Hindawi
Computational Intelligence and Neuroscience
Volume 2023, Article ID 6758406, 11 pages
https://doi.org/10.1155/2023/6758406

Research Article

A Privacy Protection Framework for Medical Image Security without Key Dependency Based on Visual Cryptography and Trusted Computing

Denghui Zhang,[1,2] Lijing Ren,[2] Muhammad Shafiq ,[1] and Zhaoquan Gu[2,3]

[1]*Cyberspace Institute of Advanced Technology, Guangzhou University, Guangzhou 510000, China*
[2]*Department of New Networks, Peng Cheng Laboratory, Shenzhen 518055, China*
[3]*School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China*

Correspondence should be addressed to Muhammad Shafiq; srsshafiq@gmail.com

Received 9 August 2022; Revised 30 September 2022; Accepted 24 November 2022; Published 31 January 2023

Academic Editor: Inam Ullah

The development of mobile Internet and the popularization of intelligent sensor devices greatly facilitate the generation and transmission of massive multimedia data including medical images and pathological models on the open network. The popularity of artificial intelligence (AI) technologies has greatly improved the efficiency of medical image recognition and diagnosis. However, it also poses new challenges to the security and privacy of medical data. The leakage of medical images related to users' privacy is emerging one after another. The existing privacy protection methods based on cryptography or watermarking often bring a burden to image transmission. In this paper, we propose a privacy-preserving recognition network for medical images (called MPVCNet) to solve these problems. MPVCNet uses visual cryptography (VC) to transmit images by sharing. Benefiting from the secret-sharing characteristics of VC, MPVCNet can securely transmit images in clear text, which can both protect privacy and mitigate performance loss. Aiming at the problem that VC is easy to forge, we combine trusted computing environments (TEE) and blind watermarking technologies to embed verification information into sharing images. We further leverage the transfer learning technology to abate the side effect resulting from the use of visual cryptography. The results of the experiment show that our approach can maintain the trustworthiness and recognition performance of the recognition networks while protecting the privacy of medical images.

1. Introduction

Recently, smart medicine has become an attractive field of applications with the development of 5G, IoT, and AI [1]. Telemedicine has come a long way in detecting and diagnosing diseases remotely, which means that medical images are transmitted more frequently over the open network. The gradual combination of modern medicine with computer technology, communication technology, and multimedia technology has provided patients and doctors with fast medical diagnoses. These technologies have greatly improved the accuracy of medical diagnosis and relieved patients who enjoy the convenience of digital medicine. The storage and analysis of medical images are gradually moving to the cloud, which is a prerequisite for efficient cooperation in remote diagnosis and resource sharing [2].

Medical imaging devices facilitate doctors' diagnosis and treatment. Benefiting from advances such as 5G in wireless communications and IoT in the industrial Internet, it is now easier to capture and transfer images in the medical field [3]. However, these novel technologies pose new challenges to patients' privacy [4]. Medical images are usually characterized by huge data volume and high security, which contain a large amount of personal information when they are not desensitized. For example, German security firm Greenbone Networks discovered 24 million leaked medical images in 2020, and two months later, the number of exposed scans exceeded 1.19 billion. Frequent medical data

leakage incidents have seriously violated users' privacy and endangered social security. Therefore, privacy protection has to become a requisite for e-healthcare systems [5]. Although different types of data protection algorithms have been proposed, most of them are aimed at text or digital media, which is different from medical images in information quantity and scale [6]. It is often unnecessary to encrypt all pixels to securely transmit an image [7].

Traditional encryption methods based on public keys not only need complex computation but also need additional key management [8]. A common obstacle to these methods is that once the key is lost, we cannot restore the secret of encrypted information [9]. Edge image capture devices generally have limited computing power, which cannot meet the demand for real-time encryption of large volumes of images. These shortcomings limit the spread of AI technology in healthcare. Strengthening the security and protection of privacy of medical images without performance degeneration is urgent. Big data is the key to the success of medical AI. The proposal of a lightweight image protection scheme is the main challenge in this article [10].

Visual cryptography (VC) [11, 12], which is a secret-sharing technology aimed at images, can be applied to digital devices with limited computing in an untrusted networking environment. Using the threshold and secret-sharing features of VC, we can use simple Boolean operations to achieve real-time encryption of large amounts of image data while removing the dependence on keys. However, the VC-generated share is not easy to manage and is vulnerable to attackers.

To preserve the privacy of data when sharing medical data, we present the MPVCNet (medical privacy-oriented VC-based recognition network) to address illegal access and identity forgery for sharing of medical images. We first construct a verifiable and expansion-free visual cryptography scheme to migrate easily forged and size-expanded shares by combining the VC and TEE technologies, where TEE is used to ensure that remote sharing operation is not tampered with [13]. Then, we can securely store and transmit medical images distributed with the proposed scheme. Since the separated sheets do not reveal any feature of the original biomedical data, we can transmit sensitive images among public networks in plain view. MVPCNet eliminates complex computing operations and key management workload in the traditional public key or watermark protection method through the software and hardware cooperation scheme.

2. Related Work

VC is a secret-sharing mechanism for images. Since it was first proposed in 1995, it has become emerging research in the field of image security [11, 14, 15]. Unlike the text-oriented secret-sharing mechanism, which divides the sum into two addends, VC splits secret pixels into multiple subpixels. All black pixels will remain once overlapping sharing blocks, although degrading white secret pixels, the difference of contrast between black and white pixels makes human eyes still recognize the features of the restored image [16]. Because a single pixel is encrypted into a larger color block, the size of the encrypted and decrypted image will expand. The noise-like shares are also often exploited by malicious users.

Because of the low-pass filtering characteristic of human eyes, an image can be recognized even if pixels in the local area change. The probabilistic VCS uses a pixel instead of a color block to encrypt the secret pixel with this feature, thus keeping the size of the images consistent [17, 18]. Although the wrong pixel may be selected at one time, the probability ensures that when many pixels are gathered, the displayed image has the same probability density as the original image; that is, the two images are similar. If we encrypt one color block at a time instead of one pixel, we can also keep the size of the decrypted image inconvenient [19]. By formulating the correspondence between secret blocks and color blocks before and after encryption, we can decrypt higher-quality secret images [10]. However, this mechanism often leads to more shares, which requires more storage space. The software-based encryption scheme is hard to avoid the attack on the hypervisor and remote operating system.

Since images generated by VC are meaningless, the attacker can forge shares without being detected. To address this limitation, researchers have proposed many CIVCS (cheating immune VCS) [20–22]. If there is no checksum information, malicious users can forge shares and deceive users. So, CIVCS often requires additional pixels or more shares to embed verification information, which can burden the cost of VC.

Big data and cloud computing technologies have solved two main challenges in the spread of medical AI [23, 24]. A growing number of smart medical systems facilitate people's daily life [25], which has become an important research direction in healthcare. The authors [26] use deep learning algorithms to identify diabetic retinopathy lesions, which are the most common sequelae of diabetes and can lead to blindness. The experimental results show that the AI algorithm outperforms medical experts and can extract lesion features from fine textures. Due to the difficulty and scarcity in the acquisition of medical image annotation data, transfer learning [27] has become a very common technique for medical image recognition. TransFusion [25] investigates the evolution of representations of different models and hidden layers during training and the advantage of feature independence of migration learning to accelerate convergence.

3. MPVCNet

With the introduction of data security regulations and increasing awareness of the limitations of AI technology, it is imperative to ensure the privacy of personal information when enjoying the convenience of AI technology [4]. To address these limitations, we will first present a size-invariant and verifiable VCS and then propose a secure and trustworthy image recognition network in this section.

3.1. The List of Abbreviations. Table 1 provides the list of abbreviations used in this article.

Table 1: The list of abbreviations used in this article.

Abbreviations	Terms
MPVCNet	Medical privacy-oriented VC-based recognition network
VC	Visual cryptography
IoT	Internet of things
VCS	VC scheme
CIVCS	Cheating immune VCS
SGX	Software guard extensions
SSIM	Structure SIMilarity
PSNR	Peak signal-to-noise ratio
ROC	Receiver operating curve

3.2. Backgrounds

3.2.1. Visual Cryptography Schemes (VCSs). The secret-sharing scheme, which is also called a threshold scheme, was originally used to provide a solution when many participants needed to share a secret. Only the number of participants who reach a threshold can recover sensitive information such as pathological and CT images. Participants less than the threshold cannot reveal any information about the original image even if they conspire.

As an extension of secret-sharing technology toward image encryption, the principle of VC is to divide an image into n unrelated shares (also called sheets), which are independent of each other and distributed to n individuals for safekeeping. If $m, m \le n$, image cannot reconstruct the original image, thus achieving image security. Image encryption schemes based on secret-sharing have the advantages of high security and simple computation.

VC essentially makes use of the contrast characteristics of human eyes to color. For white pixels, the subpixels of shares 1 and 2 are the same, while for black pixels, the subpixels of shares 1 and 2 are complementary. When superimposing, all-black decrypted pixels can be obtained, while the white decrypted pixel contains two black pixels, but its brightness is still higher than that of the black decrypted pixel.

Equations (1) and (2) illustrate the encryption and decryption process of a (2, 2)-VCS. When splitting a secret image into two meaningless shares, the original image can only be restored if two shares are obtained simultaneously. The $C_k, k \in \{0, 1\}$ is the encryption matrix for a black pixel (0) and white pixel (1), which can be generated by rotating the base matrix S_k. The top and bottom rows of the black matrix (C_1) can be superimposed to obtain 4 black pixels, and the white matrix (C_1) can be superimposed to obtain 2 black pixels. So, the decrypted image looks darker. Therefore, we can visually experience a significant difference at grayscale levels when sharing images using the overlay.

$$S^0 = \left[\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right], S^1 = \left[\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}\right], \tag{1}$$

$$C_0 = \left\{\text{permution of } S^0\right\}, C_1 = \left\{\text{permution of } S^1\right\}. \tag{2}$$

3.2.2. Intel SGX. Intel Software Guard Extensions (SGX) is one of the most popular TEE technologies [28]. It is an instruction extension proposed by Intel to provide hardware-level protection for user code and data based on the Intel CPU architecture. SGX is a specific implementation of TEE to prevent other applications, including privileged operating systems, from tampering with the user code and sniffing information about applications running in a protected environment (called enclave). Neither privileged nor unprivileged software can access the enclave. SGX technology consists of two core mechanisms: isolated execution and remote authentication.

(1) Isolated Execution. An enclave can protect the confidentiality and integrity of the processes running in it, which is defended from attacks by other processes on the host, and blocks hardware-oriented attacks. As shown in Figure 1, SGX uses an inverted sandbox model to protect applications and data from being skewed. Instead of identifying and isolating all malwares on the platform, SGX protects legitimate software by encapsulating its security operations in an enclave and protecting it from malware attacks.

SGX provides only a limited amount of memory (128 MB), and excluding the driver, the space available to the user enclave is even less, equal to about 93 MB [29]. Therefore, at runtime, SGX will load data into the enclave from the CPU to memory as needed or export to memory in ciphertext [30]. The key is stored in hardware and is not accessible even by the operating system. The malicious program can also only sniff the encrypted ciphertext data from memory. When the user program needs to access the previous data, the data in memory are paged back to enclave and the decryption operation is performed. The whole process is transparent to the user. Because the whole encryption and decryption process is performed in hardware, the impact on program performance is small.

(2) Remote Authentication (RA). Intel SGX provides a remote authentication mechanism for applications running on remote or cloud platforms. Using the remote authentication mechanism, before the client encrypts and transmits sensitive data to the application running in an enclave, the client will verify the credibility of the remote environment and application. The enclave can indicate to the challenger that its identity has not been altered while running in a trusted Intel SGX environment [31].

In a complete authentication process, a server runs in a trusted SGX environment. At first, the client challenges the server to prove its identity. If an application wants to transmit data to an enclave or receive the execution result of the enclave in confidence, it requires the establishment of a secure transmission channel between an external application and the enclave. The principle is like TLS (transport layer security) of HTTPS in a browser. First, the client and server

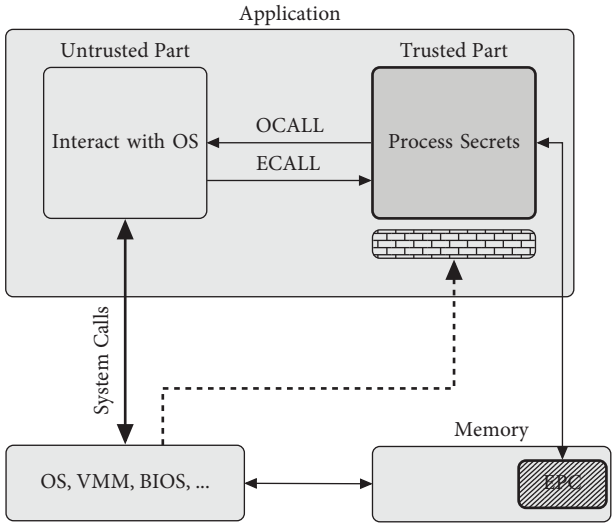

Figure 1: The workflow of Intel SGX.

use the key exchange algorithm such as the Diffie–Hellmann key exchange to negotiate the key of the asymmetric encryption algorithm, and then both communication parties use this key to encrypt messages for communication. After passing the Intel attestation service (IAS) verification, the client provides sensitive data to the server in digital copyright protection or approves the operation performed by the server by using the encryption channel established in the authentication process [32]. Please refer to documents [33] for more details about Intel SGX.

3.2.3. Verifiable Visual Cryptography. We chose the random grid (RG) method [15] as the backbone scheme; that is, C_0 and C_1 are used as the underlying base matrix. To distribute each column of collections to a participant, we shuffle the subpixel in a column and distribute one of them to a participant instead of the whole column. Although there are more complex VCS [34], which often require more shares and can reduce the quality of the restored image, the advantage of VCS is that it distributes subpixels to corresponding participants; that is, the number of subpixels received by participants is the same as that of encrypted secret pixels, so VCS can keep the image size. While the origin VCS distributes subpixels to one of the participants, it substantially expands the size of the shared image. The encryption is shown in Algorithm 1, which is explained in more detail as follows:

(1) Decomposition. Like other color images, medical images often have three (red, green, and blue) gray-level channels. The RGB mode is additive color, which is based on the superposition of light. The red light plus green light and blue light are equal to the white light in nature. This color mode is suited to appropriate devices like displays. While printers utilize ink to absorb light to display a specific color, this color mode is called the subtractive mode, which also has three (cyan, magenta, and yellow) channels. The cyan pigment blends magenta and yellow pigments to obtain black (gray) in printing. The digital color images are stored in the RGB mode. We first need to decompose the original image into three CMY gray-level channel images. This can be archived from RGB channels with the following transformation to get complementary colors: $C = 255 \setminus R, M = 255 \setminus G, Y = 255 \setminus B$.

(2) *Digital Half-Toning.* The original VCS only can encrypt white and black pixels. While the gray-level channel images have 256 levels (as a reference, the binary image has only two gray levels). Halftone is a commonly-used digital image processing technology, which simulates continuous tone images by differing the frequency of a small amount of color, and the visual effect of the quantized image is like the original image at a certain distance. Halftone technology is based on the low-pass filtering feature of human eyes and color characteristics to achieve optimal image reproduction in monochrome or binary color equipment. When viewing spatially close parts of an image, eyes will form the effect of continuous tone in general. Error-diffusion [35] is a commonly used halftone algorithm that can be described as follows:

$$\begin{cases} p_{x-1,y+1} = p_{x-1,y+1} + qe \times c[1,0], \\ p_{x+1,y} = p_{x+1,y} + qe \times c[0,2], \\ p_{x+1,y} = p_{x,y+1} + qe \times c[1,1], \\ p_{x+1,y} = p_{x+1,y+1} + qe \times c[1,2], \end{cases} \tag{3}$$

$$qe = p_{x,y} - T_{x,y}, \tag{4}$$

$$\begin{cases} T_{x,y} = 0, & \text{if } p_{x,y} < \tau, \\ T_{x,y} = 1, & \text{if } p_{x,y} \geq \tau, \end{cases}$$
$$c = \begin{bmatrix} p_{\text{old}} & p & \frac{7}{16} \\ \frac{3}{16} & \frac{5}{16} & \frac{1}{16} \end{bmatrix}, \tag{5}$$

where qe is the optimal quantization error and it usually is a threshold function truncated by τ. The coefficients c are optimal arrays to make diffusion smoother, which are [7/16, 3/16, 5/16, 1/16]. Considering the gray value of the neighborhood pixels, the output image is smoother and the noise is less than that of the random or jitter method.

(3) Encryption. The encryption process is applied to each halftone channel image separately. When processing a pixel p_{ij}, we will set the corresponding pixel S_{1ij} in *share1* to be a white or black pixel randomly. If p_{ij} is a white pixel, we set the corresponding pixel S_{2ij} in *share2* to be the same as S_{1ij}. While if p_{ij} is a black pixel, we will set S_{2ij} to be the inverse of S_{1ij}. For example, for the black, we will reinterpret S_{2ij} as a white pixel if S_{1ij} is also a black pixel. While if S_{1ij} is a black pixel, S_{2ij} will be a white pixel. After generating and sharing images S_1 and S_2, we will generate a corresponding verification message. First, a random number R is generated so that the receiver can use R to perform pair operations when receiving multiple shares. The user can also use a GUID

Data: A secret image S
Result: Share images S_1, S_2 dispatched to participants P_1, P_2
$S_h \leftarrow$ the half-toned secret image got with the error-diffusion algorithm from equation (1)
for each pixel p_{ij} in S_h **do**
  if p is black **then**
    $d_1, d_2, \cdots, d_n \leftarrow n$ randomly generated black or white pixels
    $S_{nij} = d_n, n$ is the number of shares
  end
  else
    $d \leftarrow$ a randomly generated black or white pixel
    $S_{nij} = d$
  end
end
$R \leftarrow$ a random number
for each share S_i **do**
  $A \leftarrow S_i \parallel R$
  $Q \leftarrow$ a generated quote that signs the hash value of A
  $AR \longleftarrow Q \parallel R$
for each bit b_j in AR **do**
    $p_j \leftarrow$ the corresponding pixel in S_i
    $\mathrm{LSB}(p_j) \leftarrow b_j$
  end
end

Algorithm 1: A verifiable visual cryptography scheme.

(globally unique identifier) to indicate pairing if it can be uniquely identified in this system. The enclave (E), which is the trusted part of the encryption program, and q*uote* enclave (Q) in the host will first perform local authentication. E asks Q to sign the hash value of the spliced bitstream $A = S_i \| R$, and we embed A into the quote structure. Since the hash value is fixed, the length of the final generated quote structure is also fixed. Finally, the authentication code $AR = Q \| R$ is generated, and each bit in AR is assigned to LSB (lowest significant bit) for each pixel, in turn, starting from the first pixel of the corresponding share. For example, if $AR[i] = 1$ and $S_1[i]$ are white pixels, then we will set $\mathrm{LSB}(S_1[i])$ to 1, the grayscale of the shared pixel $S_1[i]$ will be reassigned to 255; that is, it is still a white pixel. While if $AR[i] = 0$, $S_1[i]$ will be reassigned to 254 which is received by setting the lowest bit of 255 to 0. When the pixel in S_1 is black, the grayscale value of the corresponding pixel becomes 0 or 1. So, the range of pixels in shares after embedding an authentication code value is $[0, 255, 1, \text{and } 254]$. The embedding process of the authentication information into the share S_2 is similar.

(4) Check and Decryption. When restoring the original image, we can ask participants P_1 and P_2 to show their sheets. Before overlaying shares to recover the secret image, we first need to extract the verification information from the received shares. Since the *quote* structure and the paired random number (R) are both of fixed length, the embedding information is hidden in the first n (the sum of the byte lengths of Q and R) bits of the received shares. The extraction of verification information is an inverse process of the previous embedding process. Taking S_1, for example, if the ith bit of S_1 is 1 or 255, the ith bit will be 1 in the corresponding verification information (AR). While if S_1 is 0 or 254, we will assign $AR[i]$ to 0. Once the extraction process is completed, we will get the quote and pairing information and recover the pixels with a grayscale of 1 or 254 in shares to 0 and 255, respectively. We take received images as trusted shares only if they satisfy the following conditions: (i) the pairing information of the shares is the same; (ii) the validity of Q is requested to be verified by IAS, and the validity of the IAS report is verified by the local certificate. If the report is valid and the signature information of this report is consistent with the locally stored IAS certificate information and the *MRENCLAVE* ID extracted from Q is in the list of trusted enclave IDs, then we think Q is trusted; (iii) the hash value of the joined bitstream of Q and R is consistent with the embedded hash information in Q. After passing all of the previous checks, we can print out the two shares and superimpose them together or simply perform a Boolean *AND* operation to reveal secret images. Note that we perform Boolean *AND* while the Boolean *OR* operation is executed in the original scheme. This is because the original scheme uses 1 for black pixels and 0 for white pixels, whereas white pixels are generally represented by 1 or 255 and black pixels by 0 in normal image files.

3.2.4. The Recognition Network of Medical Images. In the preceding section, VC provides a safe and effective encryption scheme for the transmission and analysis of large-scale medical image data sets, which eliminates the key dependence on traditional encryption methods. Although the application of VC can efficiently transmit medical images, the quality of restored images may be lossy. This may break recognition performance [36]. In this section, we will

use the strong feature extraction ability of deep learning to propose a high-accuracy recognition network of noisy or blurred medical images.

Traditional image recognition methods obtain recognition models through elaborate feature extraction methods, which are limited by environments such as lighting, contrast, and pose, although human eyes can accurately recognize friends around us, regardless of the external conditions. With the construction of large-scale datasets and increasing computing power, deep learning-based image recognition methods are now capable of recognizing complex scenes with high performance even beyond human eyes [37].

Deep learning-based models can be automatically trained with datasets to extract complex mapping relationships between input images and target labels. However, in the process of image recognition, to maximize the recognition performance of neural networks, it is generally necessary to first design the network structure and loss function according to the features of datasets and then continuously train on the cloud platform until a stable network model is obtained [38, 39]. This process is often time-consuming, and recognition performance is still influenced by the selection of hyperparameters in the training process.

With the further development of machine learning methods, we can now directly reuse the model parameters learned from large-scale datasets. Note that our lossy medical image dataset does not have the same feature distribution as generic datasets such as ImageNet, and it is still not feasible to reuse existing models and weights directly. To issue a problem, we can add a fully connected layer at the end of the network and then fine-tune the model to match the size and class of the medical image data. Then, we can use the method in the previous section to encrypt and decrypt the original training dataset one by one to obtain a recognition model for the lossy dataset. Since this approach considers both the features and weights learned from a large-scale generic dataset and the characteristics of the model itself, therefore we can efficiently obtain a high-accuracy recognition network of noisy or blurred medical images. We only need to add a new classifier to the pretraining model, which will be trained from scratch [40].

Previous experiments [25] show that feature reuse mainly occurs at the lowest level. We conducted a weighted transfusion operation to speed up the training process; that is, only a part of the pretraining weights (corresponding to a group of consecutive layers) was transferred, and the rest weights were randomly initialized. Compared with that of complete transfer learning, the convergence speed of the indoctrination network will further accelerate.

4. Results and Discussion

To assess the effectiveness and security of the method proposed in this paper, we will test the performance of the proposed framework in signing encryption and images on the prototype system developed using the C++ language based on the SGX SDK. We evaluate the latency details of the encrypted and signed multiple images by using a modified SGX-SSL benchmarking tool. We performed experiments on an Intel desktop CPU with 16 GB RAM running on the Linux system. We both test the performance of the verifiable VCS on a normal host and a host that supports SGX features.

Figure 2 shows that the normal operation had a higher performance where no security enhancements are adopted. There are two reasons why the performance of signing in TEE is weaker than the plain signing: (i) one is the loss due to the use of encryption memory in TEE, which is inevitable in order to prevent the privileged system from sniffing; (ii) the other is the extra function calls resulting from the refactoring of the original code to allow the original code to run in TEE. However, the loss from direct function calls is still less than the loss from *ECALL/OCALL* in the enclave running on a host that allows the SGX feature.

Figure 3 shows the latency between the TEE-based approach presented in this paper and other commonly used image encryption methods. Although SGX has a higher performance payload than conventional transmission, the loss is lower than the encryption methods, including RSA and TLS. The more complex the encryption process, the higher the efficiency of encryption, which further shows the importance of adopting visual cryptography and the TEE software and hardware coprotection method. The performance difference can be attributed to two reasons. (i) SGX is an extended instruction set integrated with the CPU that can call hardware instructions to speed up the encryption and decryption process. (ii) SGX distinguishes between trusted and untrusted codes; not only is the normal operation split into two sets of code, but the parameters must be toggled and copied back and forth between the two running spaces, which enhances security at an extra cost. So, encryption, decryption, and signing operations with SGX are more time-consuming than in a normal environment, but the average loss of 8× is within reasonable bounds.

We further develop an application using the TensorFlow [41] framework to test the recognition network model rapidly. We first select the diabetic retinopathy (DR) detection dataset to identify signs of diabetic retinopathy in eye images, which consists of 35126 DR images with the size 786×512. The DR is the first cause of blinding. We choose another *BreakHis* [42] (breast cancer histopathological) dataset which contains 7909 breast histopathological images from 82 patients to further evaluate the method proposed. *BreakHis* not only advances the research of binary classification of benign and malignant but also advances the research of pathological classification (multiclassification), which is very significant in the clinic. Since the size of the images in the dataset varies, the shape of the input features of the neural network needs to be consistent. Therefore, to simplify the experiment, we used the smallest image in the dataset as the benchmark and then used data enhancement to obtain the preprocessed dataset with $S = \{1, 2, 4, 5\}$; of these, 70% were used as a training set, and 30% as a test set shows the data distribution of the DP dataset for different levels. The third and fourth columns of Table 2 are the average *PSNR* (peak signal-to-noise ratio) from half-toned and recovery images in each class according to equation (4). Due to the use of half-toned and color confusion

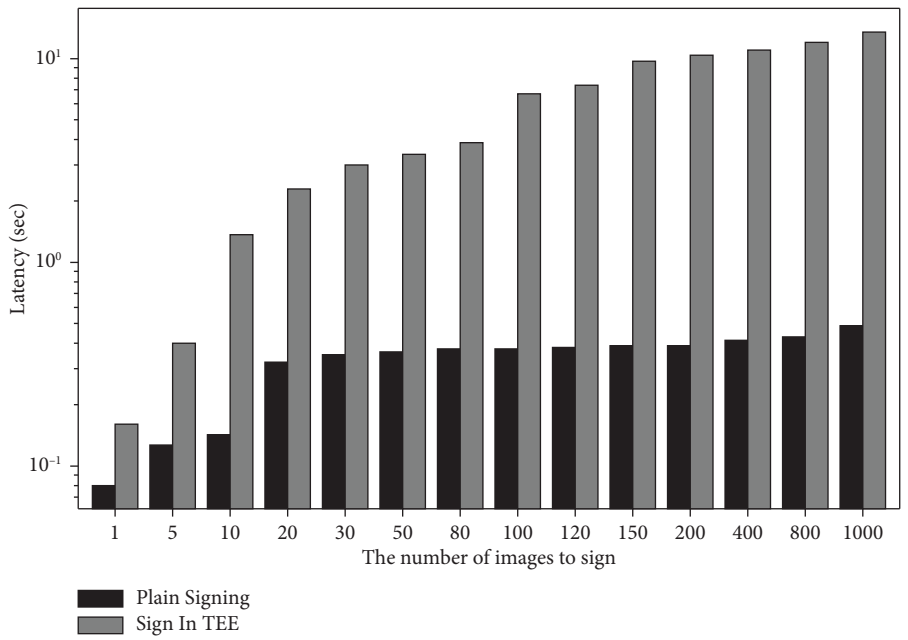

Figure 2: Comparison of the signing performance between normal and TEE environments.

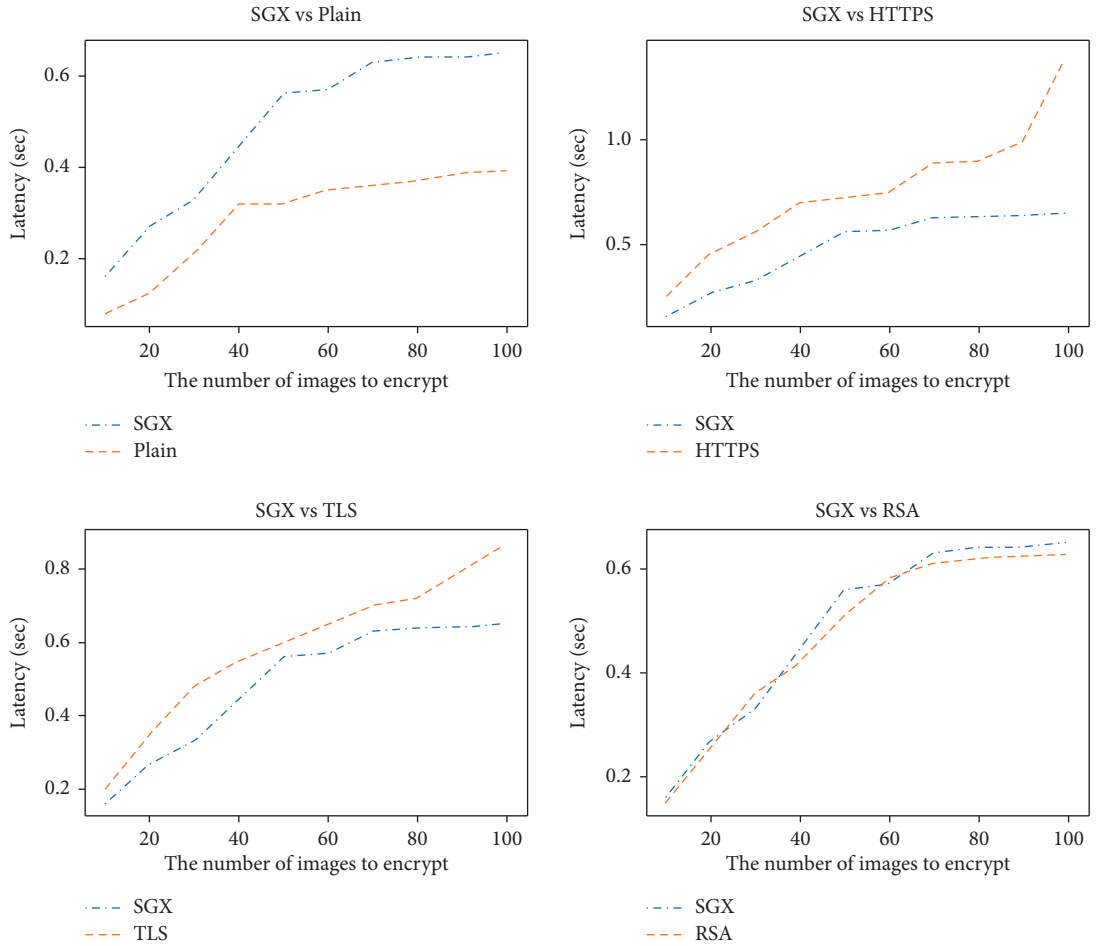

Figure 3: Comparison of SGX and other encryption methods.

Table 2: Data distribution and image quality of the DP dataset.

Class	Count	Halftone's PSNR	Our PSNR	Our SSIM	Wu's SSIM
No DR (0)	25810	27.77	27.77	7.26	2.07
Mild (1)	5292	27.76	27.71	3.34	1.32
Moderate (2)	2443	28.25	28.00	3.88	1.51
Severe (3)	873	28.17	28.13	1.78	0.92
Proliferative DR (4)	708	27.83	27.76	11.08	3.11
Total/average	35126	27.95	27.87	5.47	1.79

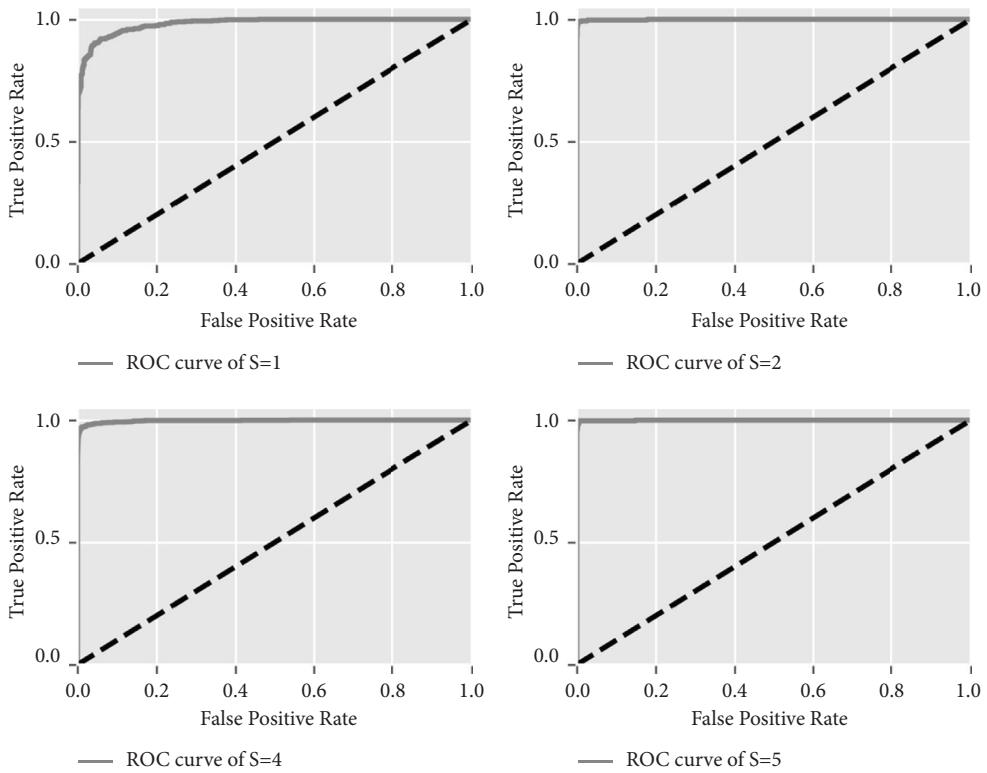

Figure 4: The ROC curves of the *BreakHis* dataset from normal images.

technologies, the quality of restored images is degraded. The image quality is almost equal to the PSNR quality of the halftoned image. The *PSNR* values are about 28, which does not affect the visual effect. We chose the *SSIM* (Structure SIMilarity) metric to further evaluate the decrypted image quality with the proposed method.

As shown in Table 2, the SSIM distinguishes images with global noise such as half-tones better than PSNR. Because our verifiable VCS is based on the random grid method, which is a lossless encryption method, although some image information is lost because of embedding TEE-based authentication information, the decryption quality of our method is still better than Wu's method [18], and the SSIM metric is about 3× that of Wu's method.

We choose another *BreakHis* [42] (breast cancer histopathological) dataset which contains 7909 breast histopathological images from 82 patients to further evaluate the method proposed. *BreakHis* not only advances the research of binary classification of benign and malignant but also advances the research of pathological classification (multiclassification), which is very significant in the clinic. Since the size of the images in the dataset varies, the shape of the input features of the neural network needs to be consistent. Therefore, to simplify the experiment, we used the smallest image in the dataset as the benchmark and then used data enhancement to obtain the preprocessed dataset with $S = \{1, 2, 4, 5\}$, and of these, 70% were used as a training set and 30% as a test set.

In the process of medical image recognition, an important factor is the resolution of the image. For example, the classic *BreakHis* provides four scales of datasets, and the visual effects of encryption and decryption of images with different scales are different. To test the performance of the proposed method in multiscale datasets, we show ROCs with four resolutions in Figures 4 and 5, respectively.

Finally, we compare the recognition performance of the normal dataset with the restored dataset using the *ResNet18* [43] backbone on the *BreakHis* dataset. Figure 4 (normal images) and Figure 5 (recovery images) show the recognition performance toward the *BreakHis* dataset in the ROC curves. The AUCs from normal and recovery datasets are 0.95 and 0.93, respectively. Despite a slight decline in

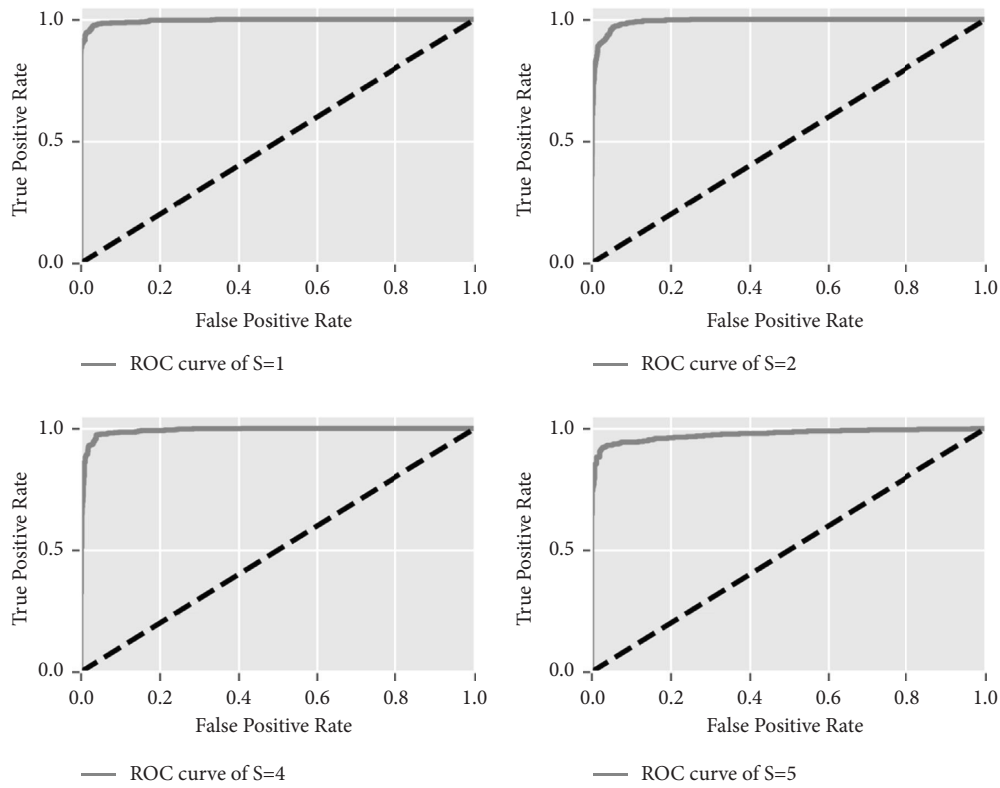

Figure 5: The ROC curves of the *BreakHis* dataset from recovery images.

performance, this experiment indicates that our find-tuned can archive high recognition even in the loss medical image dataset.

4.1. Security Analysis. As shown in the MPVCNet section, the proposed MPVCNet in this paper is divided into four parts, among which decomposition and digital half-toning belong to preprocessing, which will not affect the security of the encrypted image. In the encryption section, the existing visual cryptographic algorithms are reused, and these algorithms have been proven to be secure. To prevent the shared data from being forged, we use TEE technology to sign the encrypted shares. The whole signing process is executed in a trusted enclave. The signature key can only be accessed by a verified enclave. Even if the operating system is attacked or hijacked, it will only lead to the failure of the signature process and it is impossible to generate forged shares. In the final encryption process, the signature information will be extracted first, and then IAS will be requested for verification. Once the signature is found to not match, the system will terminate immediately. MPVCNet organically combines VC with TEE and digital signature, which ensures the efficiency and security of the whole image transmission process.

The proposed scheme maintains the visual recovery property of the VC. If the user does not have a computing device, we can still recover secret images by printing and stacking shares directly without extracting the validation information. Because we only use the lowest of the shared image to store the authentication information, the final lost gray value is less than 1/255. The complete certificate chain including local certificates, remote IAS, trusted enclave list, and pairing information during the authentication process can prevent malicious users from tampering with shared information and thus falsifying medical data. Unlike previous TLS or HTTPS-based encryption, this scheme does not require the preestablishment of a trusted channel. The entire transmission can be performed offline except that a request to the IAS is required to verify the validity of the quote. This feature facilitates the transmission of medical images since many of which are transmitted by papers or transparencies. Furthermore, unlike other trusted third-party-based authentication methods, this mechanism uses the TEE feature to remove the reliance on third-party certificate authorities. The locally maintained *MRENCLAVE* list can also be used for revocation. If an enclave explodes with a vulnerability, we can remove the ID of an enclave from the list, and then the shares generated by that enclave will not be trusted.

5. Conclusions

To address security issues after the spread of smart healthcare, we explore the possibility of using visual cryptography to preserve the privacy of medical data in this document. Simple VCS provides an effective method to allow distributed storage of images, which avoids centralized storage risks in the cloud environment. VCS is suitable for smart devices with low computing power. While existing VCS for imparting privacy of

biometric templates violates the principle, by encrypting a medical image pixel by pixel and issuing subpixels in one column of selection matrices to the corresponding participants, we archive visual cryptography without expanding the image size and keeping the computation-free feature of VCS. Although there have been various cheating-immune VCSs, we use the characteristics of TEE to effectively ensure that the sharing operation has not been maliciously tampered with, thus fundamentally ensuring the security of sharing images.

In future works, we will investigate reasons why we can keep the recognition performance for encrypted images by fine-tuning a neural network even when there is noise in the decrypted image. This reaffirms the powerful self-learning capability of deep learning and illustrates the non-interpretable machine learning since the machine learning model can maintain high recognition accuracy despite the perceived degradation of quality by the human eye. Second, we are working on class activation mapping to understand and reveal the decision-making process of the neural network on encrypted data. [5].

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (62250410365), the National Key Research and Development Program of China (2019YFB1706003), the Major Key Project of PCL (PCL2022A03), the Guangdong Key R&D Program of China (2019B010136003), the Guangdong Higher Education Innovation Group (2020KCXTD007), the Guangzhou Higher Education Innovation Group (202032854), the Guangdong Province Universities and Colleges Pearl River Scholar Funded Scheme (2019), the Guangdong Basic and Applied Basic Research Foundation of China (2022A1515011542), and the Guangzhou Science and Technology Program of China (202201010606).

References

[1] F. Al-Turjman, M. H. Nawaz, and U. D. Ulusar, "Intelligence in the Internet of Medical Things era: a systematic review of current and future trends," *Computer Communications*, vol. 150, pp. 644–660, 2020.

[2] H. Liang, M. Li, Y. Chen, T. Yang, Z. Xie, and L. Jiang, "Architectural protection of trusted system services for SGX enclaves in cloud computing," *IEEE Transactions on Cloud Computing*, vol. 9, no. 3, pp. 910–922, 2021.

[3] M. Eichelberg, K. Kleber, and M. Kämmerer, "Cybersecurity in PACS and medical imaging: an overview," *Journal of Digital Imaging*, vol. 33, no. 6, pp. 1527–1542, 2020.

[4] F. Alsubaei, A. Abuhussein, V. Shandilya, and S. Shiva, "IoMT-SAF: Internet of medical Things security assessment framework," *Internet of Things*, vol. 8, Article ID 100123, 2019.

[5] Y. Tan, J. Qin, L. Tan, H. Tang, and X. Xiang, "A survey on the new development of medical image security algorithms," in *Cloud Computing and Security*, X. Sun, Z. Pan, and E. Bertino, Eds., vol. 11065, pp. 458–467, Springer International Publishing, Berlin, Germany, 2018.

[6] M. Taheri, S. Mozaffari, and P. Keshavarzi, "Face authentication in encrypted domain based on correlation filters," *Multimedia Tools and Applications*, vol. 77, no. 13, pp. 17043–17067, 2018.

[7] T. M. Thanh and K. Tanaka, "An image zero-watermarking algorithm based on the encryption of visual map feature with watermark information," *Multimedia Tools and Applications*, vol. 76, no. 11, pp. 13455–13471, 2017.

[8] Z. Gu, H. Li, S. Khan et al., "IEPSBP: a cost-efficient image encryption algorithm based on parallel chaotic system for green IoT," *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 1, pp. 89–106, 2022.

[9] Z. Gu, L. Wang, X. Chen et al., "Epidemic risk assessment by a novel communication station based method," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 1, pp. 332–344, 2022.

[10] Y. Nakatsuka, A. Paverd, and G. Tsudik, "PDoT: private DNS-over-TLS with TEE support," in *Proceedings of the 35th Annual Computer Security Applications Conference*, pp. 489–499, San Juan, PR, USA, December 2019.

[11] M. Naor and A. Shamir, "Visual cryptography," in *Advances in Cryptology — EUROCRYPT'94*, A. De Santis, Ed., vol. 950, pp. 1–12, Springer Berlin Heidelberg, Berlin, Heidelberg, 1995.

[12] L. Tan, K. Liu, and X. Yan, "Robust visual secret sharing scheme applying to QR code," *Security and Communication Networks*, vol. 2018, pp. 1–12, 2018.

[13] M. S. Islam, M. S. Ozdayi, L. Khan, and M. Kantarcioglu, "Secure IoT data analytics in cloud via Intel SGX," in *Proceedings of the 2020 IEEE 13th International Conference on Cloud Computing (CLOUD)*, pp. 43–52, Beijing, China, October 2020.

[14] J. Mohan and D. R. Rajesh, "Enhancing home security through visual cryptography," *Microprocessors and Microsystems*, vol. 80, Article ID 103355, 2021.

[15] H.-C. Chao and T.-Y. Fan, "XOR-based progressive visual secret sharing using generalized random grids," *Displays*, vol. 49, pp. 6–15, 2017.

[16] Y. Dong, X. Huang, and G. Ye, "Visually meaningful image encryption scheme based on DWT and schur decomposition," *Security and Communication Networks*, vol. 2021, 16 pages, 2021.

[17] Y. Ren, F. Liu, T. Guo, R. Feng, and D. Lin, "Cheating prevention visual cryptography scheme using Latin square," *IET Information Security*, vol. 11, no. 4, pp. 211–219, 2017.

[18] X. Wu and C.-N. Yang, "Probabilistic color visual cryptography schemes for black and white secret images," *Journal of Visual Communication and Image Representation*, vol. 70, Article ID 102793, 2020.

[19] D. Zhang, H. Zhu, S. Liu, and X. Wei, "HP-VCS: a high-quality and printer-friendly visual cryptography scheme," *Journal of Visual Communication and Image Representation*, vol. 78, pp. 103–186, 2021.

[20] L. Zhang, X. Dang, L. Feng, and J. Yang, "Efficient secret image sharing scheme with authentication and cheating prevention," *Mathematical Problems in Engineering*, vol. 2021, 11 pages, 2021.

[21] X. Jia, D. Wang, Q. Chu, and Z. Chen, "An efficient XOR-based verifiable visual cryptographic scheme," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 8207–8223, 2019.

[22] G. Selva Mary and S. Manoj Kumar, "A self-verifiable computational visual cryptographic protocol for secure two-dimensional image communication," *Measurement Science and Technology*, vol. 30, no. 12, Article ID 125404, 2019.

[23] D. Silver, J. Schrittwieser, K. Simonyan et al., "Mastering the game of Go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, 2017.

[24] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: a unified embedding for face recognition and clustering," in *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823, Boston, MA, USA, June 2015.

[25] M. Raghu, C. Zhang, J. Kleinberg, and S. Bengio, "Transfusion: understanding transfer learning for medical imaging," 2019, https://arxiv.org/abs/1902.07208.

[26] V. Gulshan, L. Peng, M. Coram et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, 2402 pages, 2016.

[27] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[28] T. Hunt, Z. Zhu, Y. Xu, S. Peter, and E. Witchel, "Ryoan: a distributed sandbox for untrusted computation on secret data," *ACM Transactions on Computer Systems*, vol. 35, no. 4, pp. 1–32, 2018.

[29] F. McKeen, I. Alexandrovich, I. Anati, and D. Caspi, "Intel® software guard extensions (Intel® SGX) support for dynamic memory management inside an enclave," in *Proceedings of the Hardware and Architectural Support for Security and Privacy 2016 on - HASP 2016*, pp. 1–9, Seoul, Republic of Korea, June 2016.

[30] Y. Shen, Y. Chen, K. Chen, H. Tian, and S. Yan, "To isolate, or to share?: that is a question for Intel SGX," in *Proceedings of the 9th Asia-Pacific Workshop on Systems*, pp. 1–8, Jeju Island, Republic of Korea, August 2018.

[31] W. Zheng, Y. Wu, X. Wu et al., "A survey of Intel SGX and its applications," *Frontiers of Computer Science*, vol. 15, no. 3, Article ID 153808, 2021.

[32] M. U. Sardar, R. Faqeh, and C. Fetzer, "Formal foundations for Intel SGX data center attestation primitives," in *Formal Methods and Software Engineering*, S.-W. Lin, Z. Hou, and B. Mahony, Eds., vol. 12531, pp. 268–283, Springer International Publishing, Berlin, Germany, 2020.

[33] F. McKeen, I. Alexandrovich, A. Berenzon, C. Rozas, and H. Shafi, "Innovative instructions and software model for isolated execution," in *Proceedings of the 2nd International Workshop on Hardware and Architectural Support for Security and Privacy - HASP '13*, Tel-Aviv, Israel, June 2013.

[34] D. Wang, F. Yi, and X. Li, "Probabilistic visual secret sharing schemes for grey-scale images and color images," *Information Sciences*, vol. 181, no. 11, pp. 2189–2208, 2011.

[35] H.-K. Chu, C.-S. Chang, R.-R. Lee, and N. J. Mitra, "Halftone QR codes," *ACM Transactions on Graphics*, vol. 32, no. 6, pp. 1–8, 2013.

[36] Z. Gu, W. Hu, C. Zhang, H. Lu, L. Yin, and L. Wang, "Gradient shielding: towards understanding vulnerability of deep neural networks," *IEEE Trans. Netw. Sci. Eng.,* vol. 8, no. 2, pp. 921–932, 2021.

[37] B. Zhu, Z. Gu, Y. Qian, F. Lau, and Z. Tian, "Leveraging transferability and improved beam search in textual adversarial attacks," *Neurocomputing*, vol. 500, pp. 135–142, 2022.

[38] M. Shafiq, Z. Tian, A. K. Bashir, A. Jolfaei, and X. Yu, "Data mining and machine learning methods for sustainable smart cities traffic classification: a survey," *Sustainable Cities and Society*, vol. 60, Article ID 102177, 2020.

[39] M. Shafiq, Z. Tian, A. K. Bashir, X. Du, and M. Guizani, "CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine-learning techniques," *IEEE Internet of Things Journal*, vol. 8, no. 5, pp. 3242–3254, 2021.

[40] O. Russakovsky, J. Deng, H. Su et al., "ImageNet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[41] M. Abadi, P. Barham, J. Chen, Z. Chen, and X. Zhang, *TensorFlow: A System for Large-Scale Machine Learning*, USENIX Association, California, CA, USA, 2016.

[42] E. Decencière, X. Zhang, G. Cazuguel et al., "Feedback on a publicly distributed image database," *Image Analysis and Stereology*, vol. 33, no. 3, 231 pages, 2014.

[43] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 4278–4284, California, CA, USA, February 2017.