EDG-Net: Encryption and Decryption based Gan-attention Network for CT images in the Internet of Medical Things and Telemedicine

Authors: chen, Dr. kai, Ms. Hongyu Gao, Gao, Dr. Yuan, Nie, Dr. Mu, Li, Yuchen, Xie, Shipeng, Yan, Prof. Wei, Chen, Prof. Yang, Dr. Yan Xi, chen, Dr. kai

Date: 2025-07-02T11:48:19+00:00

Abstract

CT images provide medical practitioners with a scientific and intuitive rationale for the diagnosis of clinical diseases. The Internet of Medical Things (IoMT) and telemedicine facilitate the preservation, transmission, and application of medical data, driving the sharing of medical data, especially medical images. Encryption and decryption of CT images distributed in the IoMT and telemedicine are becoming critical because they contain a large amount of private patient-sensitive information and are vulnerable to third-party attacks, resulting in information exposure and privacy leakage. In this paper, we propose an Encryption and Decryption based Gan-attention network (EDG-Net) for CT images in the IoMT and telemedicine. EDG-Net consists of a generator, two discriminators, a domain transfer of attention, and adaptive normalization. In addition, a double encryption and decryption strategy is introduced by EDG-Net to effectively improve the security of the ciphertext image and the fidelity of the decrypted plaintext image. Specifically, during the encryption or decryption phase, the generator transforms the CT images mutually in the plaintext and ciphertext domains. Two discriminators to identify and modify the differences between these two domain transformations, especially to improve the accuracy of the reconstruction during decryption. The parameters of the trained encryption and decryption network are considered as the secret keys of encryption and decryption. Qualitative and quantitative analysis of public and private datasets demonstrates the superior performance of EDG-Net regarding encryption security and robustness, as well as decryption accuracy.

Full Text

Preamble

EDG-Net: Encryption and Decryption based GAN-Attention Network for CT Images in the Internet of Medical Things and Telemedicine

Kai Chen^a, Hongyu Gao^d, Yuan Gao^e, Mu Nie^f, Yuchen Li^b, Shipeng Xie^a,*, Wei Yan^c,**, Yang Chen^b,a,*, Yan Xi^g

^{aSchool} of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
^{bSchool} of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China, and Laboratory of Image Science and Technology, the School of Computer Science and Engineering, Southeast University, Nanjing 210096, China
^{cDepartment} of Neurosurgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, Jiangsu, China
^{dDepartment} of Pharmacy, Jinling Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing 210096, China
^{eThe} Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan), and Qilu University of Technology (Shandong Academy of Sciences), Jinan 250013, China
^{fSchool} of Integrated Circuits, Anhui Polytechnic University, Wuhu 241000, China
^{gThe} Shanghai First-Imaging Information Technology Co., Ltd., Shanghai 201315, China

Abstract

CT images provide medical practitioners with a scientific and intuitive rationale for the diagnosis of clinical diseases. The Internet of Medical Things (IoMT) and telemedicine facilitate the preservation, transmission, and application of medical data, driving the sharing of medical data, especially medical images. Encryption and decryption of CT images distributed in the IoMT and telemedicine are becoming critical because they contain a large amount of private patient-sensitive information and are vulnerable to third-party attacks, resulting in information exposure and privacy leakage. In this paper, we propose an Encryption and Decryption based GAN-attention Network (EDG-Net) for CT images in the IoMT and telemedicine. EDG-Net consists of a generator, two discriminators, a domain transfer of attention, and adaptive normalization. In addition, a double encryption and decryption strategy is introduced by EDG-Net to effectively improve the security of the ciphertext image and the fidelity of the decrypted plaintext image. Specifically, during the encryption or decryption phase, the generator transforms the CT images mutually in the plaintext and ciphertext domains. Two discriminators identify and modify the differences between these two domain transformations, especially to improve the accuracy of the reconstruction during decryption. The parameters of the trained encryption and decryption network are considered as the secret keys of encryption and decryption. Qualitative and quantitative analysis of public and private datasets demonstrates the superior performance of EDG-Net regarding encryption security and robustness, as well as decryption accuracy.

Keywords: Computed tomography (CT), CT imaging techniques and applications, Internet of Medical Things (IoMT), telemedicine, deep learning

1. Introduction

IoMT and telemedicine have become a rising demand and emerging technology in the healthcare industry with applications in remote diagnosis, early diagnosis of diseases, emergency advice from doctors, and collection of long-term data from patients for improved health monitoring [1]. Computed Tomography (CT) images disseminated in the Internet of Medical Things (IoMT) are widely used in the diagnosis of various diseases by utilizing X-ray beams to scan the human body in layers and computer processing to produce detailed images of the internal structure of the body, which are featured by fast imaging speed and high spatial resolution [2, 3, 4, 5]. As shown in Fig. 1, the main circulation of medical images saved and transmitted is in the hospital intranet and in the hospital extranet. When a patient is scanned by CT, MR, ultrasound, and other common medical imaging devices, the medical images are stored in the picture archiving and communication system (PACS) in the hospital intranet. When a physician diagnoses a patient, PACS retrieves the required images from the digital hospital intranet database and transmits the images to the physician's workstation, which works with the patient information from the hospital information system (HIS). For remote diagnosis, medical images of patients examined at the primary hospital need to be shared with specialists at partner hospitals in different regions. Whether in PACS and HIS on hospital intranets or in the distribution of remote diagnostics based on the IoMT, medical images still have some critical security issues when storing, transmitting, and reviewing medical images that retain sensitive private patient information. If an internal or external attacker has the ability to compromise PACS, HIS, and IoMT systems, it can easily lead to the leakage of patients' private information. Therefore, encryption protection of CT images is a growing concern for researchers [6, 7, 8, 9].

With the growing research, medical image encryption can be categorized into traditional medical image encryption, chaos-based medical image encryption, and deep learning-based medical image encryption. Traditional medical image encryption algorithms are mainly based on pixel disambiguation in the spatial domain and encryption techniques based on the transform domain. Pixel disambiguation based on the null domain of medical images leaves the pixels of the digital image unchanged, making it easy for third parties to find disambiguation patterns on statistical analysis and steal secrets. Transform domain-based medical image encryption is the transformation of a null domain image into the transform domain, which is divided into different coefficient matrices based on different parameters and frequencies, but transform domain-based medical image encryption can cause distortion of the image if it is not combined with spatial domain based encrypted images.

In recent years, chaos-based medical image encryption exploits the chaotic sequences generated by chaotic systems to disrupt and diffuse the plaintext image at the pixel level, thus changing the original pixels of the plaintext for image encryption [10, 11, 12, 13, 14]. Deng encrypted medical images by combining two different one-dimensional chaos to form composite chaos, and then dislocating or processing the image ranks [15]. Sweldens et al. proposed an improved algorithm of chaotic image encryption based on the logistic mapping. The characteristics of the algorithm are that the security of the encryption algorithm is improved by expanding the encryption space, and the shortcoming is that the encryption and decryption period is long [16]. Maure et al. used logistic mapping to generate a pseudo-random sequence to generate a pseudo-random matrix based on a pseudo-random matrix construction pattern to achieve image position dislocation and then used operations or dissimilarity to achieve the purpose of spreading pixels [17].

In addition, DNA coding and compression-aware techniques are also applied to chaotic image encryption [18, 19]. In 2018, Wu et al. proposed a DNA-based chaotic image encryption method. The use of DNA encoding greatly enhances the effectiveness of image pixel value dislocation and diffusion, but the complex encoding rules lead to high time complexity and slow operation, which is not conducive to the real-time transmission of multiple images [20]. Chai et al. used a compression-aware technique to compress the plaintext image at a certain compression rate instead of the noise-like ciphertext image obtained by traditional encryption methods, which is equivalent to visually encrypting it at the same time [21].

With the empowerment of deep learning technology in image processing and computer vision, the development of digital image encryption based on deep learning has been expedited. Dong et al. employed a binary tree neural network to exchange keys in groups and generate a cipher stream to encrypt and protect the message [22]. Fang et al. generated encryption keys using a deep convolutional generative adversarial network in conjunction with a four-dimensional hyperchaotic system, and then scrambled and diffuse encryption of the image [23]. Ding et al. improved the security of medical image encryption by generating keys for medical image encryption via generative adversarial networks [24]. Ding et al. encrypted medical images by directly converting plaintext medical images into noisy images with ciphertext style through a recurrent generative adversarial network [25].

Medical image encryption in IoMT and telemedicine faces four core challenges: traditional encryption's security-fidelity trade-offs, chaotic-inspired methods' efficiency limitations, deep learning encryption's performance dilemmas, and

IoMT-specific cross-domain risks. Traditional spatial-domain pixel permutation methods fail to disrupt statistical patterns, making them susceptible to frequency analysis attacks [26], while transform-domain encryption causes image distortion without spatial processing, violating medical fidelity requirements [27]. Chaotic systems suffer from limited key space and parameter sensitivity, with DNA-based encryption incurring time complexity, failing to meet real-time IoMT transmission demands [28, 29]. Traditional methods fail to adapt to fluctuating bandwidth, and feature variability across medical devices hindering model generalization [30] and reducing cross-device encryption effectiveness [31].

In this paper, we make the following contributions:

- A network consisting of a generator, two discriminators, a domain transfer of attention, and adaptive normalization for medical image encryption and decryption called EDG-Net has been developed by employing image-style translation deep learning techniques for medical image encryption.
- A double encryption strategy (DEDS) is introduced by EDG-Net. During the encryption phase, DEDS facilitates the fine-grained transformation from plaintext to ciphertext images and strengthens the security of ciphertext images; during the decryption phase, DEDS enables the receiver to accurately decrypt the ciphertext image into a plaintext image to ensure the fidelity of medical images.
- A novel secret key generation method for encryption and decryption of medical images is developed. We consider the parameters of the trained EDG-Net as the secret key for encryption and decryption. Since EDG-Net is a highly nonlinear deep learning network with random initials, the parameters of the trained EDG-Net change with each training of the network, so the capacity space of the encryption and decryption secret keys is large and sensitive, which is a one-time pad.
- A multi-level consistency loss function is introduced for encryption and decryption of medical images, which can effectively combine the attention module to recover the key information and tiny structures of ciphertext medical images adaptively during the decryption process and improve the accuracy of decryption.
- A private dataset of medical images for encryption and decryption is collected. Three state-of-the-art encryption methods are employed to encrypt the collected plaintext medical images into ciphertext image labels. Qualitative and quantitative results demonstrate that the dataset is useful for training, validating, and testing deep learning-based encryption algorithms for medical images.

The rest of this paper is organized as follows. Section II gives the mathematical model of EDG-Net and the details of the neural network architecture. The experimental result and analysis are performed in Section III. In Section IV, we will discuss some related issues and make conclusions.

2.1. One-step One-time Medical Image Double Encryption and Decryption Strategy

In this paper, we develop a one-step one-time medical image double encryption and decryption strategy as shown in Fig. 2. We are given a set of images for doctors to make clinical diagnoses in plaintext domain P and a different set of images that disrupt patient organ information in the ciphertext domain C. When encrypting, we want to train a mapping En: P → C such that the output $\hat{c}$ = En(p), p P, is indistinguishable from images c C by an adversary called Discriminator C trained to classify $\hat{c}$ apart from c. In theory, this objective can induce an output distribution over $\hat{c}$ that matches the empirical distribution pdata(c). The optimal G thereby translates the domain P to a domain $\hat{C}$ distributed identically to C. Meanwhile, we have a mapping De: C → P such that the output $\hat{p}$ = G(c), c C guarantee that an individual input x and reconstruction output y is paired in a correspondence De(En(p)) = c and En(De(c)) = p. In addition, we introduce two adversarial discriminators DC and DP, where DC aims to distinguish between images p P and translated images {G(c)}. In the same way, DP aims to discriminate between images c C and {G(p)}. The decryption strategy is exactly the same as the encryption strategy, only plaintext images and ciphertext images need to be swapped. The network architecture parameters of the encrypted network are taken as the secret key, and the receiver can reproduce the decrypted network by receiving the trained encrypted network parameters and inputting the cipher text into the decrypted network with the same architecture as the encrypted network to realize the decryption process.

2.2. The Detail of EDG-Net

2.2.1. Generator of EDG-Net

The generator employed during encryption and decryption has the same network architecture as shown in Fig. 3. The generator consists of an encoder Es (2 down-sampling and 4 residual blocks), an auxiliary classifier, and a decoder Ds (4 residual blocks and 2 up-sampling). During the downsampling phase, the input image x Xp increases the number of channels to 64 dimensions by a 7 × 7 convolution kernel with a step size of 1 and a padding of 3. Then the number of channels is increased to 256 dimensions by a 3 × 3 convolution kernel with step size 2 and a padding of 1, and the image size is reduced from 256 × 256 to 64 × 64. The downsampled feature maps passes through 4 same residual blocks, which are realized by a convolution with padding of 1, step size of 1 and convolution kernel of 3 × 3.

The encoded feature maps as shown in Fig. 4(a) Es(x) = {E1s … Ens} are subjected to global average pooling and global max pooling in the auxiliary classifier for plaintext domain and ciphertext domain binary classification, respectively. A weight Ws = {Ws1 … Wsn} is given to each channel of the encoded feature map by binary classification, and this weight determines the importance of the corresponding feature of this channel to realize the attention mechanism under the attention feature map Ea(x) = {Ea1, Ea2, Ea3 … Ean}, k = 1, 2, …, n. The DTA subjected to a multilayer perceptron to determine the α and β to be used for the adaptive normalization (AN). AN as shown in Fig. 4(b) combines the advantages of instance normalization (IN) and layer normalization (LN). IN assumes no channel correlation, it normalizes globally but does not preserve the content structure well, but IN ensures that the channels are not correlated and only normalizes the map itself, which well compensates for the shortcomings of LN. The features are finally fed to a decoder consisting of four residual blocks and two up-sampling to obtain the output.

2.2.2. Discriminators of EDG-Net

The proposed EDG-Net employs a global discriminator and a local discriminator as shown in Fig. 5 during the encryption and decryption processes. The global discriminator and the local discriminator consist of residual blocks, a class activation map, and a classifier. The difference between the global discriminator and the local discriminator is that the global discriminator performs two more convolutions of the features with a step size of 2, a padding of 1, and a convolution kernel of 4 × 4 to compress the features more deeply. The DAT is incorporated in both the local discriminator and the global discriminator. Although the DAT does not do the classification of plaintext and ciphertext domains, the inclusion of the attention module is beneficial for discriminating the authenticity of the images. The attention maps facilitate fine-tuning by focusing on the differences between the real and fake images in the target domain.

2.2.3. Loss Function

The multi-level consistency loss in EDG-Net is as follows:

Gp→c, Gc→p, p, c Dp, Dc, Dp, Dc $ $1 Llsgan + $ $2 Lcycle + $ $3 Lidentity + $ $4 Lcam,

Llsgan = Lp→c lsgan + Lc→p lsgan
Lcycle = Lp→c cycle + Lc→p cycle
Lidentity = Lp→c identity + Lc→p identity
Lcam = Lp→c cam + Lc→p cam

where Llsgan is adversarial loss function, Lcycle is cyclic loss function, Lidentity is identity loss function, and Lcam is class activation mapping loss. Lp→c lsgan, cycle, cam as an example of the encryption process (Lc→p lsgan, cycle, cam during decryption process calculated in exactly the same way). The adversarial loss Llsgan matches the distribution of the plaintext domain image to the distribution of the ciphertext domain image. The adversarial loss is designed as follows:

lsgan = (Ex Xc[(Dc(x))^2] + Ex Xp[(1 - Dc(Gp→c(x)))^2])

Generators in encryption and decryption employ a cyclic consistency constraint. This cyclic consistency constraint ensures that the ciphertext image can be successfully decrypted to the original plaintext domain after the plaintext image is encrypted into the ciphertext image. The cyclic consistency constraint is designed as follows:

cycle = Ex Xp[‖x - Gc→p(Gp→c(x))‖$_1$],

The generator in encryption and decryption utilizes an identity loss function that applies an identity consistency constraint to the generator. Given a plaintext image, it is guaranteed that the color distribution is similar after generating a ciphertext image using the generator. The identity loss function is designed as follows:

identity = Ex Xt[‖x - Gp→t(x)‖$_1$],

By exploiting the information from the auxiliary classifier, the attention mechanism is positively influenced and thus helps the generator and discriminator in gaining insight into what is the biggest difference between the plaintext and ciphertext domains. The CAM loss function is designed as follows:

cam = -(Ex Xp[log(p(x))] + Ex Xc[log(1 - p(x))])
cam = Ex Xc[(Dc(x))^2] + Ex Xp[(1 - Dc(Gp→c(x)))^2]

3.1. Datasets and Experiment Environment

A private dataset is collected to train and validate the outstanding performance of the proposed EDG-Net. The dataset contains 1639 digital medical images with a size of 256 × 256 from the Nanjing First Hospital, China, with the approval of the Institutional Review Board and patient consent forms. The number of images is increased to 6556 by using data enhancement with rotation of 90°, 180°, and 270°. Three state-of-the-art methods [32, 33, 34] are employed for the encryption of 6556 medical images. 19468 medical images are used as the training set for the private dataset, 100 images as the validation set, and 100 images as the test set.

The public dataset containing 150,000 medical images is employed to demonstrate the superior performance of the EDG-Net. The US National Institutes of Health provides this dataset, and all images are from 30,805 unique patients with an original image size of 1024 × 1024. The same three encryption methods and dataset generation methods applied to private datasets are also available for public datasets. 149800 medical images are used as the training set for the private dataset, 100 images as the validation set, and 100 images as the test set.

This study implemented all deep learning-based algorithms using the Paddle framework. The networks were trained and tested on a computer with configurations: CPU is Intel Core i9-9900KF @ 3.60GHz; GPU is NVIDIA RTX 3090 with 24 GB memory. The total epoch was set to 100 and 10,000 iterations in each epoch. The Adam algorithm with $ $1 = 0.5 and $ $2 = 0.999 was adopted to optimize the proposed EDG-Net. The batch size is set to 1 and the initial learning rate was set to 0.0001 and slowly decreased to 0.00001.

3.2. Results and Analysis of the Comparison Experiment

CCPL [35], CNN [36], CycleGAN [37], P2P [38], and UGA [39] were taken as comparison methods to verify the encryption and decryption performance of EDG-Net. The structural similarity index (SSIM), Peak Signal-to-Noise Ratio (PSNR), Inception Score (IS), Fréchet Inception Distance (FID), and Kernel Inception Distance (KID) were used as metrics to measure the effectiveness of EDG-Net encryption and decryption.

3.2.1. Results and Analysis of Encryption Results

Encryption results from comparison experiments on the public dataset and the private dataset are shown in Fig. 6. From Fig. 6, it can be observed that the ciphertext image of CNN and CCPL after medical image encryption cannot effectively protect the patient's tissue information. CYC_{GAN}, P2P, and UGA all provide some protection for patients' organizational information, while CYC_{GAN}'s cipher image style and distribution on public data differs significantly from the reference cipher image. The medical image encrypted by UGA is still able to see the chest rib information although the image style is similar to that of the cipher domain image, while the P2P and EDG-Net encrypted medical images provide full protection of the patient's chest tissue structure. The same phenomenon occurs for private dataset encryption results. Comparing the images encrypted by P2P and EDG-Net, the medical images encrypted by P2P are distributed with similar segmented stripes of pixels with some regularity. The image encrypted by EDG-Net does not have the regular segmented stripe-like pixel distribution like the image encrypted by P2P, and it most closely matches the style and pixel distribution of the reference ciphertext, while also maintaining a certain degree of differentiation from the style of the ciphertext, which makes it difficult to be locked and attacked by a third party.

The quantitative metrics on public and private datasets for encryption and decryption of the different methods in the comparison experiments were shown in Tab. 1. From the quantitative encryption results, EDG-Net obtained the highest PSNR and SSIM on both public and private datasets, which indicated that the medical images encrypted by EDG-Net had the smallest disparity and the highest similarity with the reference ciphertext medical images. EDG-Net achieved the lowest FID and KID on both public and private datasets, denoting that the distance between the ciphertext medical reference image and the ciphertext image encrypted by EDG-Net is the smallest, and the ciphertext generation model of EDG-Net has the best performance and the best encryption effect. The IS of EDG-Net encrypted medical images on public and private datasets are 1.43 and 1.67, respectively. This means that the KL dispersion of the category probability distributions of the cipher images generated by EDG-Net during the encryption process with the edge distributions of all the categories is the largest, and EDG-Net is outperforming all the cipher generation models in the comparison experiments.

3.2.2. Results and Analysis of Decryption Results

Fig. 7 illustrated the receiver receiving a medical image encrypted with a secret key decrypted from the public and private datasets. As shown in Fig. 7, the ciphertext images decrypted by CNN, CCPL, and CYC_{GAN} fail to effectively recover the tissue structure and organ details of the original medical images. The medical images decrypted by P2P, UGA, and EDG-Net maintain a high consistency in pixel distribution and image style with the original medical images. Compared with P2P, medical images encrypted and decrypted by EDG-Net have better restorability and fidelity in terms of details such as tissue structure boundaries, rib morphology, and so on.

To further compare the performance of P2P, UGA, and EDG-Net, the local regions of interest in Fig. 7 were picked for analysis as illustrated in Fig. 8. As illustrated in Fig. 8, the ROIs within the circles labeled in Fig. 8(c1)-(f3) after being encrypted and decrypted by the UGA on the public dataset show a significant change in the location information as compared to the other methods. As shown in Fig. 8(b3), the medical image after decryption by P2P over-sharpens the ROIs in the original medical image, which makes details such as blood vessels that are not obvious in the original image become overly clear. As illustrated in Fig. 8(d5), the images encrypted and decrypted by EDG-Net can more accurately recover the tiny structural information of the bones and blood vessels, while guaranteeing the same contrast and brightness as the original image with the same degree of similarity.

The quantitative metrics of decrypted images after each comparison method on public and private datasets in the comparison experiments are listed in Tab. 2. The results of this quantitative analysis also argue that the image quality of medical images encrypted and decrypted by P2P, UGA, and EDG-Net in Fig. 7 is superior to CNN and CCPL. The decrypted image of EDG-Net takes the maximum PSNR and SSIM on both the public dataset and on the private dataset. The IS of the decrypted images of P2P, UGA, and EDG-Net on both public datasets and private data is larger than that of CNN, CCPL, and CYC_{GAN}. Since KL dispersion ignores the distribution of real medical image data, FID and KID metrics are further introduced to demonstrate the capability of the decryption model. As with the IS metrics, in both FID and KID metrics, the images generated by P2P, UGA, and EDG-Net outperform CNN, CCPL, and CYC_{GAN}. Compared with EDG-Net, the FIDs of P2P and UGA on public and private datasets are about half of those of EDG-Net. EDG-Net has the optimal KID on both public and private datasets. It is indicated that during the decryption process, EDG-Net has the best plaintext generation effect, and the similarity between the decrypted plaintext and the original medical image is the strongest, and the decryption result of EDG-Net is the most accurate.

3.3. Results and Analysis of the Ablation Experiment

The ablation experiment is performed to investigate the effectiveness of the double encryption and decryption strategy (DEDS), domain transfer of attention (DTA), and adaptive normalization (AN) in EDG-Net. First, a deep-learning-based medical image encryption and decryption network (ED-Net) without DEDS, DTA, and AN is used as a baseline. Then, ED-Net+DEDS, which only introduces DEDS, is employed for double encryption and decryption of medical images. Next, ED-Net+AN with only AN incorporated was employed to demonstrate the role of AN in improving the accuracy of EDG-Net in transferring ciphertext pixels and plaintext pixels to each other, while assisting both the ciphertext generation model and the plaintext generation model in flexibly controlling the number of changes to the shapes and textures of features in both the ciphertext domain and the plaintext domain. Finally, ED-Net+DTA with DTA only is added to examine the ability of DTA to locate the position of key features in EDG-Net based on the classification results of the source and target domains to improve the security during the encryption phase and the accuracy during the decryption phase of EDG-Net.

Fig. 9 shows the qualitative analysis of each algorithm after encryption and decryption on public and private datasets. As can be seen from the encryption results in Fig. 9, there are some subtle differences in EDG-Net while maintaining a similar pixel style to that of the reference cipher image, but overall the encryption results seem to be better than the other methods in the ablation experiments. As shown in Fig. 9(c1)-(c6), the ciphertext image encrypted by ED-Net+DEDS can expose part of the tissue contour boundary information. The medical images encrypted in Fig. 9(d1)-(d6) can demonstrate the feature information of the ribs inside the chest, indicating that the separate introduction of AN on the baseline model does not fully provide secrecy to the plaintext. During the decryption stage, the decrypted images of ED-Net, ED-Net+DEDS, and ED-Net+AN deviated significantly from the original medical images. The decrypted images of ED-Net and ED-Net+DEDS can only decrypt fuzzy tissue shape information from the ciphertext image, which has little clinical significance for medical diagnosis at the receiving end. With the addition of AN, the decrypted image is much improved compared to ED-Net and ED-Net+DEDS, and can restore the organizational information of some key features, but the accuracy is still far from enough. ED-Net+DTA effectively decrypts global and local organizational information, with a lack of clarity recovery on the overall image. EDG-Net provides accurate decryption of medical images while maintaining secure encryption.

Tab. 3 illustrates the quantitative results of each method in the ablation experiments on the public dataset. The baseline ED-Net in the ablation experiments scored the lowest PSNR and SSIM during the encryption and decryption phase on both the public and private datasets. With the incorporation of DEDS, AN, and DTA individually, the PSNR and SSIM of the model are improved, which indicates that the quality of the encrypted ciphertext image and the quality of the decrypted plaintext image of the model are both improved. In terms of plaintext generation capability, with the introduction of DEDS, AN, and DTA, the plaintext generation modeling capabilities are improved during the decryption stage. It can be seen that due to the empowerment of AN and DTA, EDG-Net can effectively locate the key positions of the conversion between ciphertext pixels and plaintext pixels, and simultaneously can transform some medical image texture level and shape features to enhance the robustness of the encryption or decryption.

3.4. Security Analysis

3.4.1. Histogram Analysis of Encrypted and Decrypted Images

Histogram analysis is an analytical method used to describe the distribution of pixels in a digital image, which determines the pixel distribution of an image by counting the pixel grey values and the number of pixels per grey value. When encryption-protected medical images circulate in the public channel, third-party attackers can target the image information by comparing the histograms before and after encryption through histogram analysis. A robust encryption system should have a high degree of similarity in the histograms of different images encrypted. As shown in Fig. 10, the pixel distributions of the histograms of the two plaintext reference images (c1) and (d1) and the ciphertext histograms encrypted by each encryption method in the comparison experiments are uniformly distributed except for (a4), (b4), (a5), and (b5). This phenomenon indicates that CYC_{GAN} and P2P are vulnerable to attack by third parties through histogram analysis, and the encryption algorithms perform poorly in terms of reliability. Although the ciphertexts encrypted by CNN, CCPL, and UGA are in a uniform distribution, the encrypted ciphertexts perform poorly in qualitative evaluation, exposing part of the organization and failing to guarantee the security of the encryption algorithm. During the decryption phase, the pixel distribution of the plaintext images decrypted by each decryption algorithm differs significantly from that of the plaintext reference image except for P2P and EDG-Net. The histograms of (c5), (d5) and (c7), (d7) decrypted by P2P and EDG-Net are similar in shape to the plaintext reference images (c1) and (d1). This indicates that the decryption accuracy of P2P and EDG-Net is high, and the decrypted plaintext has some clinical diagnostic guidance. In summary, in the comparison experiments, P2P has some decryption accuracy but performs poorly in encryption security and reliability, and CNN, CCPL, and UGA perform poorly in security, reliability, and accuracy during both encryption and decryption phases.

3.4.2. Pixel Correlation Analysis of Encrypted and Decrypted Images

The pixel correlation of a digital image describes the relationship between neighboring pixels of an image. A proper medical image encryption algorithm encrypts a cipher image that should be random, with no significant correlation of pixels in the cipher image. The plaintext image before encryption reflects the overall or local key information such as the structural shape of tissues and organs, boundary contours, lesions, and so on, therefore the plaintext image has a distinct pixel correlation. Fig. 11 illustrates the pixel correlation between the encrypted ciphertext and the decrypted plaintext for each method in the comparison experiment. In terms of the correlation of the encrypted ciphertext image, the pixel correlation of the ciphertext image encrypted by CYC_{GAN} and P2P is optimal, and the pixel correlation coefficient is almost close to 0. However, the quality of the CYC_{GAN} decrypted plaintext is poor and the correlation between the plaintext and the pixel correlation of the original plaintext reference image is very different from the correlation shown in Fig. 11(c1) and (d1). The P2P encrypted ciphertext image is closer to the pixel correlation of the plaintext reference image, but the quality of its decrypted image is too sharpened for the overall image. This phenomenon causes unimportant details in the original image to be sharpened, which can easily mislead the clinician's diagnosis. Although the correlation of pixels of ciphertext and plaintext images of CNN, CCPL, and UGA is high, the quality of their encrypted and decrypted ciphertext and plaintext images is poor.

3.5. Algorithmic Complexity

Three universally used metrics, floating-point operations (FLOPs), number of parameters, and throughput, were adopted to compare the complexity of different DL-based encryption algorithms. The results listed in Tab. 4 were calculated on the public dataset. As can be seen from Tab. 4, the EDG-Net network has the largest floating-point operations and a larger number of parameters for the network. This is due to the generators and discriminators of the plaintext and ciphertext domains in EDG-Net. The encoder, decoder, and DTA in EDG-Net require operations such as convolution, upsampling, and downsampling, so they consume more floating-point operations. EDG-Net has larger network parameters because of the addition of DEDS, DTA, and AN, which is about twice as large as UGA and three times as large as CYC_{GAN}. The parameters of the network after the training of the encrypted and decrypted network are considered as the secret key for the receiver. EDG-Net is second only to UGA in terms of speed of testing once it has been trained on the dataset. Considering the accuracy, reliability, and security of the encryption and decryption algorithms, the model complexity and computation time of EDG-Net are completely acceptable.

4. Discussion and Conclusion

In this work, we proposed Encryption and Decryption based GAN-attention Network, EDG-Net, for information protection of CT images in IoMT and telemedicine. The proposed EDG-Net consisted of plaintext and ciphertext generators, discriminators, and domain transformation of attention and adaptive normalization. To improve the security of encryption and the accuracy of decryption, the double encryption strategy was adopted on EDG-Net. In the comparison experiments, the qualitative and quantitative results and analyses of the public and private datasets demonstrated that the encryption quality of the ciphertext image after encryption by EDG-Net was higher than that of other encryption methods, and the accuracy of the decrypted plaintext was also higher than that of other algorithms. The ablation experiments demonstrated the irreplaceable role of the adopted double encryption strategy, domain transfer of attention, and adaptive normalization in EDG-Net. The security, robustness, and reliability of the EDG-Net encryption system were also verified by histogram analysis and pixel correlation analysis. Finally, a complexity analysis was proved to determine the practical applications of EDG-Net. Compared to other methods, the proposed EDG-Net performs well in all the above experiments, boasting significant encryption and decryption capabilities, and maintaining the robustness and reliability of the encryption system when facing attacks from third parties. However, the algorithmic complexity of the proposed EDG-Net was more complex than other algorithms for encrypting images due to the complex encoding and decoding processes, and the model was more time-consuming. However, since the key was trained as a network parameter, the more network parameters, the larger the key space, and the higher the security performance of the encryption algorithm. Finally, EDG-Net is effective in guaranteeing the security of images with clinical diagnostic significance sent by the sender in a specific remote diagnosis. When the receiver receives the ciphertext and key (trained network parameters), they can effectively decrypt the ciphertext image accurately, and the decrypted image can meet the needs of clinical diagnosis. EDG-Net can effectively promote the development of remote diagnostic medical treatment for acute heart attack, acute stroke, and other emergency diseases in underdeveloped areas of medical services such as mountainous areas and island medical treatment, and can effectively ensure the safety and fidelity of clinical IoMT-based examination reports downloaded outside the hospital, medical teaching, and other practical intelligent medical application scenarios.

Although the EDG-Net demonstrated encouraging improvement in CT image encryption for IoMT and telemedicine, some issues are still to be noticed. Limitations and future research: (1) EDG-Net's training time is too long. EDG-Net's training time is less than optimal and the model parameters are large. In the future, we will work on maximizing the retention of model parameters to ensure the secret key time while minimizing the training time of EDG-Net, and more advanced pre-training strategies and deep-learning models will be utilized to solve this problem. (2) Deep learning approach is adopted in EDG-Net to accomplish encryption and decryption tasks, problems such as loss of information are generated at the stage of extraction of features of deep learning, so EDG-Net is not compared with traditional lossless encryption and decryption algorithms. In the future, we will work on developing a lossless deep learning-based CT image encryption and decryption network. (3) EDG-Net can only accomplish image encryption in medical IoT and remote diagnosis, but not hide key information on medical images. Therefore, the development of a joint encryption network to develop new key information-hiding algorithms to establish a high-dimensional and integral medical image protection framework for medical images will be the next focus of research. (4) Existing datasets are limited in quantity and ignore the clinical needs in real-life scenarios. In contrast, medical Internet-based telemedicine has to face very rich clinical scenarios and 2D or 3D medical images generated by different devices such as CT, MRI, DR, ultrasound, endoscopy, and pathology slides should be widely collected and labeled. Therefore, the comprehensive consideration of constructing a new type of medical image dataset under multiple realistic scenarios will become the next research focus.

5. Acknowledgment

The authors would like to thank the Nanjing General PLA Hospital for providing clinical data, and data are adherent to the tenets of the Declaration of Helsinki.

References

[1] A. K. S. A. B, M. A. A, G. B. B. B. C, An intelligent learning approach for improving ECG signal classification and arrhythmia analysis, Artificial Intelligence in Medicine 103.
[2] Y. Wu, S. Zhao, Q. S. H. R. W. J. L. H. M. S., Two-stage contextual transformer-based convolutional neural network for airway extraction from CT images, Artificial intelligence in medicine 143 (Sep.) (2023) 1.1
[3] M. Baygin, O. Yaman, P. D. Barua, S. Dogan, T. Tuncer, U. R. Acharya, Exemplar darknet19 feature generation technique for automated kidney stone detection with coronal CT images, Artificial intelligence in medicine (May) (2022) 127.
[4] M. Chung, J. Lee, S. Park, C. E. Lee, Y. G. Shin, Liver segmentation in abdominal CT images via auto-context neural network and self-supervised contour attention, Artificial Intelligence in Medicine 113 (3) (2021) 102023.
[5] Z. Wu, G. Shi, Y. Chen, F. Shi, L. Luo, Coarse-to-fine classification for diabetic retinopathy grading using convolutional neural network, Artificial Intelligence in Medicine 108 (2020) 101936.
[6] M. Xu, M. Islam, L. Bai, H. Ren, Privacy-preserving synthetic continual semantic segmentation for robotic surgery, IEEE Transactions on Medical Imaging.
[7] Y. Shen, A. Sowmya, Y. Luo, X. Liang, D. Shen, J. Ke, A federated learning system for histopathology image analysis with an orchestral stain-normalization GAN, IEEE Transactions on Medical Imaging 42 (7) (2022) 1969-1981.
[8] J. Shi, Y. Zhang, Z. Li, X. Han, S. Ding, J. Wang, S. Ying, Pseudo-data based self-supervised federated learning for classification of histopathological images, arXiv preprint arXiv:2205.15530.
[9] X. Gong, A. Sharma, S. Karanam, Z. Wu, T. Chen, D. Doermann, A. Innanje, Ensemble attention distillation for privacy-preserving federated learning, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021, pp. 15076-15086.
[10] R. Ye, W. Guo, A chaos-based image encryption scheme using multi-modal

skew tent maps, Journal of Emerging Trends in Computing and Information Sciences 4 (10) (2013) 800-810.

[11] R. Ye, W. Zhou, A chaos-based image encryption scheme using 3D skew tent map and coupled map lattice, International Journal of Computer Network and Information Security 4 (1) (2012) 38.

[12] W. Zhang, K.-w. Wong, H. Yu, Z.-l. Zhu, An image encryption scheme using reverse 2-dimensional chaotic map and dependent diffusion, Communications in Nonlinear Science and Numerical Simulation 18 (8) (2013) 2066-2080.

[13] G. Alvarez, S. Li, Some basic cryptographic requirements for chaos-based cryptosystems, International journal of bifurcation and chaos 16 (08) (2006) 2129-2151.

[14] A. A. Abd El-Latif, L. Li, N. Wang, Q. Han, X. Niu, A new approach to chaotic image encryption based on quantum chaotic system, exploiting color spaces, Signal Processing 93 (11) (2013) 2986-3000.

[15] B. Deng, B. D. Jawerth, G. Peters, W. Sweldens, Wavelet probing for compression-based segmentation, in: Mathematical Imaging: Wavelet Applications in Signal and Image Processing, Vol. 2034, SPIE, 1993, pp. 266-276.

[16] W. Sweldens, The lifting scheme: A custom-design construction of biorthogonal wavelets, Applied and computational harmonic analysis 3 (2) (1996) 186-200.

[17] U. Maurer, K. Pietrzak, The security of many-round Luby-Rackoff pseudo-random permutations, in: International Conference on the Theory and Applications of Cryptographic Techniques, Springer, 2003, pp. 544-561.

[18] L. Liu, L. Zhang, D. Jiang, Y. Guan, Z. Zhang, A simultaneous scrambling and diffusion color image encryption algorithm based on Hopfield chaotic neural network, IEEE Access 7 (2019) 185796-185810.

[19] X. Chai, Z. Gan, K. Yang, Y. Chen, X. Liu, An image encryption algorithm based on the memristive hyperchaotic system, cellular automata and DNA sequence operations, Signal Processing: Image Communication 52 (2017) 6-19.

[20] J. Wu, X. Liao, B. Yang, Image encryption using 2D Hénon-sine map and DNA approach, Signal processing 153 (2018) 11-23.

[21] X. Chai, H. Wu, Z. Gan, Y. Zhang, Y. Chen, K. W. Nixon, An efficient visually meaningful image compression and encryption scheme based on compressive sensing and dynamic LSB embedding, Optics and Lasers in Engineering 124 (2020) 105837.

[22] T. Dong, T. Huang, Neural cryptography based on complex-valued neural network, IEEE transactions on neural networks and learning systems 31 (11) (2019) 4999-5004.

[23] P. Fang, H. Liu, C. Wu, A novel chaotic block image encryption algorithm based on deep convolutional generative adversarial networks, IEEE Access 9 (2020) 18497-18517.

[24] Y. Ding, F. Tan, Z. Qin, M. Cao, K.-K. R. Choo, Z. Qin, DeepKeyGen: a deep learning-based stream cipher generator for medical image encryption and decryption, IEEE Transactions on Neural Networks and Learning Systems 33 (9) (2021) 4915-4929.

[25] Y. Ding, G. Wu, D. Chen, N. Zhang, L. Gong, M. Cao, Z. Qin, DeepEDN:

A deep-learning-based image encryption and decryption network for internet of medical things, IEEE Internet of Things Journal 8 (3) (2020) 1504-1518.

[26] M. K. Khan, N. Alrajeh, S. U. Khan, Internet of medical things (IoMT) security and privacy: A survey of recent advances and enabling technologies, ACM Computing Surveys 55 (4) (2022) 1-37.

[27] L. Zhang, Y. Wang, Y. Li, D. Shen, Attention-based generative adversarial network in medical imaging: A narrative review, Computers in Biology and Medicine 149 (2022) 105948.

[28] S. T. Ahmed, A. Mahmood, M. Anwar, R. Bakhsh, N. Javaid, Medical image encryption: A comprehensive review, Computers 12 (8) (2023) 160.

[29] S. Gupta, S. Mathur, Security and privacy in IoMT-based digital healthcare: A survey, in: Smart Computing and Communications, Springer, 2021, pp. 547-558.

[30] H. Guan, M. Liu, Domain adaptation for medical image analysis: A survey (2021). arXiv:2102.09508.

[31] K. Lata, L. R. C. E. Maddi, Deep learning for medical image cryptography: A comprehensive review, Applied Sciences 13 (14) (2023) 8295.

[32] X. Liao, M. A. Hahsmi, R. Haider, et al., An efficient mixed inter-intra pixels substitution at 2bits-level for image encryption technique using DNA and chaos, Optik-International Journal for Light and Electron Optics 153 (2018) 117-134.

[33] Y. Wu, Y. Zhou, J. P. Noonan, S. Agaian, Design of image cipher using latin squares, Information Sciences 264 (2014) 317-339.

[34] K. A. K. Patro, B. Acharya, Secure multi-level permutation operation based multiple colour image encryption, Journal of information security and applications 40 (2018) 111-133.

[35] Z. Wu, Z. Zhu, J. Du, X. Bai, CCPL: Contrastive coherence preserving loss for versatile style transfer, in: European Conference on Computer Vision, Springer, 2022, pp. 189-206.

[36] L. A. Gatys, A. S. Ecker, M. Bethge, Image style transfer using convolutional neural networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 2414-2423.

[37] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 2223-2232.

[38] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125-1134.

[39] J. Kim, M. Kim, H. Kang, K. Lee, U-GAT-IT: Unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation, arXiv preprint arXiv:1907.10830.