IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 14 September 2024, accepted 3 October 2024, date of publication 10 October 2024, date of current version 21 October 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3477469

RESEARCH ARTICLE

Coverless Steganography for Face Recognition Based on Diffusion Model

YUAN GUO AND ZIQI LIU
School of Computer Science and Technology, Heilongjiang University, Harbin 150080, China
Corresponding author: Ziqi Liu (q373161930@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 62473278; in part by Heilongjiang Province Natural Science Foundation under Grant LH2021F056; and in part by Heilongjiang Provincial Education Department, Grant/Award, under Grant 1355091130.

ABSTRACT As a highly recognizable biometric face recognition technology, it has been widely used in many identity verification systems. In order to enhance the protection of personal privacy and ensure the safe transmission and sharing of sensitive information without affecting the user experience, this paper proposes an innovative coverless steganography framework for face recognition images based on diffusion model. The framework firstly extracts face features and generates masks containing these features. Then, combined with conditional diffusion model and text key, a deterministic Denoising Diffusion Implicit Model (DDIM) is used to sample coverless steganography images. Secret images can also be recovered in high quality with DDIM Inversion technology. A large number of experiments show that compared with the existing methods, this approach has markedly enhanced the quality of steganographic and restored images. The face recognition rate of the restored image is more than 96%, which can effectively replace the original image for face recognition. The detection accuracy of this method is 55.25% on the steganographic detection tool, which is closer to random guessing and can resist steganographic analysis. It ensures the higher security of hidden images and solves the limitation of existing methods in protecting the privacy of face images. Moreover, it is shown how to achieve controlled local steganography with a custom mask, which enhances the controllability and flexibility of the method. In conclusion, the proposed method outperforms traditional steganography in security, controllability and robustness, and provides an effective technical scheme for steganography protection of face recognition images without additional training.

INDEX TERMS Face recognition, coverless steganography, diffusion model, DDIM.

I. INTRODUCTION

Facial recognition technology is a branch of computer vision tasked with identifying and confirming a person's identity by analyzing facial characteristics. [1]. In the past few years, it has become a successful application of pattern recognition and machine learning. Today, facial recognition technology is extensively employed across numerous applications, from security and surveillance to entertainment and social media [2], [3]. Because face recognition is greatly affected by external environmental factors, it poses a challenge to the privacy protection and security of face images. Steganography, as a popular technique nowadays, aims to hide secret information in the host medium [4]. Whereas image steganography hides information into containers in an imperceptible way, only a trusted recipient can recover the information from the steganography [5]. Image steganography specializes in disguising secret messages as images, which can provide a high degree of security and privacy protection for face recognition. Today, image steganography is used in many fields, including copyright protection, digital watermarking, secure information transmission, and digital forensics [6], [7], [8], [9].

The traditional cover-based image steganography scheme hides secret information in the cover image by subtly changing its statistical properties, usually by transforming the secret message in both spatial and adaptive domains [10]. Among the common techniques, the Least Significant Bit

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

(LSB) method alters the least significant bits of pixel values to embed data, preserving the visual quality while enabling a high capacity for data hiding [11]. Prediction error methods exploit the differences between predicted and actual pixel values to insert information, adapting to the image content for enhanced security [12]. Histogram-based methods modify pixel value histograms, subtly shifting them to encode data without significant visual changes [13]. Secret sharing methods divide and distribute data across the image, requiring multiple parts to reconstruct the secret, thereby increasing security [14]. Modular operations use arithmetic computations to embed data, providing robustness against statistical analysis [15]. Lastly, quantization-based methods adjust the quantization levels of image components to hide information, balancing imperceptibility and payload [16]. With the rise of deep neural networks, deep steganography such as autoencoder networks or invertible neural networks (INN) are widely used to hide data [17], [18], [19]. These techniques each offer unique advantages in maintaining the fidelity and security of the embedded data within steganographic practices. Therefore, in recent years, many coverless data hiding methods have been proposed to solve the problem, and secret messages can be hidden without any modification to the cover image [20]. It is fundamentally resistant to steganography and significantly improves security. Current coverless steganography methods are often implemented using frameworks such as CycleGAN and encoder-decoder model [21], [22]. However, the control over container images produced by current coverless techniques remains restricted, typically confined to embedding bits within the container image while overlooking the intricacies of concealing more complex secret images [23]. In general, current steganography methods, whether based on overlay or without overlay, are faced with challenges and limitations in pursuing a good unity of security, controllability and robustness.

In recent years, diffusion model has been widely used in many image fields because of its excellent generating ability. Inspired by diffusion model, the limitations of existing methods are further improved. At present, the generation model based on diffusion is very popular [24], [25]. The concept involves deterministically processing the image by incrementally introducing noise to the dataset, then learning to reverse this process. It aims to diffuse and then invert the noise trajectory under the same conditions, ultimately producing high-quality data. Its core function is to gradually reduce the quality of data before rebuilding it back to its original form or generating new data. Simultaneously, the diffusion model possesses several distinct characteristics, including the capability to execute tasks without any samples [26], [27], [28], good control of the generation process [29], [30], natural robustness to noise in the image [27], [31], [32], and image-to-image generation [33], [34], [35]. Diffusion model has shown broad application prospects in many fields due to its characteristics of gradually de-noising and generating high quality data. The powerful control ability of conditional diffusion model makes steganographic images highly controllable. At the same time, the diffusion model also has natural robustness. Even if the steganography image undergoes degradation or interference during transmission, it is still possible to uncover the primary content of the secret image.

Therefore, in order to further overcome the current security challenges of face recognition. A face recognition image steganography framework based on diffusion model is proposed, aiming to achieve a more secure, controllable and robust face image steganography. Specifically, many attributes of diffusion model are combined, and the robustness of face image steganography is improved. At the same time, DDIM Inversion, a technique using Denoising Diffusion Implicit Models, is employed to achieve a coverless steganography framework [36]. This ensures that hidden images have higher security and play a more important role in information security and privacy protection.

The main contributions of this dissertation are as follows:

- In this paper, a novel face image coverless steganography technique based on diffusion model is proposed: facial feature masking is combined with conditional diffusion model, and DDIM is used to achieve inversion. A steganography framework is implemented specifically for face images without any additional complex training process.
- Stable Diffusion inpainting is proposed to realize the coverless steganography of face image, so that the generated steganography and recovered image have higher quality. It is applied to face recognition image to ensure higher security of hidden image, which improves the limitation of privacy protection of face image by existing methods. Moreover, the custom mask is combined to achieve controllable local steganography, which enhances its controllability and flexibility.
- Experimental results on the CelebA-HQ and FFHQ public datasets demonstrate that this method offers considerable advantages over existing approaches in terms of both network environment and real-world degradation. It can resist steganalysis well and successfully achieve better reconstruction quality. The Face++ face recognition model obtains more than 96% confidence and closer Distance, which can effectively replace the original image for face recognition, and has higher robustness and security.

The remainder of this paper is organized as follows. Section II provides a detailed review of the related work, emphasizing the evolution of steganography techniques and their applications in face recognition. Section III describes the proposed steganography framework, detailing the integration of facial feature masking with the diffusion model. Section IV presents the experimental setup and results, demonstrating the effectiveness of the proposed method. Section V

summarizes the results of this study. Finally, Section VI looks into the future of image steganography.

II. RELATED WORK

A. STEGANOGRAPHY

Traditional image steganography is categorized into spatial and frequency domains. In the spatial domain, methods like the Least Significant Bit (LSB), Pixel Value Difference (PVD), Histogram Shift, Multibit Plane, and Palette-based steganography directly alter pixel values, offering high capacity but potential visual distortion [11], [13], [37], [38], [39]. To enhance robustness, frequency domain methods such as Discrete Cosine Transform (DCT), Discrete Fourier transform (DFT), and Discrete Wavelet Transform (DWT) encode information in the image's frequency components [40], [41], [42]. These traditional steganography methods do not consider the characteristics of the medium, and use fixed, unchanging rules to insert information. The lack of flexibility and adaptability of such methods makes steganalysis tools more likely to find anomalies by studying these rules, causing statistical suspicion. Deep learning significantly enhances image steganography. Zhu et al. developed HiDDeN, an end-to-end trainable model for both steganography and watermarking that optimizes capacity, confidentiality, and robustness [19]. Shi et al. combined GAN with LSB in Ssgan for better security [43]. Zhang et al.'s SteganoGAN set a new benchmark with a payload of 4.4 bits per pixel and robust detection evasion [44]. Baluja and Zhang et al. demonstrated methods for hiding full-size secret images [16], [45]. Peng et al. introduced a superior steganography model based on denoising diffusion probability [46]. Dinh et al., HiNet, and Jia et al. utilized invertible neural networks (INN), enhancing the flexibility and efficiency of steganographic applications [47], [48], [49].

Coverless image steganography, which embeds data without altering the original media, is evolving as a secure communication method. Zhou et al. introduced a coverless image steganography method using puzzle images created from secret messages [50]. Mu and Zhou proposed using visual vocabulary trees for retrieving steganographic images that share patches with secret images [51]. Liu et al. leveraged DWT sequence mapping and DenseNet for image retrieval [52]. Luo et al. developed a coverless data hiding method using image block matching and the DenseNet model [53]. Lu et al. introduced a non-overwritten hiding method using unsupervised learning [54]. Recently, Yu et al. developed a reversible image conversion technique utilizing a diffusion model, marking its first use in steganography [55].

These methods often rely on complex image retrieval and matching techniques, which not only increases the difficulty of implementation, but also may affect the efficiency of embedding and the feasibility of real-time communication. In addition, the practicality and efficiency of these technologies have yet to be validated and optimized in the face of large-scale practical applications.

B. FACE RECOGNITION

Recent advancements in deep learning have significantly enhanced face recognition technology, making traditional methods based on manually engineered features obsolete. DeepFace by Taigman et al. demonstrated that deep learning models, when trained on extensive datasets, deliver superior recognition performance [56]. He et al. developed a deep residual network that not only enables deeper neural networks with high accuracy but also enhances the feature representational capacity [57]. Hu et al. introduced a feature recalibration method that optimizes channel importance, enhancing significant features while reducing lesser ones [58]. Chen et al. tailored MobileNet for face recognition, striking a balance between speed and accuracy [59]. Li et al. developed a novel coverless information transmission method using CNNs for morphed face recognition, integrating secure data hiding with facial recognition [60]. Abusham et al. introduced an image encryption method enhancing facial recognition systems by preventing spoofing [61]. Liu et al. advanced adversarial attack algorithms targeting face identity, incorporating diffusion models for imperceptible manipulations, representing significant progress in adversarial machine learning for face recognition [62].

Advanced face recognition technologies are susceptible to adversarial attacks and exhibit biases inherent in their training datasets. These issues compromise their reliability and fairness, posing significant challenges for their application in sensitive areas like security and personal identification.

C. DIFFUSION MODELS

The diffusion model, initially proposed by Sohl-Dickstein et al. in 2015 [63], has become widely popular for its robust generative capabilities in domains such as image generation, restoration, and translation [26], [27], [28], [30], [34], [35]. Its versatility enhances and manipulates digital images effectively. However, the model's primary shortcoming is the prolonged training and inference times, prompting research into optimization techniques [36]. The latent diffusion model (LDM) enhances efficiency by supporting high-resolution synthesis for general conditional inputs like text or bounding boxes [30]. Additionally, the text inversion scheme improves model controllability by reconstructing user-specified concepts from a few images [64]. Techniques like masked image editing and DreamBooth enable personalized content creation, while cue-based image editing technologies allow for prompt-driven image manipulation [65], [66], [67]. Huang et al.'s work explores advanced methods for controlling diffusion models [68]. The potential for diffusion-based, coverless image steganography looks promising due to the model's rapid development and powerful generation abilities [55].

Diffusion models, despite their impressive generative capabilities, suffer from high computational costs and long training times, hindering their adoption for real-time applications. Inconsistencies in image quality and the complexity of

model tuning further challenge their practical usability and widespread implementation.

III. METHOD

In the following sections, the implementation of the approach using the diffusion model and the FaceParsing model is detailed [69]. Specifically, In section A the relevant definitions in steganography are given.In section B, the principle of DDIM is analyzed. In Section C, face image steganography without overlay in combination with a mask is described in depth. In section D, security is analyzed. In section E, the calculation formula of face recognition is given.

A. RELEVANT DEFINITIONS IN THIS STEGANOGRAPHY

Before detailing the methods of this article, let's first clearly define what constitutes an image steganography task, as shown in Figure 1. This task involves four types of images: the secret image, the secret image mask, the steganographic image, and the recovered image. Two key processing processes: the hiding process and the revealing process. The secret image is what want to hide. To precisely control this process, the FaceParsing model is used to extract the mask from the secret image. This mask, along with the secret image, is embedded in the steganographic image through a hiding process. When steganographic images are transmitted over the Internet, the image quality may decrease due to various factors, resulting in a degraded steganographic image. Even so, revealing process can still be combined with mask to recover the recovered image from degraded steganographic image, maintaining semantic consistency of the content. The framework design proposed in this paper focuses on several key attributes:

1) Higher image quality: With mask, it is possible to precisely control the content of the generated image and avoid interference from external factors such as the background. This ensures that steganographic images are not only rich in content, but also have good visual quality.
2) Security: Hiding methods are designed that are difficult to detect by the naked eye, even when faced with steganographic tools.
3) Robustness: Even if the degraded steganographic image is slightly different from the original, revealing process can still generate semantically consistent restored images from it.

B. DDIM FOR IMAGE REVERSIBLE TRANSFORMATION

The DDIM is a diffusion model that utilizes deterministic inference to generate high-quality images. This model aims to improve the generation process of traditional diffusion models by reducing randomness, thereby enhancing the quality and efficiency of the generated samples.

DDIM defines its diffusion model through two main phases: the forward phase and the reverse sampling phase. In the forward phase, the model gradually adds noise to a clean image, simulating the process of the image becoming progressively distorted. Specifically, the forward process in DDIM [36] can be described by the following (1) equations:

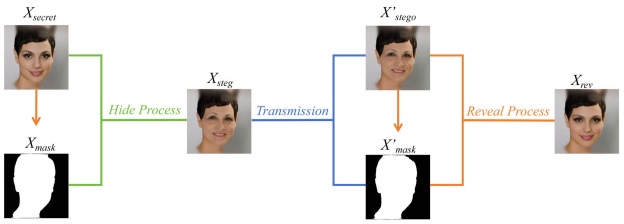

FIGURE 1. The definition and composition of steganography of face image.

$$x_t = \sqrt{\alpha_t} x_{t-1} + \sqrt{1-\alpha_t}\epsilon, \quad \epsilon \sim \mathcal{N}(0,1) \tag{1}$$

where α_t is a pre-defined noise level parameter, ϵ is the random noise sampled from a standard Gaussian distribution, and x_t is the image state at time step t. The range of the time step t is from 1 to T, denoted as $[1, T]$.

In the reverse sampling phase, the model adopts the inverse process, gradually restoring the clean image by estimating and removing the noise. This process not only reduces random variations during generation but also improves the clarity and detail representation of the image, thereby generating more realistic and high-quality images. The reverse sampling process of DDIM can be described by the following (2) equations:

$$\begin{aligned} x_s &= \sqrt{\bar{\alpha}_s} f_\theta(x_t, t) + \sqrt{1-\bar{\alpha}_s - \sigma_s^2}\epsilon_\theta(x_t, t) + \sigma_s \epsilon, \\ f_\theta(x_t, t) &= \frac{x_t - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(x_t, t)}{\sqrt{\bar{\alpha}_t}} \end{aligned} \tag{2}$$

In the Denoising Diffusion Implicit Models (DDIM), ϵ is sampled from a Gaussian distribution $\mathcal{N}(0, 1))$ with σ_s^2 representing the noise variance. The function $f_\theta(., t)$ utilizes a pre-trained noise estimator $\epsilon_\theta(., t)$, and $\bar{\alpha}_s$ is defined as the product of α_i from 1 to t: $\bar{\alpha}_s = \prod_{i=1}^{t} \alpha_i$. DDIM uniquely allows for non-adjacent sampling steps, meaning t and s can take any two steps where $s < t$, enhancing the flexibility and speed of the sampling process. Additionally, setting the noise variance σ_s to zero in (2) makes the DDIM sampling deterministic. In this scenario, the outcome is completely defined by the initial value x_T, which acts as a latent encoding. This deterministic process can also be described through the framework of an ordinary differential equation (ODE [36]), where an ODE solver is employed to resolve the corresponding ODE, providing a systematic approach to model the diffusion process.

The diffusion model is implemented using deterministic DDIM, which not only simplifies the model's complexity but also enhances its predictability and controllability. Pre-trained noise estimators are relied upon to precisely control the noise removal process, thereby optimizing the quality and efficiency of the entire image generation. With this method,

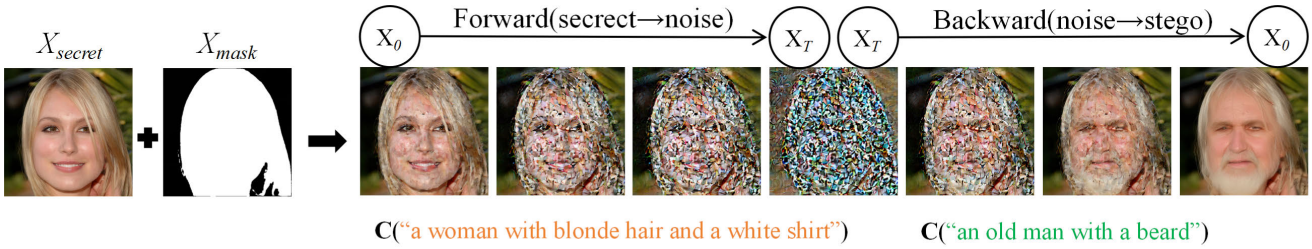

FIGURE 2. A conditional diffusion model is employed for image translation across various scenarios. In this instance, two distinct prompts are utilized to transform an image of a woman into an image of an old man.

it is possible to effectively accelerate the image generation process while maintaining the high quality of the generated images. The formula representing sampling to process using the pre-trained noise estimator is as shown in Equation (3):

$$x_0 = \text{ODESolve}(x_T; \epsilon_\theta, T, 0) \tag{3}$$

In the approach to image genteration using diffusion models, a conditional diffusion model that leverages conditions to steer the generation process is employed. Specifically, conditions such as text descriptions and masks are incorporated into the model to guide the precise crafting of image content. As illustrated in Figure 2, the process involves transforming an image of a woman into an image of an elderly man. Initially, during the forward sampling stage, the Algorithm (1) is applied to introduce noise into the female image, leading to an intermediate noisy state. Subsequently, for the backward sampling phase detailed in Algorithm (2), a specific text condition (prompt: "an old man with a beard") is input. This text prompt serves as a directive for the noise removal process, facilitating the generation of an image depicting an elderly man with a beard. Both the text condition (c) and mask (X_{mask}) are utilized as input conditions. The sampling process that iteratively refines the image from the noisy state (x_T) back to the clean state (x_0) is executed using the pre-trained noise estimator ϵ_θ as shown in Equation (4):

$$x_0 = \text{ODESolve}(x_T, X_{mask}; \epsilon_\theta, c, T, 0) \tag{4}$$

Algorithm 1 The Hide Process

Input: The pre-trained conditional diffusion model for the secret image X_{secret}, the mask X_{mask}, the noise estimator X_{noise}, the sampling time step T, and two different conditions K_{pri} and K_{pub} as private and public keys.
Output: Steganographic image X_{stego} for hiding secret image X_{secret}.
$X_{mask} = \text{GetMask}(X_{secret})$
$x_{noise} = \text{ODESolve}(X_{secret}, X_{mask}; \epsilon_\theta, K_{pri}, 0, T)$
$x_{stego} = \text{ODESolve}(X_{noise}, X_{mask}; \epsilon_\theta, K_{pub}, T, 0)$
return X_{stego}

To achieve reversible image transformation, the DDIM Inversion method based on deterministic DDIM is employed. As the name suggests, this method converts the image to potential noise and then back to the original image. The concept draws on the approximation of forward and backward differentials used in solving ordinary differential equations. Specifically, in deterministic DDIM, it enables steps s and t as described in Algorithm (1), where $s$and t represent discrete time steps in the diffusion process, enabling the model to navigate the noise addition and removal processes in non-sequential order. Equation (2) permits any two steps (s < t and s > t).When $s < t$, the method initiates the backward process, and when $s > t$, it triggers the forward process. Due to the parallel nature of the trajectories in both forward and backward processes, the initial and resultant images are closely aligned, rendering the intermediate noise x_T as an effective inverted latent variable. The formula is shown in Equation (5):

$$\begin{aligned} x_T &= \text{ODESolve}(x_0, X_{mask}; \epsilon_\theta, c, 0, T), \\ x'_0 &= \text{ODESolve}(x_T, X'_{mask}; \epsilon_\theta, c, T, 0) \end{aligned} \tag{5}$$

DDIM Inversion describes the transformation where the original image x_0 is converted to a latent code x_T, and subsequently, this latent code x_T is reverted back to the original image, with the output image being denoted as x'_0 and approximately equal to x_0. Using the DDIM Inversion method, a reversible connection between the image and its latent noise is created. By utilizing the image translation framework constructed with deterministic DDIM, the entire reversible image transformation can be completed through two DDIM Inversion cycles. This technique not only serves as the core of the coverless image steganography framework but also is key to ensuring the reversibility of the steganography process. The reversibility of this method means that even in complex image processing, the integrity and accuracy of the image content can be maintained.

Algorithm 2 The Reveal Process

Input: There may be degradation through the steganographic image X'_{stego} after transmission, according to the X'_{mask} generated by X'_{stego}, a pre-trained conditional diffusion model with noise estimator ϵ_θ, sampling time steps T, private key K_{pri}, and public key K_{pub}.
Output: The revealed image X_{rev}.
$X'_{mask} = \text{GetMask}(X'_{stego})$
$x'_{noise} = \text{ODESolve}(X'_{stego}, X'_{mask}; \epsilon_\theta, K_{pub}, 0, T)$
$x_{rev} = \text{ODESolve}(X'_{noise}, X'_{mask}; \epsilon_\theta, K_{pri}, T, 0)$
return X_{rev}

FIGURE 3. A conditional diffusion model is opted that accommodates conditional inputs to steer the outcomes of generation. Furthermore, deterministic DDIM is employed as the sampling approach and two distinct conditions specified by the model (K_{pri} and K_{pub}) serve as the private and public keys, respectively.

C. FACE STEGANOGRAPHY BASED ON DIFFUSION MODELS

The framework is constructed around a conditional diffusion model that incorporates a noise estimator, utilizing a mask and two distinct conditions as inputs to the diffusion process. In our implementation, these conditions function as private and public keys, denoted as K_{pri} and K_{pub} respectively, with the complete workflow depicted in Figure 3. This paper is divided into two main sections to discuss our coverless steganography framework: the hiding process and the revealing process.

Hiding Process: During the concealment phase, the transformation between the secret image X_{secret} and the steganographic image X_{stego} is facilitated via the deterministic DDIM's forward and backward processes. To ensure variability in the images pre- and post-transformation, the pre-trained conditional diffusion model is engaged with differing conditions for each process. These conditions also serve dual roles as private and public keys (K_{pri} and K_{pub}). Specifically, a generated mask X_{mask} from the original secret image is used to control the depiction of people independently from the background and other elements, employing K_{pri} in the forward process and K_{pub} in the reverse. The resulting steganographic image X_{stego} is then sent across the Internet, accessible to all potential recipients. This setup hinges on the effectiveness of the conditions: the private key outlines the content of the secret image, while the public key influences the steganographic image's content. In this model, the public key is inferable from the steganographic image itself, thus, it need not be transmitted separately. Conversely, the private key is crucial for accurate image recovery and must remain confidential.

Revealing Process: In the recovery stage, it is assumed the steganographic image X'_{stego} has been transmitted online and possibly altered. The recipient utilizes the same conditional diffusion process with the corresponding keys, employing a reverse sequence to the hiding process, to restore the original secret image. This involves regenerating a control mask from the steganographic image X'_{stego}, now called X'_{mask}, using K_{pri} in the forward process. Unlike the hiding phase, where K_{pub} is used forward and K_{pri} backward, the revealing phase adjusts these roles. This method of coverless image steganography doesn't require training or fine-tuning the diffusion model specifically for steganography tasks; rather, it leverages the inherent reversible image transformation capabilities of DDIM Inversion. The forthcoming section will delve into this framework's specific applications and operational details, demonstrating its efficacy in safeguarding the privacy and security of image content in real-world scenarios.

D. SECURITY ANALYSIS

The proposed approach may raise some security concerns. For example, what happens if the recipient can guess the private key? What is the quality of steganographic image generation? Is it easy for people to see that there might be secret information hidden in the images? To answer these questions, the details are explained from the following two aspects (as shown in Figure 4).

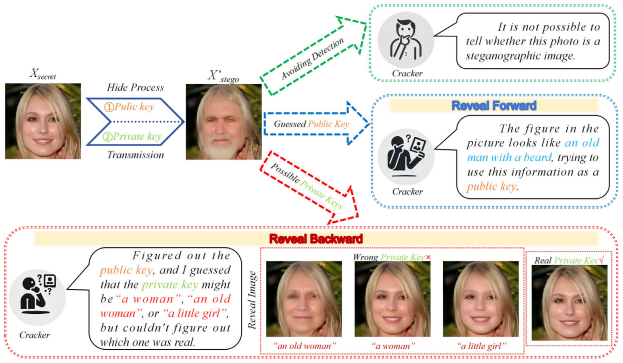

FIGURE 4. Different problems that receivers may encounter during the recovery process after receiving steganographic images are simulated to further illustrate security.

Image quality and security of private key: Our steganographic image is produced using an efficient diffusion model, ensuring that the visual quality remains high regardless of the accuracy of the input private key. Typically, it is challenging to identify any irregularities in the image with the naked eye. Although the recipient may attempt to use the exhaustive method to test the private key. The security of this method is improved by the fact that it is impossible to clearly identify the actual secret image among the many potential recovered images. This strict protection of private keys ensures that the security of information is not compromised even in cases where the key is partially guessed.

The concealment of steganographic images: Since the steganographic images are produced by diffusion models, their visual quality is guaranteed by the generative priors inherent in the model. Unlike traditional steganography methods that rely on covering media, our coverless steganography technique ensures that the steganographic image does not contain any detectable clues. This method not only hides the secret information in the image, but also makes the Area Under the Curve (AUC) value close to the random prediction 0.5 under the detection of steganalysis tools, so that these information cannot be discovered or extracted by traditional image analysis tools. This presents a great challenge for the recipient to identify other images hidden in the steganographic image or to successfully recover the secret image with existing technical means or analytical methods.

E. FACE RECOGNITION

When comparing the original/recovered image and the test image, Cosine Similarity is usually used to measure the distance between the two images. Cosine distance is used to calculate the similarity between two images, and the value ranges from 0 to 1. A value of 1 indicates complete dissimilarity (maximum distance), while a value of 0 indicates complete similarity. The smaller the distance, the higher the similarity between the images. The formula is shown in Equation (6):

$$\text{Distance} = 1 - \frac{A \cdot B}{\|A\|\|B\|} \tag{6}$$

where A represents the feature vector of the original/recovered image, and B represents the feature vector of the test image. $A \cdot B$ represents the dot product of the two vectors, and $\|A\|\|B\|$ are the norms (or lengths) of vectors A and B, respectively.

The confidence score reflects the system's confidence in the face recognition match results, usually expressed as a percentage. The higher the score, the more accurate the system believes the match between the two images is. The confidence score represents the similarity between two images, where the similarity score is calculated based on the similarity of the feature vectors, and the confidence percentage is obtained after normalization. The formula is shown in Equation (7):

$$\text{Confidence} = \frac{\text{Similarity Score}}{\text{Max Similarity Score}} \times 100 \tag{7}$$

where Similarity Score represents the matching score between two images, and Max Similarity Score represents the maximum possible matching score (usually a predefined upper limit).

IV. EXPERIMENTAL RESULTS

A. IMPLEMENTATION DETAILS AND SETUP

Experimental Setup: For the experiments, the FaceParsing model was employed to create facial masks and selected Stable Diffusion v2-Inpainting, hosted by Huggingface, as the conditional diffusion model. Deterministic DDIM Inversion was utilized for the inversion tasks, configuring both the forward and reverse processes to consist of 50 steps each. To facilitate reversible image transformation, the guidance scale was adjusted to 1.0 and the strength set to 0.99.

Data Preparation: Two face datasets were used, including CelebA-HQ [70] (CelebA-HQ is a large-scale face image dataset with 30,000 high-resolution face images selected from the CelebA dataset by following CelebA-HQ) and FFHQ [71] (a high-quality face dataset containing 70,000 high-resolution PNG face images at 1024×1024 resolution), produced a total of 240 face images (called StegFace240). The BLIP [72] model is used to analyze the image and generate the text information describing the image as the private key, and the text information is manually modified as the public key. In order to verify the superiority of our method, it was compared with several advanced image steganography methods. The comparative results underscore the effectiveness of our method, which operates without the

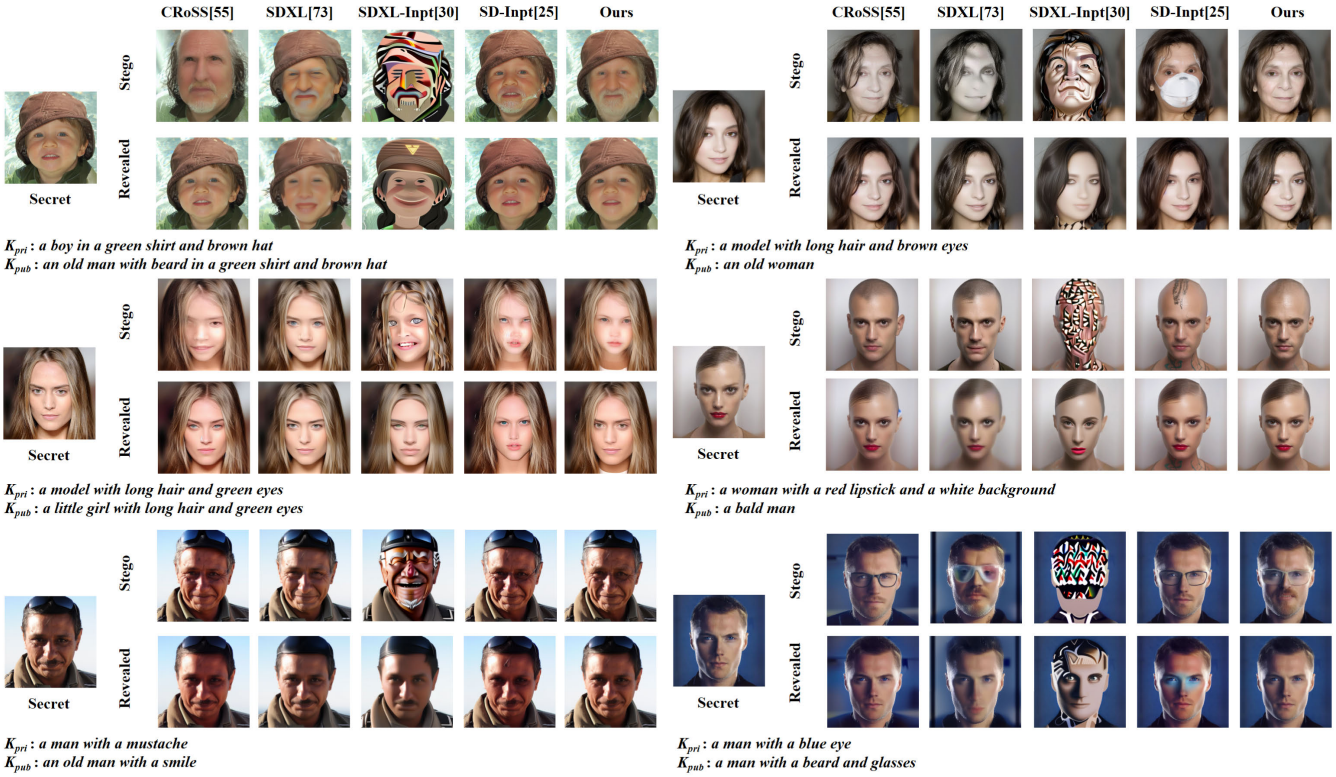

FIGURE 5. Our method compares with other methods for steganography and image recovery. It can be seen that the steganographic image generated by our method has high visual quality. Is not easy to be detected, and the recovered image has a high similarity to the secret image.

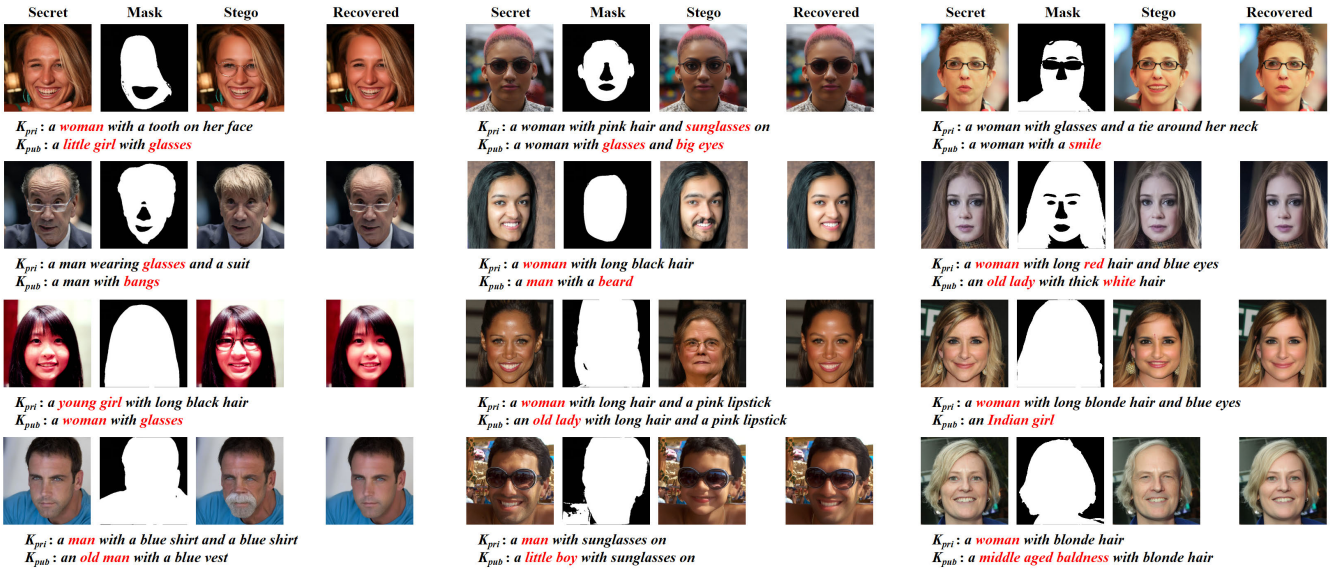

FIGURE 6. Our method combined with custom face mask control, realizes controllable local steganography and has good semantic consistency.

need for additional training. All experiments were conducted using a GeForce RTX 3090 GPU card.

B. COMPARISON WITH SOTA METHODS

In the experiments, the method was compared to different methods such as CRoSS [55], SDXL [73], SDXL-Inpt [30], and SD-Inpt [25], all of which were tested on the StegFace240 dataset. Considering the relatively few applications of Diffusion model in the field of image steganography, several versions of the Stable Diffusion model were used to perform face image steganography, which were implemented by the research team. As shown in Figure 5, the quality of steganographic and recovered images generated by various methods was compared. It can be clearly observed that steganographic

images generated using this method can efficiently hide secret images. While avoiding obvious artifacts or unrealistic image details that are almost undetectable to the naked eye.

In addition, our steganographic images support seamless modification of character characteristics such as gender, age, and facial hair with a high degree of controllability. In terms of controllability (as shown in Figure 6), our approach can perform steganography on certain areas while leaving other areas unchanged. While using the private key to precisely preserve the semantic information of the secret image, demonstrating excellent fidelity.

The method not only allows for highly accurate recovery of the secret image but also minimizes the difference between the original and recovered images. Five different metrics were used to evaluate the quality of secret and recovered images (as shown in Table 1), including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Learned Perceptual Image Patch Similarity (LPIPS), Fréchet Inception Distance (FID), and Listening Diagnostics Metrics (LDM). Among them, PSNR and SSIM, the higher the score of these two indicators, the better the quality of the recovered image. At the same time, the lower the score of LPIPS, FID and LDM, the closer the generated image is to the real image in visual perception, and the more similar it is in visual content and style. This further reduces the possibility of being identified as containing steganographic information. The results show that the method is significantly superior to other methods on the baseline.

TABLE 1. Comparison results of our proposed method and other methods on the StegFace240 dataset. The best results are highlighted in red and the second-best results are highlighted in blue.

Methods	Secret/Reverse				
	PSNR↑	SSIM↑	LPIPS↓	FID↓	LDM↓
CRoSS	23.79	0.74	0.18	48.85	11.62
SDXL	24.56	0.75	0.31	71.71	10.95
SDXL-Inpt	19.82	0.65	0.33	117.94	17.93
SD-Inpt	26.38	0.78	0.11	30.62	8.49
Ours	28.76	0.82	0.08	21.82	5.99

C. FACE RECOGNITION ANALYSIS

In order to verify the effectiveness and practicability of the proposed method, three face recognition models: DeepFace [56], FaceNet [74], and ArcFace [75] were used to systematically compare and analyze them. Specifically, the labeling accuracy and matching of facial features on the high quality public face dataset CelebA-HQ, as well as on selected recovered image datasets, were carefully evaluated. As shown in Figure 7, a series of test images is presented. These test images consist of different images of the same individual from the CelebA dataset. The distance between each test image, the original image, and the recovered image is computed, and the verification results are also displayed. Through the visual comparison in the figure, it can be observed that the matching results after restoration are consistent with the verification results of the original image, and the distance is also very close. This demonstrates that the recovered image data can effectively replace the original image data for face recognition, and also verifies the high recovery quality of the method from the side.

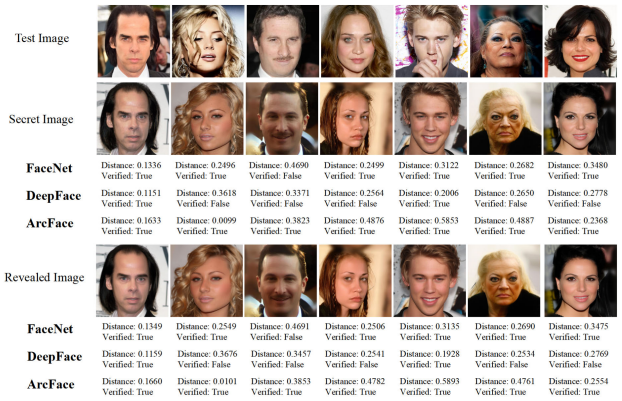

FIGURE 7. FaceNet, DeepFace and ArcFace models detect contrast Distance and Verified results. The recovered image of our method is consistent with the verification results of the original secret image, and the distance difference is small. Our method can effectively replace the original image data for face recognition.

Further, in order to verify the portability of this method in real scenes. Select two commonly used face recognition platforms, Face++ and Aliyun API, as target recognition models. The results of comparative experiments on the Stego240 dataset are shown in Figure 8, showing the confidence scores of face recognition by different methods on the two models. It can be seen from the experimental results that the face recognition rate of the recovered image and the secret image can reach more than 96% on face++, and the highest confidence is obtained. This is further proof that our approach has demonstrated excellent adaptability and superior performance in multiple real-world use environments.

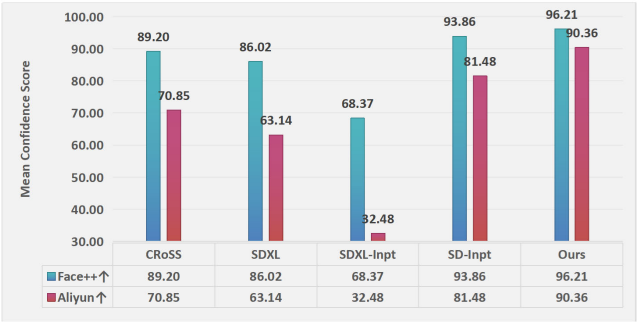

FIGURE 8. Average confidence scores based on Face++ and Aliyun.

D. STEGANALYSIS

To evaluate the security of the steganographic images, both traditional statistical methods and deep learning-based steganalysis techniques were employed to determine whether the images can withstand detection by existing steganalysis

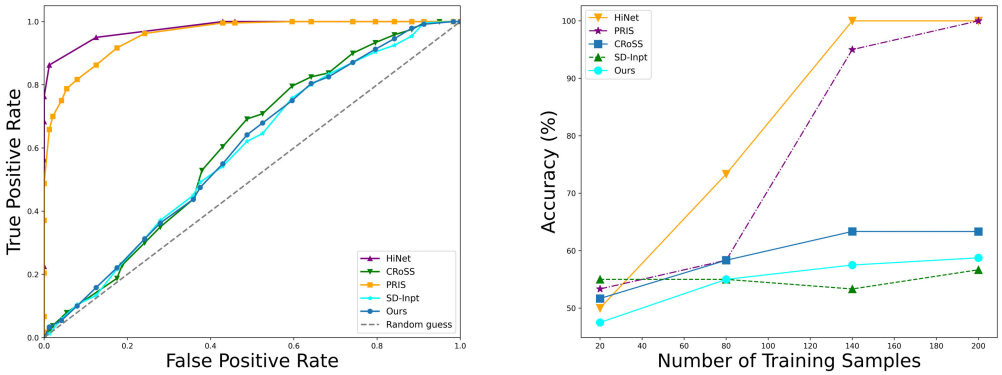

FIGURE 9. The left is the ROC curves generated by different methods under the StegExpose detector. The closer the area under the curve is to 0.5, the better the method is at ideally evading the detector. The right is the results of steganalysis using SRNet. The slower the growth of the curve and the closer the accuracy approaches 50%, the greater the resistance of the method to steganalysis.

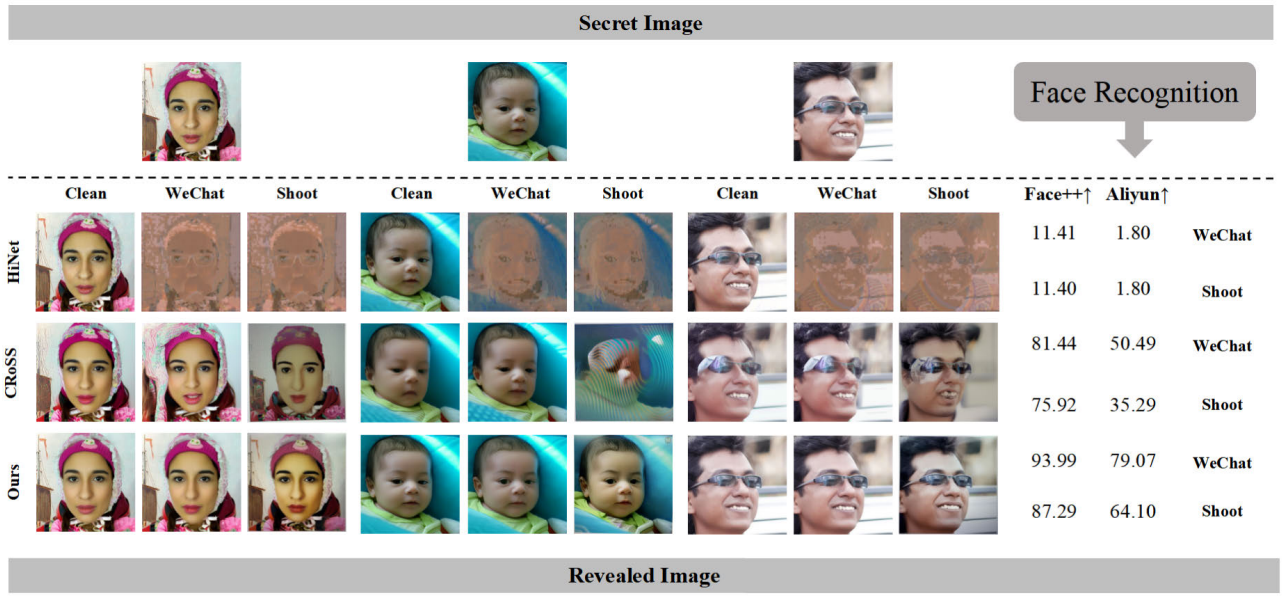

FIGURE 10. In real-world scenarios, when subjected to visual downgrades under conditions labeled "Shoot" and "WeChat," our method effectively reconstructs the contents of a secret image, whereas other methods display significant color distortion or fail entirely. Additionally, our approach maintains high confidence levels on platforms like Face++ and Aliyun.

TABLE 2. Detection accuracy of different methods on SRNet. The best results are highlighted in red and the second-best results are highlighted in blue.

Methods	Accuracy (%)±std
HiNet	77.17±0.251
PRIS	74.33±0.219
CRoSS	57.50±0.059
SD-Inpt	53.50±0.023
Ours	55.25±0.049

tools. As shown on the left side of Figure 9, the open-source steganalysis tool StegExpose [76] was used to test the anti-steganalysis capability of the model. By adjusting different detection thresholds, Receiver Operating Characteristic (ROC) curves were generated. The closer the area under the ROC curve is to 0.5, the closer the detection accuracy is to random guessing, indicating better resistance to steganalysis detection. The results clearly show that the method exhibits low detection accuracy, suggesting that the steganographic images generated by the model possess high security and can effectively deceive the StegExpose tool.

On the right side of Figure 9, the deep learning-based steganalysis tool SRNet [77] was used to test the steganographic images produced by various methods using the StegFace240 dataset. SRNet was retrained by gradually increasing the number of steganographic images used for training. The data in the figure indicates that compared to other methods, the proposed method shows significantly lower detection accuracy, further demonstrating its strong

TABLE 3. Comparison of PSNR (dB) results for our proposed method and other techniques under various levels of degradation. The best results are highlighted in red and the second-best results are highlighted in blue.

Methods	Gaussian noise			JPEG compression		
	$\sigma = 10$	$\sigma = 20$	$\sigma = 30$	QF = 20	QF = 40	QF = 80
HiNet	20.45	13.55	10.22	11.06	11.12	12.75
PRIS	23.83	18.29	14.90	12.86	13.02	15.66
CRoSS	20.78	19.10	17.49	20.73	21.36	22.96
SD-Inpt	24.04	21.40	19.65	24.16	25.09	25.79
Ours	25.96	23.99	22.48	25.49	26.87	28.15

anti-steganalysis capability. Table 2 presents the detection accuracy of different image hiding methods using SRNet. Ideally, the closer the detection accuracy is to 50%, the better the performance of the image hiding algorithm. Our method achieved a detection accuracy of 55.25%, indicating that the steganographic images are almost impossible to accurately detect as containing hidden information.

E. ROBUSTNESS ANALYSIS

To validate the robustness of the method, a series of simulated degradation experiments were conducted, including Gaussian noise and Joint Photographic Experts Group (JPEG) compression. As shown in Table 3, the method demonstrated exceptional adaptability to various levels of degradation, with the least performance decline. Notably, in the presence of Gaussian noise and JPEG compression, the method achieved the highest Peak Signal-to-Noise Ratio (PSNR) values. Even under severe conditions such as Gaussian noise with $\sigma = 30$ and JPEG compression with QF $= 20$, the PSNR values remained above 20dB and 25dB respectively, whereas other methods exhibited a significant drop in fidelity.

To further demonstrate the robustness of the approach, real-world degradation was also tested. In order to simulate the influence of network transmission, experiments were conducted to send and capture container images on the screen via WeChat network. As shown in Figure 10, under this complex degradation condition, all other methods fail entirely or show significant color distortion. In contrast, the method not only successfully reveals the general content of the secret image, but also maintains good semantic consistency with the private key. Once again proving the superiority of the method. In both extreme cases, the approach still achieves the highest confidence of 93.99% (WeChat) and 87.29% (Shoot) on Face++. The proposed method has also succeeded in maintaining higher reconstruction quality compared to the latest methods. These experimental results fully validate the effectiveness and robustness of the method under various experimental and real-world conditions.

V. CONCLUSION

An image steganography framework for face recognition based on diffusion was proposed. This framework combined an innovative mask extraction model, a conditional diffusion model, and deterministic DDIM technology. It made full use of the potential of the diffusion model in privacy protection and realized coverless steganography which was difficult to detect by traditional steganography tools. A large number of experiments showed that this method had obvious advantages in steganography and recovery compared to the existing techniques. The generated steganographic images were diverse and struck a good balance in terms of security, controllability, and robustness. In addition, after recovery, the method was consistent with the verification results of the original face recognition image. The distance could be as low as 0.0004, which was very close to the original image. Meanwhile, the confidence of the method on the Face++ face recognition model reached 96.21%, surpassing the existing methods. This technology not only guaranteed the privacy of users in the face recognition system but also realized the safe, controllable, and robust transmission of data, which was very suitable for the protection of face recognition images.

VI. FUTURE WORK

In the future, facial steganography technology based on diffusion model shows great potential in terms of security, controllability and robustness. However, there are still some limitations: pixel-level objective fidelity (such as PSNR) needs to be improved compared to traditional methods. Good at modifying a single body, but difficult to handle global content. At present, only a single image is supported, and multi-image hiding is expected in the future. We will explore new strategies to improve multi-information hiding capabilities and pixel-level fidelity, provide a more powerful solution for privacy protection in face recognition, promote the application of steganography technology in this field, and provide comprehensive protection for user privacy.

REFERENCES

[1] H. Ullah, M. U. Haq, S. Khattak, G. Z. Khan, and Z. Mahmood, "A robust face recognition method for occluded and low-resolution images," in *Proc. Int. Conf. Appl. Eng. Math. (ICAEM)*, Aug. 2019, pp. 86–91.

[2] W. Junqing, P. Changgen, T. Weijie, and W. Zhenqiang, "FaceEncAuth: Face recognition privacy security scheme based on facenet and smalgorithms," *J. Comput. Eng. Appl.*, vol. 58, no. 11, p. p93, 2022.

[3] R. Pena, F. A. Ferreira, F. Caroli, L. J. S. Silva, and H. Lopes, "Globo face stream: A system for video meta-data generation in an entertainment industry setting," in *Proc. ICEIS*, 2020, pp. 350–358.

[4] C. P. Sumathi, T. Santanam, and G. Umamaheswari, "A study of various steganographic techniques used for information hiding," *Int. J. Comput. Sci. Eng. Surv.*, vol. 4, no. 6, pp. 9–25, Dec. 2013. [Online]. Available: https://ar5iv.org/pdf/1401.5561

[5] Y. Xu, C. Mou, Y. Hu, J. Xie, and J. Zhang, "Robust invertible image steganography," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 7875–7884.

[6] S. Mathur, N. Gupta, and D. Garg, "Secure image steganography with blockchain for copyright protection," *J. Cybersecur. Digit. Forensics*, vol. 11, no. 1, pp. 101–115, 2023.
[7] D. Ariatmanto and F. Ernawan, "Adaptive scaling factors based on the impact of selected DCT coefficients for image watermarking," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 34, no. 3, pp. 605–614, Mar. 2022.
[8] C.-C. Chang and C.-T. Li, "Algebraic secret sharing using privacy homomorphisms for IoT-basedhealthcare systems," *Math. Biosci. Eng.*, vol. 16, no. 5, pp. 3367–3381, 2019.
[9] G.-D. Su, C.-C. Chang, and C.-C. Lin, "High-precision authentication scheme based on matrix encoding for AMBTC-compressed images," *Symmetry*, vol. 11, no. 8, p. 996, Aug. 2019.
[10] I. J. Kadhim, P. Premaratne, P. J. Vial, and B. Halloran, "Comprehensive survey of image steganography: Techniques, evaluations, and trends in future research," *Neurocomputing*, vol. 335, pp. 299–326, Mar. 2019.
[11] H. Olewi, M. Msallam, S. Salim, and H. Al-Behadili, "Enhanced security through integrated Morse code encryption and LSB steganography in digital communications," *Traitement du Signal*, vol. 1, no. 1, pp. 519–524, 2024.
[12] J. Wang, X. Chen, J. Ni, N. Mao, and Y. Shi, "Multiple histograms-based reversible data hiding: Framework and realization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 8, pp. 2313–2328, Aug. 2020.
[13] S. Kamil Khudhair, M. Sahu, K. R. Raghunandan, and A. Sahu, "Secure reversible data hiding using block-wise histogram shifting," *Electronics*, vol. 12, no. 5, p. 1222, Mar. 2023.
[14] G.-D. Su, Y. Liu, and C.-C. Chang, "A square lattice oriented reversible information hiding scheme with reversibility and adaptivity for dual images," *J. Vis. Commun. Image Represent.*, vol. 64, Oct. 2019, Art. no. 102618.
[15] C.-C. Chang, C.-T. Li, and K. Chen, "Privacy-preserving reversible information hiding based on arithmetic of quadratic residues," *IEEE Access*, vol. 7, pp. 54117–54132, 2019.
[16] S. Baluja, "Hiding images in plain sight: Deep steganography," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.
[17] S.-P. Lu, R. Wang, T. Zhong, and P. L. Rosin, "Large-capacity image steganography based on invertible neural networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 10816–10825.
[18] H. Yang, Y. Xu, X. Liu, and X. Ma, "PRIS: Practical robust invertible network for image steganography," *Eng. Appl. Artif. Intell.*, vol. 133, Jul. 2024, Art. no. 108419.
[19] J. Zhu, R. Kaplan, J. Johnson, and L. Fei-Fei, "Hidden: Hiding data with deep networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 657–672.
[20] J. Qin, Y. Luo, X. Xiang, Y. Tan, and H. Huang, "Coverless image steganography: A survey," *IEEE Access*, vol. 7, pp. 171372–171394, 2019.
[21] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2223–2232.
[22] Z. Zhou, H. Sun, R. Harit, X. Chen, and X. Sun, "Coverless image steganography without embedding," in *Proc. 1st Int. Conf. Cloud Comput. Secur. (ICCCS)*, Nanjing, China. Cham, Switzerland: Springer, 2015, pp. 123–132.
[23] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 6840–6851.
[24] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," 2020, *arXiv:2011.13456*.
[25] A. Lugmayr, M. Danelljan, A. Romero, F. Yu, R. Timofte, and L. Van Gool, "RePaint: Inpainting using denoising diffusion probabilistic models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11461–11471.
[26] B. Kawar, M. Elad, S. Ermon, and J. Song, "Denoising diffusion restoration models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 23593–23606.
[27] Y. Wang, J. Yu, and J. Zhang, "Zero-shot image restoration using denoising diffusion null-space model," 2022, *arXiv:2212.00490*.
[28] P. Dhariwal and A. Nichol, "Diffusion models beat GANs on image synthesis," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 34, 2021, pp. 8780–8794.
[29] C. Mou, X. Wang, L. Xie, Y. Wu, J. Zhang, Z. Qi, and Y. Shan, "T2I-adapter: Learning adapters to dig out more controllable ability for text-to-image diffusion models," in *Proc. AAAI Conf. Artif. Intell.*, 2024, vol. 38, no. 5, pp. 4296–4304.
[30] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 10684–10695.
[31] Z. Wang, Z. Zhang, X. Zhang, H. Zheng, M. Zhou, Y. Zhang, and Y. Wang, "DR2: Diffusion-based robust degradation remover for blind face restoration," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 1704–1713.
[32] A. Hertz, R. Mokady, J. Tenenbaum, K. Aberman, Y. Pritch, and D. Cohen-Or, "Prompt-to-prompt image editing with cross attention control," 2022, *arXiv:2208.01626*.
[33] R. Mokady, A. Hertz, K. Aberman, Y. Pritch, and D. Cohen-Or, "Null-text inversion for editing real images using guided diffusion models," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 6038–6047.
[34] G. Kim, T. Kwon, and J. C. Ye, "DiffusionCLIP: Text-guided diffusion models for robust image manipulation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 2426–2435.
[35] C. Meng, Y. He, Y. Song, J. Song, J. Wu, J.-Y. Zhu, and S. Ermon, "SDEdit: Guided image synthesis and editing with stochastic differential equations," 2021, *arXiv:2108.01073*.
[36] J. Song, C. Meng, and S. Ermon, "Denoising diffusion implicit models," 2020, *arXiv:2010.02502*.
[37] C.-T. Lin and D.-R. Duh, "Advanced PVD-based steganography for secure communication," *J. Netw. Comput. Appl.*, vol. 200, Jan. 2023, Art. no. 102955.
[38] S. Kang, H. Park, and J. Park, "High capacity secure dynamic multi-bit data hiding using Fibonacci Energetic pixels," *Multimedia Tools Appl.*, vol. 83, pp. 5181–5206, May 2023.
[39] S. Imaizumi and K. Ozawa, "Multibit embedding algorithm for steganography of palette-based images," in *Proc. 6th Pacific-Rim Symp. Image Video Technol. (PSIVT)*, Guanajuato, Mexico. Cham, Switzerland: Springer, 2014, pp. 99–110.
[40] M. Alotaibi, D. Al-hendi, B. Alroithy, M. AlGhamdi, and A. Gutub, "Secure mobile computing authentication utilizing hash, cryptography and steganography combination," *J. Inf. Secur. Cybercrimes Res.*, vol. 2, no. 1, pp. 9–20, 2019.
[41] A. A. Zakaria, M. Hussain, A. W. A. Wahab, M. Y. I. Idris, N. A. Abdullah, and K.-H. Jung, "High-capacity image steganography with minimum modified bits based on data mapping and LSB substitution," *Appl. Sci.*, vol. 8, no. 11, p. 2199, Nov. 2018.
[42] N. Al-Juaid, A. Gutub, and E. Khan, "Efficient implementation of multi-image secret hiding based on LSB and DWT steganography comparisons," *Arabian J. Sci. Eng.*, vol. 43, no. 2, pp. 911–920, 2018.
[43] H. Shi, J. Dong, W. Wang, Y. Qian, and X. Zhang, "SSGAN: Secure steganography based on generative adversarial networks," in *Proc. 18th Pacific-Rim Conf. Multimedia Adv. Multimedia Inf. Process. (PCM)*, Harbin, China. Cham, Switzerland: Springer, 2018, pp. 534–544.
[44] K. Alex Zhang, A. Cuesta-Infante, L. Xu, and K. Veeramachaneni, "SteganoGAN: High capacity image steganography with GANs," 2019, *arXiv:1901.03892*.
[45] C. Zhang, P. Benz, A. Karjauv, G. Sun, and I. S. Kweon, "UDH: Universal deep hiding for steganography, watermarking, and light field messaging," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 33, 2020, pp. 10223–10234.
[46] Y. Peng, D. Hu, Y. Wang, K. Chen, G. Pei, and W. Zhang, "StegaDDPM: Generative image steganography based on denoising diffusion probabilistic model," in *Proc. 31st ACM Int. Conf. Multimedia*, Oct. 2023, pp. 7143–7151.
[47] L. Dinh, D. Krueger, and Y. Bengio, "NICE: Non-linear independent components estimation," 2014, *arXiv:1410.8516*.
[48] J. Jing, X. Deng, M. Xu, J. Wang, and Z. Guan, "HiNet: Deep image hiding by invertible network," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 4733–4742.
[49] X. Jia, H. Xin, L. Gu, H. Wang, J. Sun, and W. Wan, "AFcIHNet: Attention feature-constrained network for single image information hiding," *Eng. Appl. Artif. Intell.*, vol. 126, Nov. 2023, Art. no. 107105.
[50] Z. Zhou, Y. Mu, and Q. M. J. Wu, "Coverless image steganography using partial-duplicate image retrieval," *Soft Comput.*, vol. 23, no. 13, pp. 4927–4938, Jul. 2019.
[51] Y. Mu and Z. Zhou, "Visual vocabulary tree-based partial-duplicate image retrieval for coverless image steganography," *Int. J. High Perform. Comput. Netw.*, vol. 14, no. 3, pp. 333–341, 2019.

[52] Q. Liu, X. Xiang, J. Qin, Y. Tan, J. Tan, and Y. Luo, "Coverless steganography based on image retrieval of DenseNet features and DWT sequence mapping," *Knowl.-Based Syst.*, vol. 192, Mar. 2020, Art. no. 105375.
[53] Y. Luo, J. Qin, X. Xiang, Y. Tan, Q. Liu, and L. Xiang, "Coverless real-time image information hiding based on image block matching and dense convolutional network," *J. Real-Time Image Process.*, vol. 17, no. 1, pp. 125–135, Feb. 2020.
[54] J. Lu, J. Ni, L. Li, T. Luo, and C. Chang, "A coverless information hiding method based on constructing a complete grouped basis with unsupervised learning," *J. Netw. Intell.*, vol. 6, no. 1, pp. 29–39, 2021.
[55] J. Yu, X. Zhang, Y. Xu, and J. Zhang, "CRoSS: Diffusion model makes controllable, robust and secure image steganography," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 36, 2024, pp. 1–14.
[56] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 1701–1708.
[57] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[58] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141.
[59] S. Chen, Y. Liu, X. Gao, and Z. Han, "MobileFaceNets: Efficient CNNs for accurate real-time face verification on mobile devices," in *Proc. 13th Chin. Conf. Biometric Recognit. (CCBR)*, Urumqi, China. Cham, Switzerland: Springer, 2018, pp. 428–438.
[60] Y.-H. Li, C.-C. Chang, G.-D. Su, K.-L. Yang, M. S. Aslam, and Y. Liu, "Coverless image steganography using morphed face recognition based on convolutional neural network," *EURASIP J. Wireless Commun. Netw.*, vol. 2022, no. 1, p. 28, Dec. 2022.
[61] E. Abusham, B. Ibrahim, K. Zia, and M. Rehman, "Facial image encryption for secure face recognition system," *Electronics*, vol. 12, no. 3, p. 774, Feb. 2023.
[62] D. Liu, X. Wang, C. Peng, N. Wang, R. Hu, and X. Gao, "Adv-diffusion: Imperceptible adversarial face identity attack via latent diffusion model," in *Proc. AAAI Conf. Artif. Intell.*, 2024, vol. 38, no. 4, pp. 3585–3593.
[63] J. Sohl-Dickstein, E. Weiss, N. Maheswaranathan, and S. Ganguli, "Deep unsupervised learning using nonequilibrium thermodynamics," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 2256–2265.
[64] R. Gal, Y. Alaluf, Y. Atzmon, O. Patashnik, A. H. Bermano, G. Chechik, and D. Cohen-Or, "An image is worth one word: Personalizing text-to-image generation using textual inversion," 2022, *arXiv:2208.01618*.
[65] O. Avrahami, D. Lischinski, and O. Fried, "Blended diffusion for text-driven editing of natural images," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 18208–18218.
[66] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, and K. Aberman, "DreamBooth: Fine tuning text-to-image diffusion models for subject-driven generation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 22500–22510.
[67] T. Brooks, A. Holynski, and A. A. Efros, "InstructPix2Pix: Learning to follow image editing instructions," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 18392–18402.
[68] L. Huang, D. Chen, Y. Liu, Y. Shen, D. Zhao, and J. Zhou, "Composer: Creative and controllable image synthesis with composable conditions," 2023, *arXiv:2302.09778*.
[69] L. Luo, D. Xue, and X. Feng, "EHANet: An effective hierarchical aggregation network for face parsing," *Appl. Sci.*, vol. 10, no. 9, p. 3135, Apr. 2020.
[70] P. Krähenbühl, K. He, Y. Xu, and J. Liu, "High-resolution image synthesis and semantic manipulation with conditional GANs," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Jun. 2019, pp. 8798–8807.
[71] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4396–4405.
[72] J. Li, D. Li, C. Xiong, and S. Hoi, "BLIP: Bootstrapping language-image pre-training for unified vision-language understanding and generation," in *Proc. Int. Conf. Mach. Learn.*, 2022, pp. 12888–12900.
[73] D. Podell, Z. English, K. Lacey, A. Blattmann, T. Dockhorn, J. Müller, J. Penna, and R. Rombach, "SDXL: Improving latent diffusion models for high-resolution image synthesis," 2023, *arXiv:2307.01952*.
[74] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 815–823.
[75] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 4690–4699.
[76] B. Boehm, "StegExpose—A tool for detecting LSB steganography," 2014, *arXiv:1410.6656*.
[77] I. Corley, J. Lwowski, and J. Hoffman, "Destruction of image steganography using generative adversarial networks," 2019, *arXiv:1912.10070*.

YUAN GUO received the B.S. degree in automation from Qiqihar University, Qiqihar, China, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from Yanshan University, Qinhuangdao, China, in 2004 and 2008, respectively. She was a Visiting Scholar with Johns Hopkins University, Baltimore, MD, USA, from 2012 to 2013. She is currently a Professor of computer science and technology with Heilongjiang University. Her current research interests include steganography, optical image encryption, sensor technology, and image processing.

ZIQI LIU received the B.S. degree in software engineering from Tianjin Normal University, China, in 2022. She is currently pursuing the master's degree with Heilongjiang University, Harbin, China. Her current research interests include image encryption, information hiding, and image processing.