DeepEDN: A Deep-Learning-Based Image Encryption and Decryption Network for Internet of Medical Things

Yi Ding, *Member, IEEE*, Guozheng Wu, Dajiang Chen, *Member, IEEE*, Ning Zhang, *Senior Member, IEEE*, Linpeng Gong, Mingsheng Cao, *Member, IEEE*, and Zhiguang Qin, *Member, IEEE*

***Abstract*—Internet of Medical Things (IoMT) can connect many medical imaging equipment to the medical information network to facilitate the process of diagnosing and treating doctors. As medical image contains sensitive information, it is of importance yet very challenging to safeguard the privacy or security of the patient. In this work, a deep-learning-based image encryption and decryption network (DeepEDN) is proposed to fulfill the process of encrypting and decrypting the medical image. Specifically, in DeepEDN, the cycle-generative adversarial network (Cycle-GAN) is employed as the main learning network to transfer the medical image from its original domain into the target domain. The target domain is regarded as "hidden factors" to guide the learning model for realizing the encryption. The encrypted image is restored to the original (plaintext) image through a reconstruction network to achieve image decryption. In order to facilitate the data mining directly from the privacy-protected environment, a region of interest (ROI)-mining network is proposed to extract the interesting object from the encrypted image. The proposed DeepEDN is evaluated on the chest X-ray data set. Extensive experimental results and security analysis show that the proposed method can achieve a high level of security with a good performance in efficiency.**

***Index Terms*—Deep learning, image encryption, Internet of Medical Things (IoMT), medical image.**

Manuscript received May 6, 2020; revised July 7, 2020; accepted July 24, 2020. Date of publication July 28, 2020; date of current version January 22, 2021. This work was supported in part by NSFC under Grant 61872059 and Grant 61502085, in part by the Natural Science Foundation of Guangdong Province under Grant 2018A030313354, and in part by the China Postdoctoral Science Foundation Funded Project under Grant 2015M570775. *(Corresponding author: Dajiang Chen.)*

Yi Ding is with the Network and Data Security Key Laboratory of Sichuan Province, University of Electronic Science and Technology of China, Chengdu 610054, China, and also with the Institute of Electronic and Information Engineering, University of Electronic Science and Technology of China, Guangdong 523808, China (e-mail: yi.ding@uestc.edu.cn).

Guozheng Wu is with the Department of Information Science, National Natural Science Foundation of China, Beijing 100085, China (e-mail: wugz@nsfc.gov.cn).

Dajiang Chen, Linpeng Gong, Mingsheng Cao, and Zhiguang Qin are with the Network and Data Security Key Laboratory of Sichuan Province, University of Electronic Science and Technology of China, Chengdu 610054, China (e-mail: djchen@uestc.edu.cn; glpglp@std.uestc.edu.cn; cms@uestc.edu.cn; qinzg@uestc.edu.cn).

Ning Zhang is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada (e-mail: ning.zhang@uwindsor.ca).

Digital Object Identifier 10.1109/JIOT.2020.3012452

I. Introduction

THE INTERNET of Medical Things (IoMT) is an interdisciplinary field which adopts the Internet-of-Things (IoT) technologies in the domain of medicine [1]–[3]. With the development of IoMT, many medical imaging equipment are widely connected to facilitate the process of diagnosing and treating for doctors, e.g., brain magnetic resonance imaging (MRI) for brain tumor diagnosis and the computed tomography (CT) of the lung for lung nodule detection. In IoMT, medical images are usually managed by a system called picture archiving and communication systems (PACSs) [4]. When a patient is scanned by the medical imaging equipment, the generated medical images will be first stored into the PACS. When the doctor begins to examine the patient, the PACS will retrieve the needed images from the database and transfer the images to the doctors' workstation that works with the patient information from the hospital information system (HIS). Although the PACS and HIS operate in an intranet environment, there are still some critical security issues when storing, transferring, and reviewing medical images, which preserve sensitive privacy information of patients. If an attacker, either an internal or an external attacker, has the ability to intrude the PACS or HIS, it becomes much easy to eavesdrop these medical images, resulting in severe privacy information leak of patients [5]–[8].

To safeguard the IoMT system and protect the patients' privacy, encryption and decryption can be performed on medical images, e.g., data encryption standard (DES), advanced encryption standard (AES), and the hash function [9], [10]. In addition, image encryption based on chaotic systems is also employed in the literature [11]. However, these methods are hard to achieve a good balance between security performance and encryption efficiency. Deep learning also holds great potential in dealing with this issue, where multilayer neural networks extract a hierarchy of features from raw input images. The convolutional neural network (CNN) [12], [13] has demonstrated the significant advantages in computer vision [14]–[21] as well as in the image-domain transfer [22], [23]. Transferring the image from one domain onto another can be considered as a problem of texture transfer where the goal is to learn the mapping relationship between an input image and an output image from a set of aligned image pairs. One of the most popular image-to-image transformation method is the cycle-consistent adversarial networks [24],

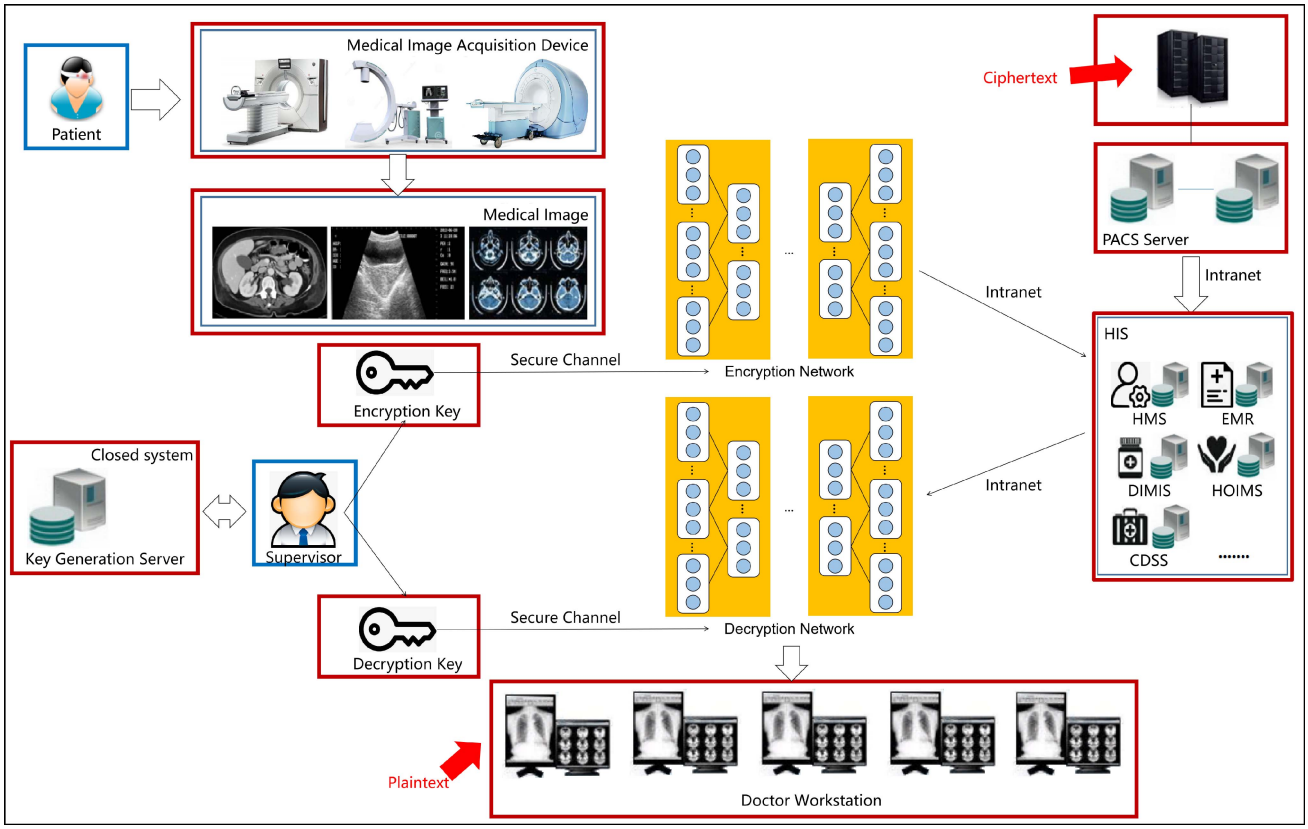


Fig. 1. Architecture of DeepEDN.

which introduces two cycle consistency losses that transform the image from one domain to the other, and then reconstruct back to the original image. In fact, the deep learning algorithm has also been adopted to solve the problem of image denoising [25].

Inspired by the above works, in this work, a deep-learning-based image encryption and decryption network (DeepEDN) is proposed for image-to-image transformation and image denoising. The novel idea is based on the following two important insights: 1) if the medical image can be transferred into other image domain that is greatly different from the original one, this medical image can be regarded as encrypted and 2) the medical image decryption process can be implemented in the manner of image denoising or image reconstruction. In DeepEDN, the cycle-generative adversarial network (CycleGAN) network is employed as the main learning network to implement the image-to-image transformation. There are two domains in the encryption process: 1) the original medical image domain and 2) the target domain, where the target domain is regarded as "hidden factors" to guide the learning model to realize the encryption process. For the encryption network, it consists of a generation network and a discriminator network. The former will generate the image similar to the target domain while the latter will promote the generation network to generate the same images as the target domain by identifying the generated images. Therefore, after processing using the encryption network, the original medical image can be converted into the target domain and becomes the ciphertext. The decryption process is similar to traditional encryption–decryption methods, which is the inverse operation of the encryption process. In practice, a reconstruction network, which is actually a decryption procedure, is used to restore the encrypted image to the original one. In DeepEDN, the parameters of the generation network is regarded as the private key for encryption while the parameters of the reconstruction network is regarded as the private key for decryption. Moreover, DeepEDN adopts the unsupervised learning to train the learning network and it does not need much labeled samples. It overcomes the data set issues in training and is beneficial to the application of deep learning in the cryptography field.

Based on DeepEDN, the PACS system is improved by employing a key generation server. As shown in Fig. 1, the key generation server is in charge of training the encryption network and the decryption network. The PACS system can call the encryption network to encrypt the medical image and then store these ciphertext images into the image database. When reviewing, the HIS system will adopt the decryption network to decrypt the ciphertext image to the original one. The encryption network and the decryption network will be transferred over the secure channel. Moreover, a region of interest (ROI)-mining network is proposed to directly extract the ROI (organ or tissue) from the encrypted medical image without decryption. More specifically, when inputting an encrypted medical image into the ROI-mining network, the interested segmented object can be

directly extracted without revealing other parts of the patient's information.

In a nutshell, the main contributions of this work are summarized as follows.

1) A novel medical image encryption and decryption network, DeepEDN, is developed to realize the encipherment process by applying deep learning in the field of image-to-image transformation. The proposed encryption method is with the large key space, one-time pad (OTP), and be sensitive to a key change. To the best of our knowledge, this work is the first work to attempt to adopt the deep learning method in the area of medical image encryption.
2) An ROI-mining network is proposed to directly extract the interested segmentation region from the encrypted medical image instead of decrypting the ciphertext image first. From the experiments, it can be found that the proposed approach can realize the data mining process directly from the privacy-protected environment.
3) Extensive experiments are conducted on the chest X-ray data set to evaluate the proposed DeepEDN. The results demonstrate that the medical image can be transmitted with a high level of security and efficiency, compared with existing medical image encryption methods. Moreover, the proposed encryption algorithm can resist various attacks, even if the attacker has known the complete process for key generation.

The remainder of this article is organized as follows. Section II gives an introduction of image encryption and deep learning. Section III presents the details of the proposed DeepEDN. Section IV analyzes the security of the proposed method. Section V shows the encryption and decryption performance and evaluates network efficiency. Finally, Section VI gives a summarization.

II. Related Work

A. Medical Image Encryption

In the literature, there are many approaches proposed for image encryption [26]–[28]. Based on the transformation of cosine number, Lima and Neto [26] proposed a novel scheme for encrypting the medical images. The proposed schema is a mathematical tool only requiring modular arithmetic so as to prevent rounding-off errors. It is a flexible method that can be applied to the medical images in the format of DICOM. Natsheh *et al.* [27] proposed a simple yet effective encryption approach for multiframe DICOM medical images. It can accelerate the encryption and decryption process by using AES. Mukhedkar *et al.* [28] demonstrated that image encryption can be done by using a faster Blowfish algorithm and image hiding technique LSB.

Chaotic maps have the characteristics of pseudorandomness, ergodicity, and initial value sensitivity. Chaotic sequences generated by chaotic maps have good characteristics of security keys. Chaotic cryptography has gradually become a new research direction of cryptography. Medical image encryption can also be performed using chaotic maps. Kanso and Ghebleh [29] proposed a novel selective chaos-based image encryption scheme to encrypt the medical image, which consists of several rounds. Each round is composed of two block-based phases, a shuffling phase and a masking phase, which both adopt chaotic cat maps to encrypt an input image. Fu *et al.* [30] presented a novel chaos-based medical image encryption scheme. In order to improve efficiency, this method adopts a bit-level shuffling algorithm as a substitution mechanism in the process of permutation. In [31], an image encryption algorithm, which is based on the wavelet function and 4-D chaotic system, is proposed to effectively protect the image. The wavelet function is used to scramble the pixel location in the image while the 4-D chaotic system is used to disturb the pixel value.

B. Image-to-Image Transfer by Generative Adversarial Networks

Based on the creative work proposed by Goodfellow *et al.* [32] in 2014, many researchers put their focus on the usage of GAN-based methods in different applications. The GAN network consists of a generator and an adversarial discriminator. The former one is used to capture the data distribution while the latter one evolves to distinguish the fake data from the real data [32]. GAN-based methods can achieve state-of-the-art results in a wide variety of applications, such as image generation [33], image segmentation [34], image super resolution [35], and image-to-image transformation [23], [41].

Arroyo *et al.* [43] used a conditional generative adversarial network (CGAN) to implement the image-to-image transformation. It is proved that this approach achieves a better performance on synthesizing photos from label maps, reconstructing objects from edge maps, and colorizing images. DualGAN [42] mechanism enables to train the learning network from two sets of unlabeled images. Taking two sets of unlabeled images as the input, DualGAN learns two reliable image transformation networks at the same time and hence can facilitate a lot of image-to-image transformation tasks. In [24], Cycle-GAN is proposed to fulfill the image transformation task with unpaired images. The Cycle-Gan simultaneously trains two sets of GAN models. One model is used to learn the mapping from class A to class B and the other one learns the mapping from class B to class A. The loss is redefined with the combination of these two mappings. The success of GAN depends on the idea of an adversarial loss, which forces the generated images to distinguish from target images. The adversarial loss is used to learn the mapping of the original images to the "target-domain images," which represents the image-to-image transformation.

Regardless of their merits, these algorithms are difficult to achieve a good balance between security and efficiency. On the one hand, as there is plenty of information in a single medical image with a high correlation among this information, when encrypting the medical image, the block encryption algorithm is with low efficiency and cannot meet the real-time requirement. On the other hand, the chaotic system usually adopts the 1-D chaotic map to generate pseudorandom sequences. Consequently, the chaotic system tends to be easy to analyze

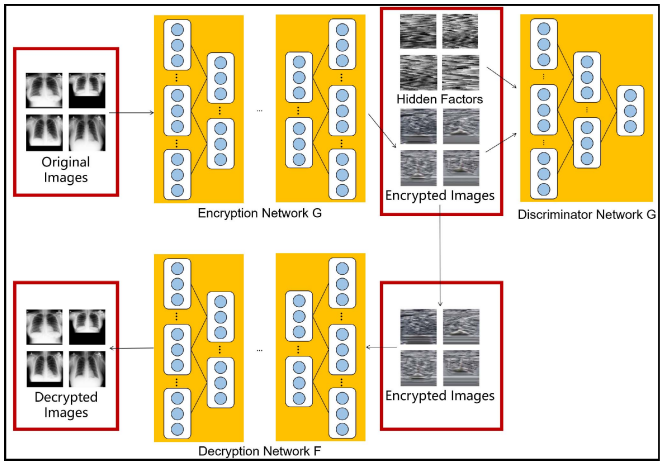


Fig. 2. Overall framework of DeepEDN.

and predict through a nonlinear prediction method based on phase-space reconstruction [43]. The deep learning algorithm has been employed in the security field [44]–[46]. However, there is no work on medical image encryption and decryption.

In this article, deep learning techniques are used to encrypt and decrypt medical images, in which parameters of the deep learning network model are regarded as the encryption and decryption keys. Due to the large key space and the complex model structure, the proposed method can achieve a high level of security with high efficiency.

III. Encryption and Decryption Network

A. Architecture of DeepEDN

The traditional GAN-based network used for domain transfer learning can only convert the original image into the target-domain image. However, facing with the image encryption and decryption, except for converting the original medical image into a ciphertext image, we also need to restore the ciphertext image to the original image, that is, image reconstruction. It can be found that the Cycle-Gan network achieves a good performance in the area of domain transformation and image reconstruction through the generating adversarial loss and cyclical consistent loss. Therefore, the Cycle Gan network is adopted as the learning network in this work.

As shown in Fig. 2, DeepEDN mainly consists of three subnetworks: 1) the encryption network G; 2) the discriminator network D; and 3) the decryption network F. The encryption network G is used to encrypt the original input images, the decryption network F is responsible for restoring the encrypted images to the original one (decrypting the image), and the discriminator network D is mainly designed for improving the performance of the encryption network by distinguishing the generated images from the images in the target domain (hiding factors). In deep learning methods, the loss function is usually used to train the model. The overall loss L of the proposed model is given as follows:

$$L = L_G + L_D + L_R \tag{1}$$

where L_G indicates the loss of the encryption network G, L_D indicates the loss of the discriminator network D, and L_R indicates the loss of the decryption network F.

1) Encryption Network and Decryption Network: The encryption network G is used to transform the original medical images into the target domain for medical image encryption. The G network begins with an initial convolution stage to spatially downsampling and encode the images, and the useful features obtained in this stage will be used for the following transformation. Then, nine residual blocks [48] are performed to construct the manifold and content features. The output images are reconstructed with two up-convolution blocks which contain a strided convolutional layer and the stride is set to 2. Finally, the prediction is exported by a 7×7 convolution kernel. In addition, the structure of the decryption network F is the same as the encryption network G.

The proposed model includes two mappings $G : X \to Y$ and $F : Y \to X$. The goal of the mapping function G is to learn how to transform the original medical images X into the images Y in target domain, and cheat the discriminator network D. When the discriminator network D cannot successfully distinguish whether an image is generated by the encryption network G or a real ciphertext image domain Y, it means that the encryption network G converts the original patient image domain X into a ciphertext image domain Y successfully. The loss L_G of the encrypted network G is

$$L_G = \min_G \left(E_{x\sim p\text{data}(x)} \log(1 - D(G(x)))\right) \tag{2}$$

where G represents an encryption network and D represents the discriminator network. The goal of L_G is to minimize the success rate of the discriminator network D for detecting the ciphertext generated by the encryption network G. In addition to the encryption, another goal of the proposed method is to ensure that the restored image reserves the texture information of the original one even it is encrypted. As shown in Fig. 2, for each image x from domain X, the reconstruction loss measures the difference between $G(x)$ and the original image, i.e., $x \to G(x) \to F(G(x)) \approx x$. The reconstruction loss L is defined as

$$\begin{aligned} L_R &= E_{x\sim p_{\text{data}(x)}} ||F(G(X) - X||_1 \\ &= E_{x\sim p_{\text{data}(x)}} \sum_{i=1}^{n} |F(G(x_i) - x_i| \,. \end{aligned} \tag{3}$$

2) Discriminator Network: The discriminator network D is used to evaluate whether the output image of the encryption network belongs to the target domain. For the discriminator network D, after processing with initial convolutional layers, two strided convolutional blocks are adopted to reduce the resolution of the image and to encode essential local features for subsequential discrimination. Then, the network employs a feature construction block and a 3×3 convolutional layer to obtain the final result. In addition, for each convolutional layer, the leaky ReLU (LReLU) with $\alpha = 0.2$ is adopted and followed with batch normalization (BN) layer.

The training of the discriminator network D is to classify the images and check whether it comes from the ciphertext domain Y or is generated by the encryption network G. The encryption network G attempts to generate an image $G(x)$ that is similar to

the image in the domain Y, while the discriminator network D aims to find the difference between transformed samples from $G(x)$ and real samples in Y. The minimizing loss L_D of the discriminator network D is equivalent to maximizing the classification accuracy of the discriminator network D, which is opposite to the goal of the encryption network G. The loss L_D is given as follows:

$$L_D = E_{x\sim p\text{data}(x)}\log D(x) + E_{x\sim p\text{data}(x)}\log(1 - D(G(x))) \quad (4)$$

where G represents the encrypted network and D represents the discriminator network. L_D and L_G in the GAN network form an adversarial relationship. When the two networks reach an equilibrium state, the discriminator network D can achieve 50% classification accuracy for both the generated ciphertext image and the real ciphertext domain image Y. In other words, the ciphertext image generated by the encryption network G is very similar to the real ciphertext domain Y so that the discriminator network D cannot distinguish them.

3) Key Generation Process: In DeepEDN, the final parameters of the network G can be considered as the private key for encryption while the parameters of the network F are regarded as the private key for decryption.

For encryption, the parameters for each convolutional layer are first randomly initialized as follows:

$$W_n = \text{random}\left[w_{n,1}, w_{n,2}, \ldots, w_{n,j}, \ldots\right] \quad (5)$$

where w_n is the nth convolutional layer and $w_{n,j}$ is the jth parameter of one convolutional layer. Therefore, the private key W for encryption is actually composed of all the parameters of each convolutional layer, and is defined as follows:

$$W = \text{consist}[W_1, W_2, \ldots, W_n, \ldots]. \quad (6)$$

When training the encryption network, the private key for encryption is continuously updated and refined with different input images through the forward propagation training process. The adversarial loss L_{gan} is calculated to measure the difference between the predicted result and the target one in "hidden factors," thereby guiding the network to train and update the private key for encryption.

Except for forward propagation, the backpropagation algorithm (BP) is also employed to pass loss of the entire network back to the convolutional layers. It is actually a gradient descent, which can further update the parameters in each layer to achieve better performance. The gradient descent can be described as

$$\begin{aligned}\theta_j &= \theta_j - \alpha \vee J(\theta) = \theta_j - \alpha\frac{\delta}{\theta_j}J(\theta)\\ &= \theta_j - \alpha\frac{\delta}{\theta_j}\frac{1}{2m}\sum_{i=1}^{m}\left(h_\theta\left(x^i\right) - y^i\right)^2\\ &= \theta_j - \alpha\frac{1}{2m}\sum_{i=1}^{m}2\frac{\delta}{\theta_j}\left(h_\theta\left(x^i\right) - y^i\right)\left(\frac{\delta}{\theta_j}\left(h_\theta\left(x^i\right) - y^i\right)\right)\\ &= \theta_j - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta\left(x^i\right) - y^i\right)\left(\sum_{i=0}^{n}\frac{\delta}{\theta_i}\theta_i x_i - \frac{\delta}{\theta_i}y^i\right) \quad (7)\end{aligned}$$

Algorithm 1 Image Encryption/Decryption

Initialization: Digitize the 255×255 image into a 255×255 matrix X_*^0. And then enter it into our 21-layer(L_c) encryption/decryption model.

1: **while** L< L_c **do**
2: **for all** element (X_1^L, X_2^L, ...) in matrix X^L **do**
3: Each pre-trained 3×3 convolution kernel W_*^L in L^{th} layer sequentially traverses the image matrix and multiplies it with the corresponding elements of the matrix($W_*^L \times X_*^L$).
4: Add the obtained nine W_*X_* to get a new predicted value X_*^{L+1} in $(L+1)^{th}$.
5: Collect all X^{L+1} and combine them into a new matrix to form the next-level feature matrix.
6: **end for**;
7: $L = L + 1$;
8: **end while**

Output: Convert the last layer of matrix X^{L_c} into an image to get the final encrypted/decrypted image.

where θ_j is the value of parameter θ in the jth training epoch. α is the learning rate and $\vee J(\theta)$ means the gradient that is passed back to the convolution layer θ in the jth training epoch.

The generation process of the private key for decryption is similar to the process of generating the privacy key for encryption, except that the initial input of the decryption network is the predicted result of the encryption network. In addition, the loss of the decryption network is the reconstruction loss, which is given in

$$\begin{aligned}L_{\text{reconstruction}} &= E_{x\sim p_{\text{data}(x)}}||F(P(X)) - O(X)||_1\\ &= E_{x\sim p_{\text{data}(x)}}\sum_{i=1}^{n}|F(P(x_i)) - O(x_i)| \quad (8)\end{aligned}$$

where $F()$ is the decryption network, $P(x)$ is the pixel x in the predicted image, and $O(x)$ is the corresponding position pixel x in the original image. The encryption network G and the decryption network F are trained in an alternative manner. When the loss becomes stable, the final parameters (privacy keys) for the encryption and decryption network can be obtained. The complete privacy-key generation process is presented in Fig. 3.

After obtaining the key, the patient's medical image can be encrypted by the encryption network G and then decrypted by the decryption network F. The proposed medical image encryption/decryption algorithm is given in Algorithm 1.

Since the GAN model is highly nonlinear and randomly initialized, and the parameters of the learning network can be totally different at different training times. In other words, the GAN network is unstable, which is its weakness when used for computer vision tasks. However, this instability has advantages for cryptography. By utilizing this instability, the proposed deep-learning-based encryption method can be regarded as an OTP method. Specifically, the parameters of the encryption network are totally different after training the network at different times. Overall, due to the depth and complex structure of the learning encryption network, the proposed framework is with higher security.

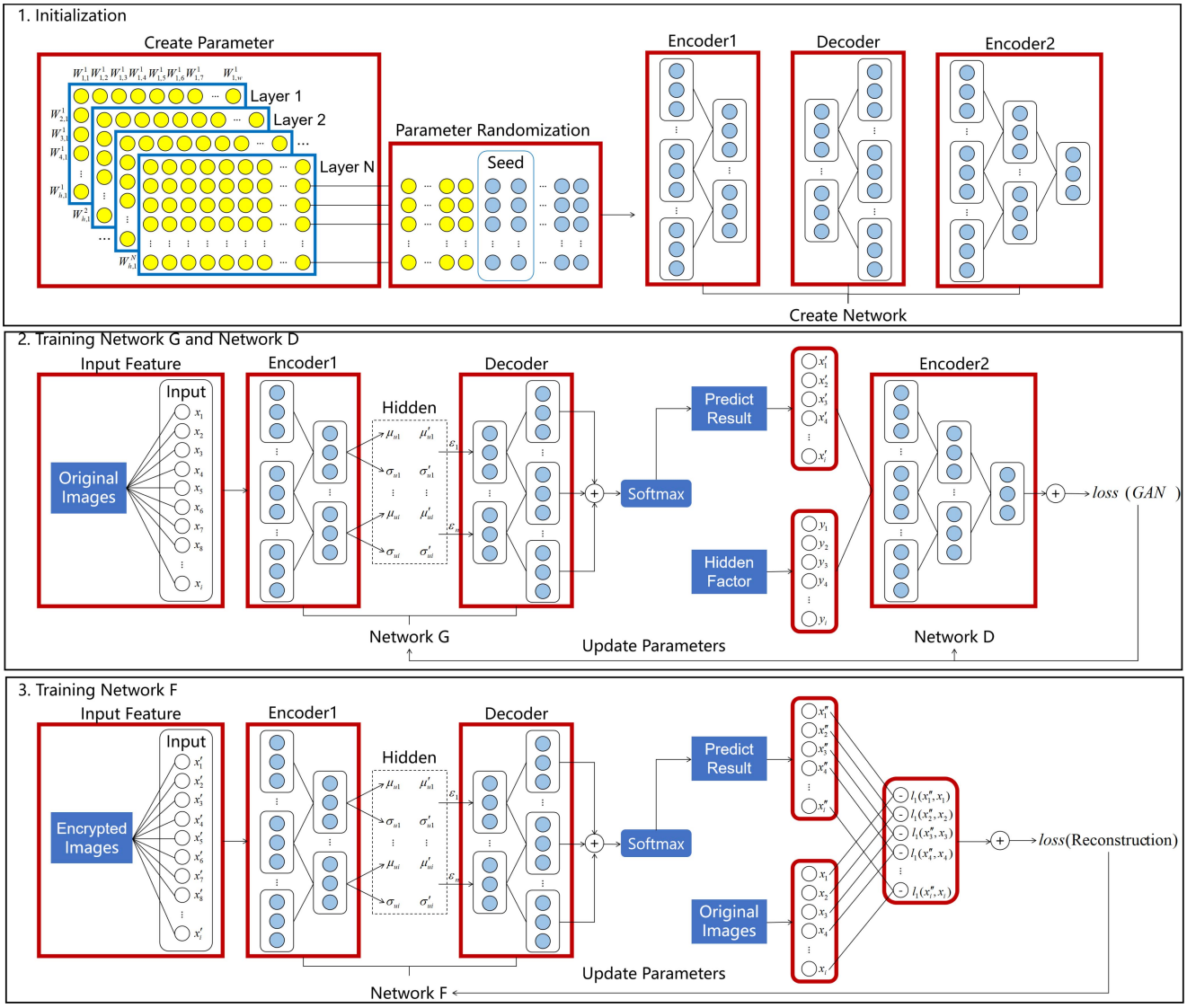


Fig. 3. Key generation process.

B. ROI Mining Network in Ciphertext Environments

Although various methods have achieved a good performance in protecting the image privacy, it is still a challenge to directly obtain the effective information in a ciphertext environment, e.g., extracting the desired ROI from the encrypted medical image. In DeepEDN, an ROI-mining network is proposed to segment the ROI from the encrypted medical image. In order to extract useful texture features in a ciphertext environment, a deeper network structure is adopted to learn semantic features to accurately segment the specific target. The input encrypted image will be processed with five blocks, and each block has a downsampling convolution. In the first block, since the convolutional kernel size is set to 3×3, each convolution operation can learn the local information from the input image. As the increasement of the network depth, more abstracted semantic information can be obtained. Finally, by combining the output results from each convolution layer, the final prediction results can be achieved.

Each block in ResNet-50 has two subblocks. One is the identity (ID) block in which the stride of each convolution layer is 1. The ID block is mainly used to extract abstract features through multilayer convolution. Since the dimensions of the input and output are the same, these feature maps can be serially connected. The other basic block is the Conv block, where the dimensions of input and output are different and it is used to change the dimension of the feature vector and to resize feature size through a strided convolutional layer. The CNN-based neural network commonly converts the image into a small feature map with many channels. However, by increasing the network layers, there will be a huge number of output channels and parameters, resulting in increased computational complexity and reduced network efficiency. Therefore, it is necessary to reduce the dimension of the Conv block before processing it with the ID block.

In DeepEDN, the proposed ROI-mining network is used to implement the medical image segmentation task in the ciphertext environment. Medical image segmentation is a key

step in medical image analysis. Its purpose is to extract useful features and segment the doctors' interested objects. The segmentation results can provide a reliable basis for clinical diagnosis and pathological research. When training the ROI-mining network, the encrypted medical image is first used as the input of the network. Then, the pixel-level segmentation labels in the corresponding medical image are adopted to supervise the training process. Finally, the model parameters are updated by the mean square error (MSE). The loss function of this segmentation model is described as

$$L_S = \frac{1}{N}\sum_{i=0}^{N}(g_i - p_i)^2 \tag{9}$$

where g_i represents the value of the ith pixel in the label and p_i is the predicted value of the ith pixel in the predicted result. N represents the total number of pixels in this image. The final training result is a high-quality splitter that can segment the medical images without decryption.

The usage of the ROI-mining network is of great significance for medical image security. It can implement data mining in an untrusted environment to securely extract specific objects, which is also beneficial for protecting the privacy of the patient. This network can further improve the security of medical image analysis and can be widely used in many medical applications.

C. Adversary Model

In DeepEDN, the most important factors of the key generation process include the structure of the model and the chosen hidden factors. If the network structure or hidden factors leaks, the attacker can train a similar encryption network by imitating the private key generation process so as to crack the ciphertext image. This kind of attack is called the imitation learning attack. This article proposes three possible adversary models for imitation learning attack: 1) the hidden factors leakage; 2) the network architecture leakage; and 3) both the hidden factors and the network architecture leakage.

1) Hidden Factors Leakage: Hidden factors leakage means that the attacker has known the hidden factors used for the encryption, and tries to employ the same hidden factors to train the attacking network with several different network architectures to decrypt the ciphertext image. There are two encryption and decryption networks with different network structures: 1) the encryption/decryption network A and 2) encryption/decryption network B. These two encryption and decryption networks are trained with the same hiding factor. If the decryption network B is able to recover the image encrypted by the encryption network A, it means that the attacker can crack the secure key by the imitation learning attack.

2) Network Architecture Leakage: The network architecture leakage assumes that only the architecture of the encryption and decryption network is leaked, and the hidden factors remain confidential. In this adversary model, the attacker can decrypt the encrypted image by training the same network structure without knowing the hidden factors. The attacker can employ different hidden factors to train the same network structure to construct different decryption networks. If the attacker is able to recover the encrypted ciphertext image, the attack is successful.

3) Both Hidden Factors and Network Architecture Leakage: The strongest adversary model is that both the network architecture and hidden factors are leaked. In such a scenario, the attacker can train the network with the same network structure and hidden factor adopted for training the encryption/decryption network. To prevent such attacks, after each training of the network, the parameters of the encryption/decryption network representing the actual private key must be completely different. It means that the proposed encryption algorithm should be similar to the OTP and can be regarded as a chaotic encryption algorithm.

IV. Security Analysis

In DeepEDN, both encryption and decryption networks are constructed with 24 layers and the number of the parameters for each network is 2 757 936. The explicit specification of the network is represented in Table I. For the ROI-mining network, a deeper resnet-50 architecture is adopted. The network structure of the ROI-mining network is given in Table II. The data set is chest X-rays [48]. This data set is provided by the U.S. National Library of Medicine to foster research in computer-aided diagnosis of pulmonary diseases. It also focuses on pulmonary tuberculosis (TB). All radiographs were acquired from the Department of Health and Human Services in the USA and Shenzhen No. 3 People's Hospital in China. The data sets contain normal and abnormal chest X-rays with manifestations of TB and corresponding radiologist readings. The proposed method is running on the Nvidia GTX 2080Ti. When training the network, it takes around 10 min for each epoch of the model.

A. Key Security Analysis

The ideal encryption scheme has the following characteristics: 1) the key space is large enough so that it can effectively resist the exhaustive attack under the premise of the existing computing power; 2) the key generated for each time should be different, i.e., the key generation should be uniform at random; and 3) the encrypted image must be highly sensitive to the key. The security of the key will be analyzed from these three characteristics in the following sections.

1) Key Space Analysis: The size of the key space determines the difficulty of an attacker encounters using an exhaustive attack. In this work, the key space of the proposed encryption algorithm is the number of parameters for the deep learning network, with a total of 2 757 936 parameters in the experiments. Each parameter or key is a floating-point number between 0 and 1, which is 32 b in the computer and can be expressed as a decimal number with ten significant digits. Therefore, the key space of the encryption model can be expressed as $(2^{32})^{2757936}$. It becomes very hard for attackers to break the system and the proposed scheme can effectively resist attacks.

2) Key Randomness Analysis: The encryption network is trained four times with the same settings. Accordingly, the

TABLE I
STRUCTURE OF THE ENCRYPTED NETWORK AND DECRYPTION NETWORK

Convolution Layer Name	Number	Size	Input Channels	Output Channels	Parameters	Total Parameters
Down Convolution1	1	7×7	3	32	4704	4704
Down Convolution2	1	3×3	32	64	18432	23136
Down Convolution3	1	3×3	64	128	73728	95864
Residual Blocks	18	3×3	128	128	2564208	2661072
Up Convolution1	1	3×3	128	64	73728	2734800
Up Convolution2	1	3×3	64	32	18432	2753232
Up Convolution3	1	7×7	32	3	4704	2757936

TABLE II
STRUCTURE OF ROI-MINING NETWORK

Convolution Layer Name	Number	Size	Input Channels	Output Channels	Parameters	Total Parameters
Block 1	2	7×7	3	64	4704	4704
Block 2	3	3×3	64	256	18432	23136
Block 3	12	3×3	256	512	73728	95864
Block 4	18	3×3	512	1024	2564208	2661072
Block 5	1	3×3	1024	2048	73728	2734800

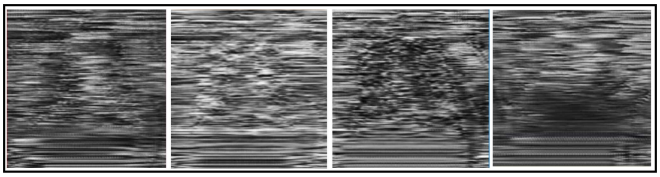


Fig. 4. Same image is encrypted with the key obtained from four networks.

TABLE III
SSIM BETWEEN TWO ENCRYPTED IMAGES

Image	A	B	C	D
A	1	0.07	0.11	0.09
B	0.07	1	0.08	0.04
C	0.11	0.08	1	0.05
D	0.09	0.04	0.05	1

parameters of these four networks are adopted as encryption keys, i.e., key A, key B, key C, and key D, respectively. The same image is encrypted with these four keys, and the encrypted images are shown in Fig. 11. Fig. 4(a)–(d) shows the results obtained by encrypting the same original image from four networks. It is clear that these four images are different. The similarity among these four encrypted images (SSIM) is calculated, and the result can be found in Table III. The SSIM index between different images is mostly lower than 0.1, which indicates that the similarity between different images is very low.

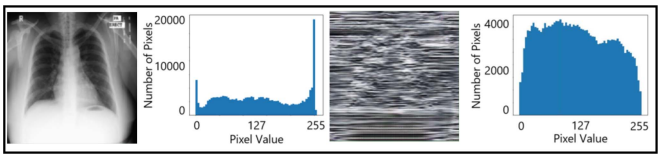
According to the experiment, it can be found that since parameters of the neural network are randomly initialized, the private keys for the medical image encryption network are totally different after every training. These difference results in different encrypted images, which are processed with different encryption networks. The idea behind this is the training of the deep learning network is not stable. Different initialized parameters can lead to a dramatic difference in final parameters in different training. It can be demonstrated that the proposed method is similar to OTP and can be regarded as one type of OTP method.

3) Key Sensitivity Analysis: Unlike traditional encryption algorithms, the error in deep learning models will be propagated among layers. In the convolution process, the *l*th pixel in the *N*th layer feature map is passed to a neighboring pixel of the $(N+1)$th layer via a 3×3 convolution kernel. When a feature point is erroneous, it will be passed to the 3×3 feature points in the next layer. As the depth of the convolutional network increases, the error of feature points will increase with two pixels for each layer. In the upsampling process, this error increases exponentially with the superposition of the deconvolution operation. The experiment assumes the attacker knows the most private keys, and only about 5% of key parameters is modified which is regarded as the unknown part. Then, the encrypted image is input to the network with new parameters, and the network cannot decrypt the ciphertext image to the original one. It means that even if only 5% of the parameters is changed, the private key cannot encrypt or decrypt the medical image correctly. In other words, it becomes very hard for attackers to guess at least 95% of the right key parameters in a key space with $(10^{10})^{2757936}$ so as to break the proposed algorithm.

B. Ciphertext Security Analysis

1) Histogram Analysis: To evaluate the performance of the proposed encryption network, the original image is shown in Fig. 5(a) and the encrypted image is shown in Fig. 5(c). Through the experiment, it can be found that the pixel distribution of the original image and the encrypted image is quite different. In Fig. 5, the pixel histogram of the original chest x-ray image has a total of 57600*(240 * 240) pixels [Fig. 5(b)], in which more than 30 000 pixels have a value of 0 and more than 5000 pixels have a value of 255. The pixel distribution of the original image is relatively concentrated. However, the distribution of encrypted medical images [Fig. 5(d)] is more uniform, which is helpful for mitigating the statistical analysis.

2) Entropy Analysis: The information entropy of the encrypted image is regarded as an effective quantitative measurement for algorithms against statistical attacks. The image information entropy represents the statistical feature of the



(a) (b) (c) (d)

Fig. 5. Pixel distribution of the original image and the encrypted image.

TABLE IV
EVALUATION OF THE ENTROPY EFFECT OF OUR NETWORK

Image Id	**1**	**2**	**3**	**4**	**5**
Entropy	7.96	7.96	7.95	7.94	7.95
Image Id	**6**	**7**	**8**	**9**	**10**
Entropy	7.97	7.95	7.96	7.96	7.95

TABLE V
NETWORK MODEL OF DIFFERENT ARCHITECTURES

Convolution Layer	**Net. A**	**Net. B**	**Net. C**	**Net. D**
Down Convolution1	1	1	1	1
Down Convolution2	1	1	1	1
Down Convolution3	1	1	1	1
Residual Blocks	18	15	12	9
Up Convolution1	1	1	1	1
Up Convolution2	1	1	1	1
Up Convolution3	1	1	1	1

grayscale distribution of the image. In an ideal case, the encrypted image should be similar to random noise, the grayscale distribution tends to be uniform, and the expected value should be 8. The information entropy formula is defined as follows:

$$\text{Entropy} = -\sum_{l=0}^{N} p(l)\log_2(p(l)) \tag{10}$$

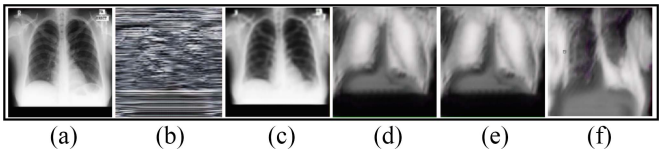
where N is the number of gray levels of the pixel value and $p(l)$ is the probability that the pixel value l appears. The entropy metric is calculated on the encrypted medical image, and the results are given in Table IV. It is clear that the image encrypted by the proposed method is close to the ideal value of 8 on information entropy. The experiments show that the images encrypted by the proposed method have the ability to resist statistical attacks.

C. Security Analysis Under Different Adversary Models

The experiments are conducted to validate whether an attacker can generate a key under three different adversary models.

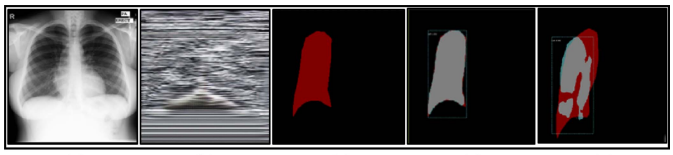
1) Hidden Factors Leakage: In this experiment, four different network structures are considered, namely, network A, network B, network C, and network D. The training conditions are kept the same. The network structure of these four networks is shown in Table V and the number represents the number of layers in each convolution layer.

The original image is encrypted by using the trained network A. The ciphertext image is then decrypted by the decryption network obtained from network A, network B, network C, and network D, respectively, to restore the original image. As shown in Fig. 6, the original image [Fig. 6(a)] encrypted by network A [the encrypted image is shown in Fig. 6(b)] can only be correctly decrypted by the decryption network A as shown in Fig. 6(c). While the image decrypted by network B, network C, and network D is visually unrecognizable and the result is shown in Fig. 6(d)–(f), respectively. The experiments show that even if the attacker knows the hidden factors, the "attack network" trained with different network structures still cannot be used to decrypt the ciphertext image.



(a) (b) (c) (d) (e) (f)

Fig. 6. Decryption performance for different networks.



(a) (b) (c) (d) (e)

Fig. 7. Mutual decryption performance between networks under different hidden factors training.

2) Network Architecture Leakage: In this experiment, different hidden factors are adopted to train the encryption network with the same network structure. All training conditions are kept the same. As shown in Fig. 7(a) and (b), two different domain images ("hidden factors A" and "hidden factors B") are chosen as hidden factors to train the network with the same architecture. Fig. 7(c) is the original image, Fig. 7(d) is the image generated by the encrypted network which is trained by "hidden factors A," and Fig. 7(e) presents the result of decrypting the ciphertext image through the decryption network trained with "hidden factors B." From the experiment, it can be found that the image generated by the encrypted network which is trained by "hidden factors A" cannot be decrypted by the network trained by "hidden factors B." Therefore, it can be proven that the "attack network" with the same architecture trained by different hidden factors cannot be used to decrypt the ciphertext image with each other. That is, even if attackers obtain the network architecture, they cannot train the decryption network to decrypt the encrypted image without knowing the hidden factors.

3) Both Hidden Factors and Network Architecture Leakage: In this experiment, the network is trained with four times under the same hidden factors and training conditions to get the networks A, B, C, and D, respectively. The experiment evaluates the decryption performance for these four networks on the same ciphertext image to verify whether the parameters generated for each network are different. As shown in Fig. 8, the gray value distribution of the image decrypted by the decryption key B, the decryption key C, and the decryption key D is completely different from the image decrypted by the decryption key A. It can be clearly found that under the same training condition, the encrypted medical image encrypted by one network cannot be decrypted by adopting the parameters

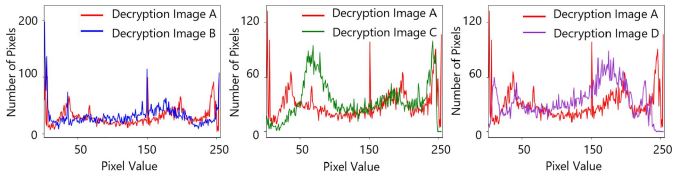


Fig. 8. Decryption performance for these four networks on the same ciphertext image.

in other network. Even if the model parameters are trained with the same network architecture and the same hidden factors, they cannot be used to decrypt the image with each other. The experiments show that even if both the network architecture and the hidden factors are leaked, and training the network under the same training conditions, the parameters of each network are totally different, i.e., the secure keys are different. It can be proven that DeepEDN is secure even if the network architecture and hidden factors are revealed.

D. Security Analysis Under Different Attack Models

1) Ciphertext Only Attack: In this type of attack, the attacker has access to a string of ciphertext, but cannot access the corresponding plaintext.

In DeepEDN, the key space of the encryption model can be expressed as $(2^{32})^{2757936}$ and it is very hard for the attacker to break down. At the same time, the privacy key generated with multiple iterations and diffusions is complex. Therefore, it is difficult to crack the ciphertext through ciphertext only attacks.

2) Known Plaintext Attack: The known plaintext attack means that the attacker knows a string of plaintext and the corresponding ciphertext. The attacker will try to decrypt the rest of the ciphertext by using these known information.

In traditional sequential pixel visiting pattern methods, concrete encryption factors, which are generally retrieved as equivalent keys, can be used to recover the received ciphertexts. Taking XOR encryption as an example, the masks calculated directly from plaintext and ciphertext are sufficient to decode the ciphertext. Typically, masks sequentially correspond to the plain pixels and the retrieved masks by plaintext attack can be directly adopted to crack other ciphertexts. However, the proposed algorithm adopted the nonsequential encryption mechanism. Without the knowledge of the pixel visiting pattern, the privacy key cannot be obtained by the attacker, thus making the plaintext attack infeasible. The proposed algorithm adopts the iteration and diffusion procedures to generate the privacy key. These kinds of producers can significantly improve the security performance and provide additional immunity of the cipher against the known plaintext attack.

3) Chosen Plaintext Attack: In this type of attack, the attacker can access the encryption device, choose a string of plaintext, and construct its corresponding ciphertext string.

Generally, an attacker can observe the change of the ciphertext image by making small changes to the plaintext image, such as changing the value of only one pixel of the ciphertext, so as to obtain the connection between the plaintext image and the ciphertext image. This type of attack is called the differential attack, which is a kind of chosen plaintext attack method. If a small change in the plaintext image can cause a huge change in the ciphertext image, this differential attack method usually fails to take effect. It indicates that the encryption algorithm can resist this chosen plaintext attack method. Here, the number of pixel change rate (NPCR) is adopted to measure the degree of image changing. NPCR refers to the rate of pixels change which indicates the ratio of different pixel values at the same position between two plaintext/ciphertext images. The definition of NPCR is as follows:

$$\mathrm{NPCR} = \frac{\sum_{i=0}^{W}\sum_{j=0}^{H} D(i,j)}{W \times H} \tag{11}$$

where W and H represent the width and height of the image, respectively. T1 and T2 represent a ciphertext image obtained by encrypting two different plaintext images, respectively. If $T_1(i,j) = T_2(i,j), D(i,j) = 1$. If $T_1(i,j) \neq T_2(i,j), D(i,j) = 0$. In the experiment, there is only about 1% different pixels between these two plaintext images. Both the original plaintext image and the plaintext image with 1% pixel value changed are input to the proposed encryption model. Then, the NPCR is used to compare the differences between these two encrypted images. The calculated average NPCR value is 94.21%, which means that the information of the plaintext image is well diffused into the ciphertext image. Since DeepEDN has good diffusion performance and is highly sensitive to the plaintext, it achieves a good performance to resist the chosen plaintext attack like the differential attack.

4) Chosen Ciphertext Attack: In this type of attack, the attacker can access the decryption device, choose a string of ciphertext, and construct its corresponding plaintext string.

Since the structure of our decryption model is exactly the same as the encryption model, the experiment for the chosen ciphertext attack is similar to that in chosen plaintext attack. In this experiment, the input of the decryption network is the ciphertext image and the NPCR is used to calculate the difference between two decrypted images. According to the experiment, it is found that when the input ciphertext image changes slightly (just 1% pixels changed), the average NPCR value between two decrypted images is 94.87%. It means that if the input ciphertext image changes slightly, the decrypted image will change dramatically. This demonstrates that the proposed algorithm has good diffusion performance and is also highly sensitive to the ciphertext. It is effective to resist the chosen ciphertext attack.

V. Experimental Results

A. Performance of Encryption and Decryption

In addition to the chest X-ray data set, we further conducted experiments on brain MRI images and ultrasound images in order to evaluate the effectiveness of our proposed method on more medical imaging devices. The brain MRI images come from the BRATS 2015 [49], [50] data set. This data set includes brain images of 274 patients, and there are four different modal T1, T1c, T2, and Flair for each patient. The ultrasound data set is collected for the never segmentation

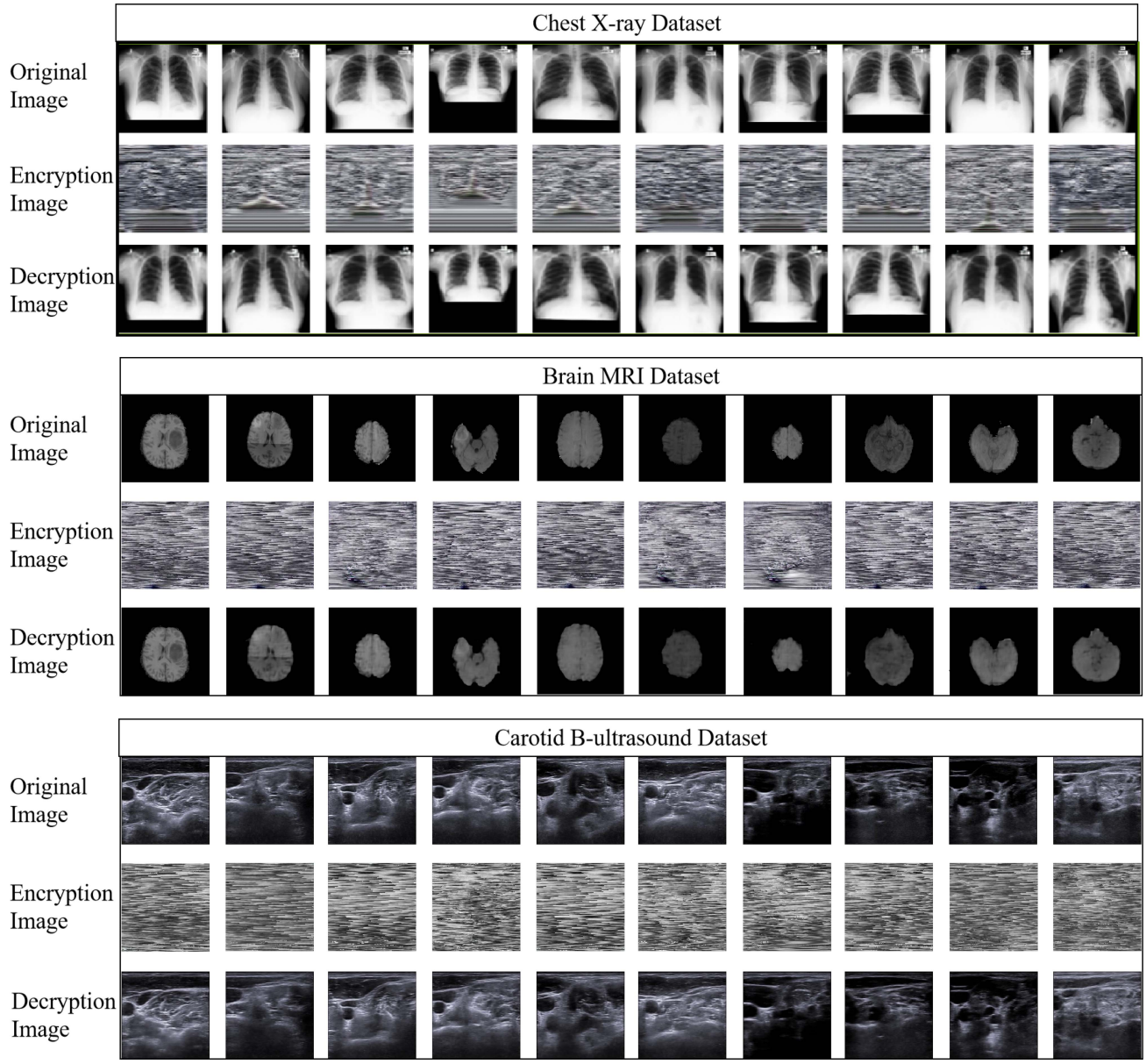


Fig. 9. Encryption and decryption performance of the proposed method.

which consists of 1055 ultrasound images. As shown in Fig. 9, the results of the proposed method for medical image encryption and decryption are presented in a visual way. It can be found that the proposed method also achieves a good encryption performance on these data sets.

It can be seen that the ciphertext image generated by the encryption network G is totally different from the original medical image and the pathology information cannot be observed. In addition, the image in the third row is decrypted from the encrypted one through the decryption network F, which can recover the detailed information of the original image and restore it to the original one. In order to evaluate the effectiveness of the decryption network, the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) are employed as evaluation metrics.

The quantitative measure of the decryption error is PSNR, which is based on the root MSE (RMSE) between the decrypted data and ground truth. It can be represented as

$$\text{PSNR} = 20 \log_{10} \frac{255}{\text{RMSE}}. \tag{12}$$

To further evaluate the performance of encryption and decryption, the SSIM is used as another metric

$$\text{SSIM}(x, y) = [l(x, y)]^{\alpha} [c(x, y)]^{\beta} [s(x, y)]^{\gamma} \tag{13}$$

where $l(x, y)$ is the brightness comparison, $c(x, y)$ is the contrast comparison, and $s(x, y)$ is the structure comparison. The closer the SSIM is to 1, the more resemblance the two images are. If this value approaches to 0, the two images are completely different. In an ideal case, the SSIM between the encrypted image and the original image is equal to 0, and the SSIM between the decrypted image and the original image is equal to 1. As shown in the second and third rows of Table VI, the SSIM between the encrypted image and the original image

TABLE VI
EVALUATION OF THE SSIM AND PSNR

Image Id	1	2	3	4	5	6	7	8	9	10
SSIM(Encrypted)	0.01	0.02	0.01	0.01	0.02	0.02	0.01	0.02	0.01	0.01
SSIM(Decrypted)	0.93	0.88	0.90	0.94	0.93	0.91	0.91	0.93	0.91	0.89
SSIM(2X)	0.90	0.92	0.90	0.92	0.89	0.91	0.88	0.90	0.91	0.90
PSNR	37.43	35.34	36.01	38.03	35.76	35.87	36.13	37.17	35.88	35.74
PSNR(2X)	35.48	35.74	35.03	35.28	34.87	36.73	34.75	34.61	36.17	34.80

is close to 0, and the SSIM between the decrypted image and the original image is close to 1.

For most of the medical image processing tasks, the image can be compressed to the one-half size of the original one to reduce the storage consumption and does not affect the doctor's diagnosis. In order to ensure that the decrypted image does not affect the doctor's diagnosis, the performance of the reconstructed image decrypted by the decryption network is also compared with the one-half compressed image. According to the experiment, it is demonstrated that the performance of the reconstructed image is equivalent to that through directly compressing the original image to half and then restoring it. In Table VI, from lines 3 to 6, 2× means that the original image is compressed to one half and then restored. At this level, humans can accurately identify the patient's organ contours and bone information from reconstructed images.

B. Performance of ROI-Mining Network

The direct extraction of interested information under ciphertext conditions is of great significance for medical image security and also for the data mining with privacy protection. The proposed ROI-mining network can segment the patient's interested organ tissue from the ciphertext image without decrypting the image first. The proposed network has the ability to realize data mining from the privacy environment by extracting the ROI from the encrypted image directly. In order to evaluate the proposed ROI-mining network, the well-known evaluation metric Dice score is adopted in here and is defined as

$$\text{Dice}(\text{GT}, \text{AT}) = \frac{\text{GT} \cap \text{AT}}{(|\text{GT}| + |\text{AT}|)/2}. \tag{14}$$

The GT represents the ground truth and the AT represents the model predictions. Fig. 10 shows the performance of the proposed ROI-mining network on the patient's left lung. It can be clearly seen that the prediction (gray ones) obtained from the model is almost the same as the ground truth (red ones). In addition, the original medical images are also adopted as the experiment data for training the same ROI-mining network, which is mainly used as the comparison. Under the same training conditions, the DICE of the segmentation network for plaintext is 0.967 while the DICE of the segmentation network for the ciphertext image is 0.962. It can be proven that the ROI-mining network can achieve a good segmentation performance on both plaintext and ciphertext images.

As mentioned before, the privacy keys of the network are totally different when training the network at different times even if all the conditions are the same. Therefore, the attacker cannot obtain the same ROI-mining network even if employing

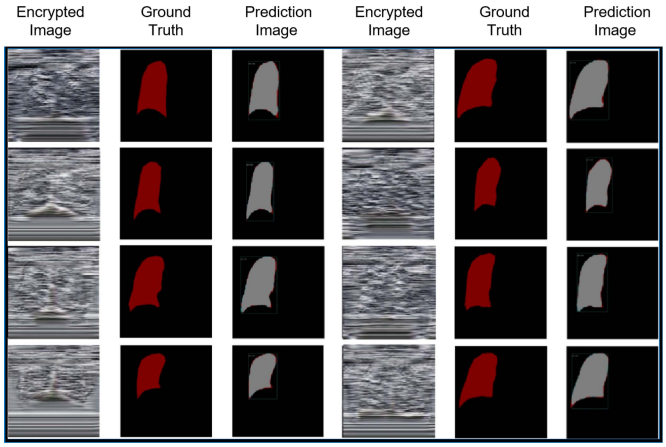


Fig. 10. Performance of ROI-mining network.

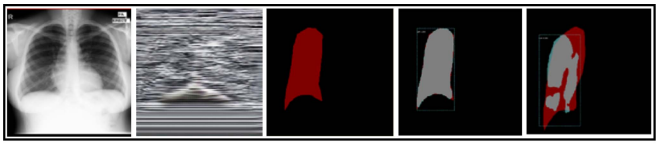


Fig. 11. Attack experiment for the proposed ROI-mining network.

the same ciphertext image for training. The experiment can be found in Fig. 11. In this experiment, Fig. 11(a) is the original image and Fig. 11(b) is the ciphertext image of Fig. 11(a). Fig. 11(c) is the ground truth for the right lung segmentation. Fig. 11(d) is the correct extraction result segmented by the ROI-mining network. Fig. 11(e) is the error extraction result segmented by the attacker.

C. Efficiency

To evaluate the efficiency of the proposed network, the running speed of encryption and decryption process on different resolution medical images is evaluated. For 256×256 resolution, the proposed network can encrypt or decrypt 14.28 medical images per second, while the speed is 3.65 images/s for encrypting or decrypting 512×512 resolution image. This encryption/decryption speed can basically meet the efficiency requirement in clinical practice. In addition, some chaotic encryption algorithms have been adopted as the comparison method for evaluating efficiency. For instance, Zhou *et al.* [36] introduced a simple chaotic system, which employs a combination of two existing 1-D chaotic maps (seed maps). Liao *et al.* [37] introduced a novel image encryption algorithm based on self-adaptive wave transmission. Wu *et al.* [39]

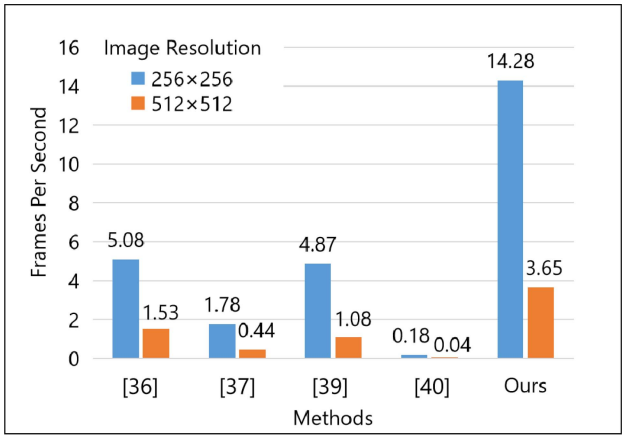


Fig. 12. Efficiency comparison between our method and other existing methods.

introduced a wheel-switch chaotic system for image encryption. In [40], the proposed method first adopts the 2-D logistic map with complicated basin structures and attractors for image encryption. This method can encrypt an intelligible image into a random-like one both from the point of view of the statistical and the human visual system.

Fig. 12 shows the comparison among the aforementioned five chaotic encryption algorithms and the proposed method. The FPS represents the number of images that can be encrypted/decrypted in 1 s. It can be found that our methods achieve the fastest encryption speed both on 512×512 resolution and 256×256 resolution images. Although the number of keys in our method is greater than the number of keys used in chaotic encryption methods, the processing time of our method is still with higher efficiency.

Compared with the block cipher, the length of the ciphertext encrypted by our proposed method is equal to the plaintext. But in some block ciphers, the length of the ciphertext is longer than the length of the plaintext, which causes storage overhead. In addition, we evaluated the efficiency of block ciphers on 512×512 resolution, where DES encrypts an image in 0.79 s and AES encrypts an image in 0.54 s. As seen in Fig. 12, our method only takes 0.27 s for each 512×512 resolution image. It indicates that the proposed method is with high efficiency.

VI. Conclusion

In this article, a novel medical image encryption and decryption method (namely, DeepEDN) is proposed by leveraging deep learning techniques, which is one of the early attempts to adopt the concept of "deep learning" for medical image encryption. The Cycle-GAN network is adopted as the learning network to encrypt and decrypt the medical image. A target domain is used to guide the learning model in the encryption process. The reconstruction network can decrypt the encrypted image to the original image (plaintext). Moreover, an ROI-mining network is proposed to directly extract the ROI from the encrypted medical image, with which DeepEDN can segment the interested organ or tissue in the ciphertext environment without decrypting the medical image. We conduct experiments on the chest X-ray data sets and the results show that the proposed algorithm can protect the medical image with a high-security level and can encrypt/decrypt the image in a more efficient way, compared with state-of-the-art medical image encryption methods.

References

[1] A. Gatouillat, Y. Badr, B. Massot, and E. Sejdić, "Internet of Medical Things: A review of recent contributions dealing with cyber-physical systems in medicine," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 3810–3822, Oct. 2018.

[2] N. Zhang, P. Yang, J. Ren, D. Chen, L. Yu, and X. Shen, "Synergy of big data and 5G wireless networks: Opportunities, approaches, and challenges," *IEEE Wireless Commun.*, vol. 25, no. 1, pp. 12–18, Feb. 2018.

[3] D. Chen *et al.*, "S2M: A lightweight acoustic fingerprints-based wireless device authentication protocol," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 88–100, Feb. 2017.

[4] B. Liu and H. Huang, "Picture archiving and communication systems and electronic medical records for the healthcare enterprise," in *Biomedical Information Technology*. New York, NY, USA: Academic, 2020, pp. 105–164.

[5] K. Zhang, J. Ni, K. Yang, X. Liang, J. Ren, and X. S. Shen, "Security and privacy in smart city applications: Challenges and solutions," *IEEE Commun. Mag.*, vol. 55, no. 1, pp. 122–129, Jan. 2017.

[6] D. Chen, N. Zhang, R. Lu, N. Cheng, K. Zhang, and Z. Qin, "Channel precoding based message authentication in wireless networks: Challenges and solution," *IEEE Netw.*, vol. 33, no. 1, pp. 99–105, Jan./Feb. 2018.

[7] D. Chen, N. Zhang, N. Cheng, K. Zhang, Z. Qin, and X. Shen, "Physical layer based message authentication with secure channel codes," *IEEE Trans. Depend. Secure Comput.*, early access, Jun. 12, 2018, doi: 10.1109/TDSC.2018.2846258.

[8] I. Natgunanathan, A. Mehmood, Y. Xiang, L. Gao, and S. Yu, "Location privacy protection in smart health care system," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3055–3069, Apr. 2019, doi: 10.1109/JIOT.2018.2878917.

[9] Y. Zhang, W. Liu, S. Cao, Z. Zhai, X. Nie, and W. Dai, "Digital image encryption algorithm based on chaos and improved DES," in *Proc. IEEE Int. Conf. Syst. Man Cybern.*, Mar. 2009, pp. 474–479.

[10] K. Chang, Y. Chen, C. Hsieh, C. Huang, and C. Chang, "Embedded a low area 32-bit AES for image encryption/decryption application," in *Proc. IEEE Int. Symp. Circuits Syst.*, Apr. 2009, pp. 1922–1925.

[11] M. Preishuber, T. Hutter, S. Katzenbeisser, and A. Uhl, "Depreciating motivation and empirical security analysis of chaos-based image and video encryption," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 9, pp. 2137–2150, Sep. 2018.

[12] Y. LeCun *et al.*, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.*, vol. 1, no. 4, pp. 541–551, Dec. 1989.

[13] L. Ale, N. Zhang, H. Wu, D. Chen, and T. Han, "Online proactive caching in mobile edge computing using bidirectional deep recurrent neural network," *IEEE Internet Things J.*, vol. 6, no. 3, pp. 5520–5530, Jun. 2019.

[14] H. Chen, Z. Qin, Y. Ding, L. Tian, and Z. Qin, "Brain tumor segmentation with deep convolutional symmetric neural network," *Neurocomputing*, vol. 392, pp. 305–313, Jun. 2020, doi: 10.1016/j.neucom.2019.01.111.

[15] Y. Ding, C. Luo, C. Li, T. Lan, and Z. Qin, "High-order correlation detecting in features for diagnosis of Alzheimer's disease and mild cognitive impairment," *Biomed. Signal Process. Control*, vol. 53, Sep. 2019, Art. no. 101564.

[16] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[17] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, Feb. 2016.

[18] B. Xiao, G. Ou, H. Tang, X. Bi, and W. Li, "Multi-focus image fusion by Hessian matrix based decomposition," *IEEE Trans. Multimedia*, vol. 22, no. 2, pp. 285–297, Feb. 2020.

[19] B. Xiao, K. Wang, X. Bi, W. Li, and J. Han, "2D-LBP: An enhanced local binary feature for texture image classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 9, pp. 2796–2808, Sep. 2019.

[20] H. Tang, B. Xiao, W. Li, and G. Wang, "Pixel convolutional neural network for multi-focus image fusion," *Inf. Sci.*, vols. 433–434, pp. 125–141, Sep. 2018, doi: 10.1016/j.ins.2017.12.043.
[21] B. Xiao *et al.*, "Follow the sound of children's heart: A deep-learning-based computer-aided pediatric CHDs diagnosis system," *IEEE Internet Things J.*, vol. 7, no. 3, pp. 1994–2004, Mar. 2020.
[22] P. Isola, J. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE CVPR*, Jun. 2017, pp. 5967–5976.
[23] A. Cherian and A. Sullivan, "Sem-GAN: Semantically-consistent image-to-image translation," in *Proc. IEEE WACV*, Oct. 2019, pp. 1797–1806.
[24] J. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE ICCV*, Oct. 2017, pp. 2242–2251.
[25] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.
[26] J. Lima and E. Neto, *Audio Encryption Based on the Cosine Number Transform*. New York, NY, USA: Kluwer, Oct. 2016. doi: 10.1007/s11042-015-2755-6.
[27] Q. N. Natsheh, B. Li, and A. G. Gale, "Security of multi-frame DICOM images using XOR encryption approach," *Procedia Comput. Sci.*, vol. 90, pp. 175–181, Jul. 2016.
[28] M. Mukhedkar, P. Powar, and P. Gaikwad, "Secure non real time image encryption algorithm development using cryptography & steganography," in *Proc. IEEE INDICON*, Jul. 2015, pp. 1–6.
[29] A. Kanso and M. Ghebleh, "An efficient and robust image encryption scheme for medical applications," *Commun. Nonlinear Sci. Numer. Simulat.*, vol. 24, pp. 98–116, Jan. 2015.
[30] C. Fu *et al.*, "An efficient and secure medical image protection scheme based on chaotic maps," *Comput. Biol. Med.*, vol. 43, no. 8, pp. 1000–1010, May 2013.
[31] W. Yu, C. Chi, X. Wei, and X. Yang, "Image encryption algorithm based on high-dimensional chaotic systems," in *Proc. Int. Conf. Intell. Control Inf. Process.*, Nov. 2010, pp. 463–467.
[32] I. J. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. NIPS*, Dec. 2015, pp. 2672–2680.
[33] J. Bao, D. Chen, F. Wen, H. Li, and G. Hua, "CVAE-GAN: Fine-grained image generation through asymmetric training," in *Proc. IEEE ICCV*, Oct. 2017, pp. 2764–2773.
[34] Y. Li and L. Shen, "cC-GAN: A robust transfer-learning framework for HEp-2 specimen image segmentation," *IEEE Access*, vol. 6, pp. 14048–14058, 2018.
[35] W. Liu, X. Liu, H. Ma, and P. Cheng, "Beyond human-level license plate super-resolution with progressive vehicle search and domain priori GAN," in *Proc. 25th ACM Int. Conf. Multimedia*, Oct. 2017, pp. 1618–1626.
[36] Y. Zhou, L. Bao, and C. L. P. Chen, "A new 1D chaotic system for image encryption," *Signal Process.*, vol. 97, pp. 172–182, Oct. 2014.
[37] X. Liao, S. Lai, and Q. Zhou, "A novel image encryption algorithm based on self-adaptive wave transmission," *Signal Process.*, vol. 90, pp. 2714–2722, Mar. 2010.
[38] Z. Hua, Y. Zhou, C. Pun, and C. L. P. Chen, "2D Sine Logistic modulation map for image encryption," *Inf. Sci.*, vol. 297, pp. 80–94, Feb. 2015.
[39] Y. Wu, J. Noonan, and S. Agaian, "A wheel-switch chaotic system for image encryption," in *Proc. Int. Conf. Syst. Sci. Eng.*, May 2011, pp. 23–27.
[40] Y. Wu, G. Yang, H. Jin, and J. Noonan, "Image encryption using the two-dimensional logistic chaotic map," *J. Electron. Imag.*, vol. 21, pp. 3014–3022, Jan. 2012.
[41] N. Wang, W. Zha, J. Li, and X. Gao, "Back projection: An effective postprocessing method for GAN-based face sketch synthesis," *Pattern Recognit. Lett.*, vol. 107, pp. 59–65, Jun. 2018.
[42] Z. Yi, H. Zhang, P. Tan, and M. Gong, "DualGAN: Unsupervised dual learning for image-to-image translation," in *Proc. IEEE ICCV*, Mar. 2017, pp. 2868–2876.
[43] D. Arroyo *et al.*, "On the security of a new image encryption scheme based on chaotic map lattices," *Chaos*, vol. 18, pp. 1–8, Aug. 2008.
[44] F. Jiang *et al.*, "Deep learning based multi-channel intelligent attack detection for data security," *IEEE Trans. Sustain. Comput.*, vol. 5, no. 2, pp. 204–212, Apr.–Jun. 2020, doi: 10.1109/TSUSC.2018.2793284.
[45] D. Chen *et al.*, "An LDPC code based physical layer message authentication scheme with prefect security," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 748–761, Apr. 2018.
[46] A. Ferdowsi and W. Saad, "Deep learning for signal authentication and security in massive Internet of Things systems," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1371–1387, Feb. 2019.
[47] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE CVPR*, Sep. 2016, pp. 770–778.
[48] S. Jaeger *et al.*, "Two public chest X-ray datasets for computer-aided screening of pulmonary diseases," *Quant. Imag. Med. Surg.*, vol. 4, pp. 475–477, Dec. 2014.
[49] B. H. Menze *et al.*, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: 10.1109/TMI.2014.2377694.
[50] M. Kistler, S. Bonaretti, M. Pfahrer, R. Niklaus, and P. Büchler, "The virtual skeleton database: An open access repository for biomedical research and collaboration," *J. Med. Internet Res.*, vol. 15, no. 11, p. e245, Oct. 2013, doi: 10.2196/jmir.2930.

Yi Ding (Member, IEEE) received the B.S. degree in software engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2008, and the Ph.D. degree in computer science from the Dublin Institute of Technology, Dublin, Ireland, in 2012.

From 2012 to 2016, he was a Postdoctoral Research Fellow with UESTC, where he has been an Associate Professor with the School of Information and Software Engineering. He is also a Researcher with the Institute of Electronic and Information Engineering, UESTC. His research interested include machine learning, deep learning, medical image processing, and computer-aided diagnosis.

Guozheng Wu received the B.S. degree in information management and decision science from the University of Science and Technology of China, Chengdu, China, in 1997, the M.S. degree in public management from Tsinghua University, Beijing, China, in 2002, and the Ph.D. degree from the University of Electronic Science and Technology of China in 2010.

He is currently a Staff with the National Natural Science Foundation of China, Beijing, China. His research interests include computer networking, information security, applied cryptography, information management, software engineering, and machine learning and its applications.

Dajiang Chen (Member, IEEE) received the Ph.D. degree in information and communication engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2014.

He is currently an Associate Professor with the School of Information and Software Engineering, UESTC. He was a Postdoctoral Fellow with the Broadband Communications Research Group, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, from 2015 to 2017. His current research interests include wireless security, physical layer security, secure channel coding, and machine learning and its applications in wireless network security and wireless communications.

Dr. Chen served as the Workshop Chair for BDEC-SmartCity'19 in conjunction with IEEE WiMob 2019 and the organizer for IoT track in conjunction with EAI CollaborateCom 2020. He also serves/served as a Technical Program Committee Member for IEEE Globecom, IEEE ICC, IEEE VTC, IEEE WPMC, and IEEE WF-5G.

Ning Zhang (Senior Member, IEEE) received the Ph.D. degree from the University of Waterloo, Waterloo, ON, Canada, in 2015.

He was a Postdoctoral Research Fellow with the University of Waterloo and the University of Toronto, Toronto, ON, Canada. He is an Associate Professor with the University of Windsor, Windsor, ON, Canada.

Dr. Zhang received the Best Paper Awards from IEEE Globecom in 2014, IEEE WCSP in 2015, the *Journal of Communications and Information Networks* in 2018, IEEE ICC in 2019, IEEE Technical Committee on Transmission Access and Optical Systems in 2019, and IEEE ICCC in 2019. He also serves/served as a track chair for several international conferences and a co-chair for several international workshops. He serves as an Associate Editor for the IEEE INTERNET OF THINGS JOURNAL, the IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING, IEEE ACCESS, and *IET Communications, and Vehicular Communications* (Elsevier), and a Guest Editor of several international journals, such as IEEE WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, and the IEEE TRANSACTIONS ON COGNITIVE COMMUNICATIONS AND NETWORKING.

Linpeng Gong received the B.E. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2016, where he is currently pursuing the M.S. degree.

His current research interests include deep learning, computer vision, digital image security, and Internet of Things.

Mingsheng Cao (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, 2008, 2011, and 2019, respectively.

He currently is a Lecturer with the Network and Data Security Key Laboratory of Sichuan Province, UESTC. His research interests include network security, pervasive computing, and machine learning and its applications in network security and pervasive computing.

Zhiguang Qin (Member, IEEE) received the B.E. degree from Yibin University, Yibin, China, in 1980, the M.S. degree from Xiangtan University, Xiangtan, China, in 1988, and the Ph.D. degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1996.

He was the Dean of the School of Software, UESTC, where he is the Director of the Key Laboratory of New Computer Application Technology and the Director of UESTC-IBM Technology Center. His research interests include wireless sensor networks, mobile social networks, information security, applied cryptography, information management, intelligent traffic, electronic commerce, distribution, and middleware.

Dr. Qin served as the General Co-Chair for WASA 2011 and Bigcom 2017.