by ACGAN has a corresponding class label, c. Hence, the synthetic image generator model G will generate the fake sample image $X_{\text{fake}} = G(z, c)$. In the training stage, the real sample X_{real} and $G(z, c)$ are fed into discriminator model D' where it outputs the probability distribution indicating if the image is real or fake, $P(\zeta = \{\text{real}, \text{fake}\}|X)$, as well as the probability distribution over the classes, $P(\rho|X)$, hence, $D'(X) = P(\zeta|X), P(\rho|X)$. Considering the two tasks, the objective functions are therefore, the log-likelihood of the image sources (i.e. real or fake), L_ζ, and the log-likelihood of the image class, L_ρ:

$$\begin{aligned} L_\sigma = E[\log P(\sigma = \text{real } |X_{\text{real}})] + \\ E\left[\log P\left(\sigma = \text{fake}|X_{\text{fake}}\right)\right] \end{aligned} \tag{3}$$

$$\begin{aligned} L_\rho = E\left[\log P\left(\rho = c|X_{\text{real}}\right)\right] + \\ E\left[\log P\left(\rho = c|X_{\text{fake}}\right)\right] \end{aligned} \tag{4}$$

During the minimax training of this model, D' and G are trained to maximize $L_\sigma + L_\rho$ and $L_\rho - L_\zeta$ respectively.

StArtGAN: Based on the ACGAN framework, we design the Steganography Art GAN (StArtGAN) architecture with the generator as an encoder of secret into images, whereas the discriminator is the decoder of images back into the secret. Unlike typical image synthesise GAN models that aim for realism, the priorities of the StArtGAN are to maximize the ambiguity of the images generated while still maintaining a high classification accuracy.

Our Generator, G takes in the class label c and 100-dimension noise vector z for generation and each forms a 128-channel 50×50 embedding which are then concatenated into a 256-channel embedding, as equal balance of the class information and noise variation to represent the secret. This embedding then goes through 3 layers of deconvolution that

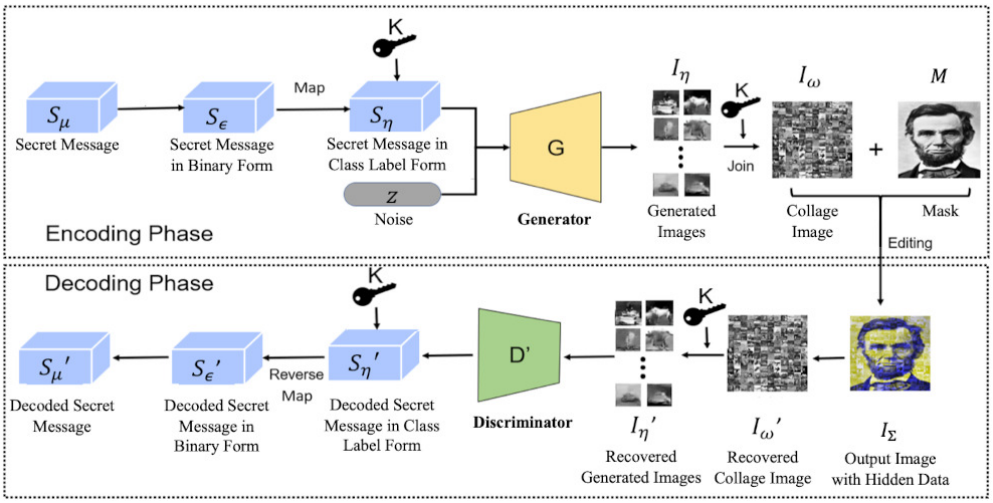

Fig. 1: The proposed framework. The upper box illustrates the flow of the encoding process while the lower box shows the decoding process.

reduces the channels while expanding the spatial dimensions to form a 32×32 single-channeled grayscale image. On the other hand, the Discriminator D' consists of two convolution layers that branches out into 2 densely connected layers that performs the classification and discrimination tasks when given the generated images.

In our context, the input c will be the encoded secret S_η to synthesize the class image using G. To demonstrate, every digit in S_η are used to synthesize a grayscale 32×32 pixel images, I_η thus transforming the secret into a series of images.

4) Duotone Effect Collage Image Information Hiding: Duotone effect is among one of the most widely used photo editing effects, which implements two contrasting colors to create a visually pleasing effect in images. The secret images generated, I_η, are in grayscale, which consists of only 1-channel.

A total of $N \times N$ I_η secret images are collected and their sequences are re-ordered by using K. After that, these secret images are joined together (viz., in raster manner) to form as a collage image I_ω.

At this step, our proposed framework further suggests to transform the collage image into duotone image and/or overlay it with another mask M to create an inconspicuous output image. The purpose of this suggestion is to avoid the attention by the eavesdropper, especially when the output image is transmitted via unsecured public communication channel.

The collage image is converted into 3-channel (i.e., RGB channels) by replicating the channels with same pixel intensities. For instance, an image pixel of value 79 in grayscale collage image is converted into $\{79, 79, 79\}$ to represent RGB channels. Duotone collage image can be easily generated by zeroing-out the other two channels of sender's choice. If sender selected red color as the main color, the RGB channels will become $\{79, 0, 0\}$, and vice versa. Mask M can be overlayed on top of the 3-channel collage image as an additional alpha channel. After the duotone collage image overlay process, output image with hidden data, I_Σ, will be sent to the recipient.

C. The Decoding Phase

On the recipient end, the pixel values in I_Σ are extracted to recover the collage image I'_ω, where $I'_\omega \sim I_\omega$, assuming that there is no information loss during the transmission. Since it is known that the secret images are of size 32×32, the secret images I'_η can be recovered in correct sequences using K (i.e., the secret key K should be shared by both sender and recipient), and similarly, $I'_\eta \sim I_\eta$.

In this phase, the receiver will already have the trained decoder, D' to decode the I'_η. D' can be the same discriminator D' from the StArtGAN, i.e. trained with G, or an externally trained model $\hat{D}$ using the same settings as D' and data samples generated by G. Hence, for every 32×32 pixels of the image I'_η, D' will perform classification to obtain S'_ϵ.

After that, Table R can then be regenerated by using K as reference to retrieve the decoded message in binary form S'_ϵ. The reverse process of Equation 1 are performed to generate the final decoded secret message S'_μ.

Notably in this framework, the decoder is re-useable as long as the same settings are used to generate the data, therefore, it will not be necessary to retrain or re-transmit a decoder from the sender to the receiver. Similar to the secret key, the receiver will already have the decoder and only the collage image containing the secret will be sent thus minimizing the transmission load.

III. Experiments and Results

In our preliminary experiments, the StArtGAN is trained and tested on a subset of the WikiArt [17] dataset containing images of art created by notable artists such as Picasso, Braque, etc. This dataset is selected for its challenging visual appearance even to a regular person's observation, thus allowing the model to generate visually ambiguous image. It should be noted that the choice of dataset will not have

TABLE II: Average classification accuracy on 1000 generated images from each class.

Training Data	Model	Accuracy
WikiArt	ResNet50	13.60%
	MobileNetv2	16.63%
	EfficientNet	20.06%
	$\hat{D}$	18.44%
	StArtGAN	**99.91%**
Synthesized Images	ResNet50	99.81%
	MobileNetv2	99.90%
	EfficientNet	99.94%
	$\hat{D}$	91.24%

TABLE III: Average clustering homogeneity of I'_η from 100 sample output images I_Σ.

Features	Dimensions	Homogeneity Score
Pixels	1024	0.6081
PCA	256	0.6083
	128	0.6104
	64	0.6081
	32	0.6177
	16	0.6255
	8	0.5996
	4	0.4447
	2	0.2660

a significant impact on the accuracy of the decoder as the learning mechanism will adapt accordingly. Specifically in this experiment, 8 classes are selected from the WikiArt dataset with over-sampling to balance the number of images from each class, totaling to 36443 images with a 80:20 ratio for training and validation. Adam optimization was employed for the training with the learning rate of 0.0001 and 0.001 for Generator and Discriminator respectively.

A. Secret Retrieval Performance

In our experiments we look into the classification accuracy as evaluation because in this work, the secret retrieval accuracy is equivalent to the classification accuracy of the Discriminator D' when given a generated image. Additionally, we also used clustering homogeneity to examine if the image classes can be inferred from a collage image made up of 16×16 synthesized images. This is to illustrate the scenario, if the collage image is deconstructed back into individual encoded images, would it be possible to deduce their classes and subsequently extract the encoded secret without prior knowledge of the class samples.

1) Classification Accuracy: In the first evaluation, we used the Generator G to generate a test set of synthesized images for classification, specifically 1000 images per class. In order to compare and validate that D' is distinctly trained for the classification, we separately trained other classifiers on the same WikiArt subset as comparison. In particular, we trained a model $\hat{D}$ with the same architecture as D' as well as adapt the notable ResNet50 [18], MobileNetv2 [19], and EfficientNet [20] models through transfer learning. Additionally, we also include an evaluation of these models trained using the generated images instead of the original WikiArt data, to illustrate that separate a classifier can be used as decoder but only if the samples from the encoder is available for training. For this, we used G to generate 4560 images for each class to produce 36480 training images.

As seen in Table II, models that are trained on the WikiArt dataset has significantly low accuracy which means they will not be able to decipher the secret (classes) even if the original training set is known. Separately, their accuracy is on par with the discriminator D' of StArtGAN, indicating a possibility of creating a separate decoder but only if the specific samples are known and used to train the models.

However, it should be noted that while an accuracy above 90% is highly desirable in most applications, it would be detrimental for secret decoding as it affects the final retrieved secret. Nonetheless, in our framework this can be easily avoided by a feedback verification through the discriminator to make sure that the secret image can be classified correctly before it is selected to be used in the collage image for information hiding purposes. By adding this additional step, the secret message can always be decoded correctly during the decoding phase.

2) Clustering Homogeneity: In this second evaluation, we create 100 samples of Output Images I_Σ, each containing 16×16 images of 32×32 pixels generated by G. We then recover each of the collage image I'_ω and subsequently the generated images I'_η. We perform clustering on I'_η of each I'_ω and then evaluate their homogeneity score based on the actual classes S_η we used to generate the initial collage I_ω. We performed two variations of clustering, the first is by directly using the 1024 pixels of I'_η as the input features, while the second is by reducing its 1024 pixel dimensions into smaller dimensions through Principal Component Analysis (PCA). The k-means algorithm is used for clustering with k-means++ initialization and the clusters amount is set as 8.

Table III shows the average homogeneity score of the clustering using the original pixels as well as lower dimensions ranging from 2 to 256. Homogeneity measures if the clusters formed contain only the members of a single class. Based on the results, we can deduce that the clusters in fact do not contain only single class samples. This indicates a significant ambiguity of I'_η, thus would pose a considerable challenge to decipher the secret encoded without the specifically trained decoder, providing an added advantage against eavesdropping.

B. Embedding Capacity

In the proposed framework as explained in the prior section, each of the generated secret images will carry $\log_2 \Lambda$ bits of secret message in binary form. In our experiments, a scrambled-collage image joined by 16×16 generated secret images are generated. Hence, in this case, the embedding capacity in the scrambled-collage image is $16 \times 16 \times \log_2(8) = 768$ bits. However, it is important to note that the embedding capacity is scalable, as we can flexibly control the size of the collage image N or the number of class Λ in the model, to accommodate the secret messages. The embedding capacity shown in the table is an example to illustrate the application of the framework.

Besides, more embedding capacity can be carried since the scrambled-collage image is extended to three channels. In

TABLE IV: Embedding capacity comparison.

Method	Absolute Capacity
Method in [8]	8 bits
Method in [10]	1536 bits
Method in [11]	4480 bits
Method in [12]	1600 bits
Method in [14]	300 bits
Method in [9]	80 bits
Method in [13]	4 words
Method in [21]	2 words
Proposed Method	768-2304 bits*

** embedding capacity is scalable.*

our experiments, two channels and three channels embedding are also explored to carry two times and three times of the calculated embedding capacity in the previous paragraph.

The embedding capacity of the proposed method and the related works are summarized in Table IV. From the table, the proposed method shows good performance in terms of the embedding capacity, since the embedding capacity can be controlled by using parameters N and Λ, as well as the utilization of multiple channels after applying duotone effect.

C. Image Quality

Figure 2 shows the output images of the embedding process in the proposed framework. On the first row of the figure, four scrambled-collage images $I_{\Sigma}^{\{1,\cdots,4\}}$, are generated using four distinctive sets secret images generated from our model. The following rows of output images are the extended images with duotone and/or overlay effect(s) using these I_{Σ}.

In the proposed framework, additional effects are suggested to enhance the inconspicuous level of the output images. The first effect that applied to the I_{Σ} is the duotone effect. Second row of Fig. 2 demonstrate the 1-channel to 3-channel image mode extension of the scrambled-collage images. On Fig. 2(e)-(g), either one of the channel, i.e., red, green or blue channel, is utilized respectively to carry the scrambled-collage image (viz., which also represents the secret images), while the intensities of other two channels are replaced with zero values. This create red, green and blue duotone scrambled-collage images. If different color of duotone is desired (apart from the sharp red, green and blue), the zero-ed out channels can be fine-tuned to create the desired duotone color effect, e.g., Fig. 2(h) shows the output of embedded red channel and fine-tuned blue channel.

We also explored the possibilities of embedding more capacity by using the proposed framework, since there are available spaces after the duotone effect. Hence, double capacity embedding on two channels and triple embedding capacity on three channels are tested and the output images are as shown in Fig. 2(i)-(l).

Finally, the experiment is completed by applying overlay effect on the duotone scrambled-collage image to enhance the aesthetic level of the output images. The output images shown in Fig. 2(m)-(p) are the examples of the overlay effect. The duo-tone output image with overlay effect, viz., in RGBA format, is a possible medium to produce an image to carry the secret inconspicuously. Alternative methods can

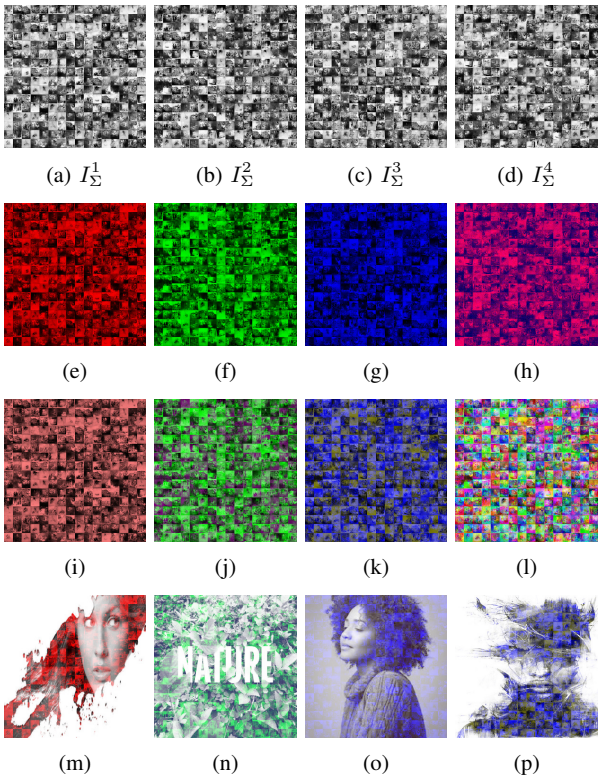

Fig. 2: Output images with hidden data generated by the proposed framework. First row shows four scrambled-collage-images, $I_{\Sigma}^{1}, I_{\Sigma}^{2}, I_{\Sigma}^{3}, I_{\Sigma}^{4}$, joined by using four different sets of generated secret images. Second row shows the duotone effect by using the proposed framework on different channels, in particular, (e) shows red duotone effect, (f) and (g) shows green and blue duotone effects, while (h) shows red duotone and fine tune on blue channel. Third row shows the outcome images if double capacity embedding is applied on two channels (refer to (i) - (k)) and triple capacity is applied on three channels (refer to (l)). Last row ((m) - (p)) shows the output images with various overlay effects.

be applied for this purpose, but essentially the secret is in the 1-channel image generated from the StArtGAN. Other arrangements, effects or masking can also be applied to the secret images, scrambled-collage image or the duotone image to accommodate the environment and communication needs. Nevertheless, based on our subjective point of view, the final output images show that the proposed framework can achieve the goal of coverless image steganography, which aims at producing inconspicuous output image with hidden data.

For the output image quality evaluation, objective measures are not applicable to this framework because this is a coverless steganography method hence there is no reference image to perform quality comparison on quality metrics such as PSNR and Structural Similarity (SSIM). Secondly, non-reference quality metrics evaluate the naturalness of realistic images, however the images produced in this framework are more towards artistically edited images, thus the evaluation would

(a) Secret Images (b) Recovered Secret Image (c) Pixel Intensities Difference

Fig. 3: The difference of the generated and recovered secret images are compared and displayed in (c).

not be suitable with these measures as well.

The recovered secret images are compared with the original secret images to prove the recovery performance. As shown in Fig. 3, their pixel intensities difference appear in black, indicates the recovered image is identical to the original secret image.

The PSNR value and SSIM index are used as a referential evidence for quality evaluation for distortion of the restored image. The PSNR expressions are given below.

$$\begin{aligned} \mathrm{PSNR} &= 10 \times \log\left(\frac{255^2}{\mathrm{MSE}}\right) \\ \mathrm{MSE} &= \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(x(i,j)-y(i,j))^2 \end{aligned} \tag{5}$$

The SSIM expression is given below.

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)\,(2\sigma_{xy} + c_2)}{(u_{ir}^2 + u_u^2 + c_1)\,(\sigma_r^2 + \sigma_u^2 + c_2)} \tag{6}$$

As a result, the PSNR and MSE values between the two images are infinity, indicating that both images are exactly identical. Therefore, we can conclude that full recovery in the collage process can be achieved by using our proposed method.

IV. Conclusion

In this work, we proposed a framework to increase the secrecy of encoded message through relabeling, scrambling and image synthesis. We demonstrated that image synthesis by our StArtGAN model is not only able to construct images from a secret message as a medium of covert communication, but the discriminator of the model can act as a dedicated decoder for accurate message recovery. Our framework is scalable by various number of classes, various size of the collage image and various channels can be utilized to control the embedding capacity. In the future, we intent to focus on enhancing the security of the proposed model and perform more study and comparison with other relevant works.

Acknowledgment

This publication shares the outcome of the undergoing project, entitled Steganography by Image Synthesis: A Novel Model for Coverless Data Hiding, supported by Fundamental Research Grant Scheme (Grant ID:FRGS/1/2019/ICT02/UM/02/1) awarded by the Malaysia Government. This research is also supported in part by the Internal Research Fund (IRFund) MMUI/210003 provided by Multimedia University.

References

[1] N. Jiang, N. Zhao, and L. Wang, "Lsb based quantum image steganography algorithm," *International Journal of Theoretical Physics*, vol. 55, no. 1, pp. 107–123, 2016.

[2] K. A. Al-Afandy, O. S. Faragallah, A. ELmhalawy, E.-S. M. El-Rabaie, and G. M. El-Banby, "High security data hiding using image cropping and lsb least significant bit steganography," in *2016 4th IEEE International Colloquium on Information Science and Technology (CiSt)*, pp. 400–404, IEEE, 2016.

[3] S. Chakraborty, A. S. Jalal, and C. Bhatnagar, "Lsb based non blind predictive edge adaptive image steganography," *Multimedia Tools and Applications*, vol. 76, no. 6, pp. 7973–7987, 2017.

[4] Y. Jia, Z. Yin, X. Zhang, and Y. Luo, "Reversible data hiding based on reducing invalid shifting of pixels in histogram shifting," *Signal Processing*, vol. 163, pp. 238–246, 2019.

[5] Zhicheng Ni, Yun-Qing Shi, N. Ansari, and Wei Su, "Reversible data hiding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, pp. 354–362, March 2006.

[6] S. Mukherjee and B. Jana, "A novel method for high capacity reversible data hiding scheme using difference expansion," *International Journal of Natural Computing Research (IJNCR)*, vol. 8, no. 4, pp. 13–27, 2019.

[7] H. S. El-sayed, S. El-Zoghdy, and O. S. Faragallah, "Adaptive difference expansion-based reversible data hiding scheme for digital images," *Arabian Journal for Science and Engineering*, vol. 41, no. 3, pp. 1091–1107, 2016.

[8] Z. Zhou, H. Sun, R. Harit, X. Chen, and X. Sun, "Coverless image steganography without embedding," in *International Conference on Cloud Computing and Security*, pp. 123–132, Springer, 2015.

[9] L. Zou, J. Sun, M. Gao, W. Wan, and B. B. Gupta, "A novel coverless information hiding method based on the average pixel value of the sub-images," *Multimedia tools and applications*, vol. 78, no. 7, pp. 7965–7980, 2019.

[10] K.-C. Wu and C.-M. Wang, "Steganography using reversible texture synthesis," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 130–139, 2014.

[11] W. K. Lee, S. Ong, K. Wong, and K. Tanaka, "A novel coverless information hiding technique using pattern image synthesis," in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 1122–1127, IEEE, 2018.

[12] Z. Qian, H. Zhou, W. Zhang, and X. Zhang, "Robust steganography using texture synthesis," in *Advances in Intelligent Information Hiding and Multimedia Signal Processing*, pp. 25–33, Springer, 2017.

[13] M.-m. Liu, M.-q. Zhang, J. Liu, Y.-n. Zhang, and Y. Ke, "Coverless information hiding based on generative adversarial networks," *arXiv preprint arXiv:1712.06951*, 2017.

[14] 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.

[15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, pp. 2672–2680, 2014.

[16] A. Odena, C. Olah, and J. Shlens, "Conditional image synthesis with auxiliary classifier gans," in *International conference on machine learning*, pp. 2642–2651, 2017.

[17] W. R. Tan, C. S. Chan, H. E. Aguirre, and K. Tanaka, "Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification," in *2016 IEEE international conference on image processing (ICIP)*, pp. 3703–3707, IEEE, 2016.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[19] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4510–4520, 2018.

[20] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning*, pp. 6105–6114, PMLR, 2019.

[21] M. Liu, M. Zhang, J. Liu, and X. Yang, "Generative steganography based on gans," in *International Conference on Cloud Computing and Security*, pp. 537–549, Springer, 2018.