E-mail: jig@aircas.ac.cn
Website: www.cjig.cn
Tel: 010-58887035
中国图象图形学报
JOURNAL OF IMAGE AND GRAPHICS

中图法分类号:TP309; TP18　　**文献标识码**: A　　**文章编号**: 1006-8961(2024)07-1934-14

论文引用格式:Chang X Q, Wang M H, You D T and Wu X J. 2024. Cryptanalysis method for chaotic image encryption system. Journal of Image and Graphics, 29(07):1934-1947(常晓琦, 王明合, 游大涛, 武相军. 2024. 面向混沌图像加密系统的密文分析方法. 中国图象图形学报, 29(07):1934-1947)[DOI:10.11834/jig.230147]

面向混沌图像加密系统的密文分析方法

常晓琦,王明合,游大涛*,武相军

河南大学软件学院,开封　475000

摘　要:**目的**　密文评估方法在衡量和增强混沌图像加密系统的安全性方面发挥着至关重要的作用。现有以密钥空间、密文密钥敏感性、像素个数变化率和统一平均变化强度等为代表的评估方法虽无法保证通过测试的加密系统一定具有非常高的安全性。而以选择明文攻击为代表的分析方法,与前者相比缺乏通用性和一致性,需要针对不同的加密系统设计不同的攻击方案。针对上述问题,本文基于深度学习模型面向混沌图像加密系统提出了一种兼具通用性和有效性的密文评估方法。**方法**　该方法的核心思路是以降噪自编码器为基础模型,使用编码器分别对图像加密方法中的扩散密文、置乱密文和完整加密密文进行深度表示,然后使用解码器以上述深度表示为输入生成相应的不同明文,最后统计该明文与真实明文间的结构相似度作为度量加密方法抵抗密码学常用攻击手段能力的量化指标。对于一个加密方法来说,不仅其完整加密密文必须完全不可破译,而且其置乱阶段和扩散阶段的密文中也必须有一项是完全不可破译的,否则表明加密方法存在严重的安全缺陷。另外,密文数据集是影响上述方法有效性的关键因素。针对该问题,本文提出了一种相关性密文生成方法,该方法充分利用了明文敏感性密钥的特性,确保了生成的密文和本文评估方法的真实性和有效性。**结果**　本文以Arnold置乱、2D-SCL(2D chaotic map based on the sine map, the chebyshev map and a linear function)加密和基于二维交叉混沌映射的量子加密为例对提出的密文评估方法进行了实验验证,实验中用到的数据集分别是MNIST(modified national institute of standards and technology database)和Fashion-MNIST。实验结果显示,本文提出的密文分析模型对上述加密方法及其各个阶段生成的密文图像表现出不同的密文分析能力:对Arnold置乱、2D-SCL扩散和量子bite置乱的密文来说,破译图像与真实图像间结构相似性指数(structural similarity,SSIM)的值均大于0.6;虽然在其他阶段的密文分析方面的效果较低,但也能破译出部分关键明文信息,呈现出较高的结构相似度。**结论**　本文提出的密文图像分析方法通过客观的评价指标数据,能够有效地评估加密方法的安全性,为提升混沌图像加密方法的安全性提供了直观有效的量化依据,具有较高的指导意义。

关键词:图像安全;密文分析;混沌图像加密系统;明文敏感性;深度学习;降噪自编码器

Cryptanalysis method for chaotic image encryption system

Chang Xiaoqi, Wang Minghe, You Datao*, Wu Xiangjun

School of Software, Henan University, Kaifeng　475000, China

Abstract: **Objective** Cryptography security analysis methods play a vital role in measuring and enhancing the security of chaotic image encryption systems. The existing ciphertext analysis methods for chaotic image encryption are generally

收稿日期:2023-04-07;**修回日期**:2023-09-20;**预印本日期**:2023-09-27

*通信作者:游大涛　youdatao@163.com

基金项目:国家自然科学基金项目(61872125);河南省科技攻关(国际科技合作类)项目(182102410051);河南省自然科学基金项目(202300410038)

Supported by: National Natural Science Foundation of China (61872125); Science and Technology Foundation of Henan Province (182102410051); Natural Science Foundation of Henan Province, China (202300410038)

divided into two categories. Although the evaluation methods based on numerical statistics, which are represented by key space analysis, sensitivity analysis of ciphertext to secret key, numbers of pixels change rate, and unified average changing intensity, have excellent versatility and consistency, the security of the test-passed encryption scheme cannot be ensured. While common attack methods in cryptography, which are represented by selective plaintext attack, can intuitively and effectively assess the security of chaotic encryption schemes, they lack versatility and consistency compared with security analysis methods, and different attack schemes need to be designed for different encryption schemes. To address the problem, this paper proposes a cryptanalysis method of chaotic image encryption system that is both versatile and effective based on denoising autoencoder. **Method** The ciphertext analysis method is improved based on the denoising autoencoder. It uses a cryptographic system with a known specific encryption steps to encrypt the plaintext image dataset and constructs the ciphertext analysis model, which takes the ciphertext image dataset as input and the original image dataset as the target data. The cryptanalysis model uses the encoder to obtain the depth representation of the diffusion ciphertext, scrambling the ciphertext and fully encrypted cipher image generated by the image encryption scheme, which is the structural features of images extracted from ciphertext images, and then uses the decoder to generate the different deciphered plaintext with the above depth representation as input. In this way, a deciphering model for a certain known encryption scheme can be trained, thereby achieving the purpose of ciphertext analysis. The effect of cryptanalysis is measured objectively and comprehensively by proposing three types of evaluation indicators suitable for ciphertext analysis based on peak signal to noise ratio(PSNR) and structural simitarity(SSIM): max PSNR(MPSNR), max SSIM(SSIM), average of PSNR(APSNR), avarage of SSIM(ASSIM), cumulative distribution of PSNR(CDPSNR), and cumulative distribution of SSIM(CDSSIM), which measure the ability of an encryption scheme to resist popular attacks in cryptography by counting the structural similarity between the generated deciphered plaintext and real plaintext. This evaluation indicator, in addition to the subjective perception of human eyes, can visually display the differences between plaintext images and deciphered images by real data and complete the evaluation of the security of encryption schemes. For the one encryption scheme, not only the fully encrypted ciphertext but also one of the ciphertexts in the scrambling stage and the diffusion stage must be completely undecipherable; otherwise, the encryption scheme has serious security flaws. In addition, the ciphertext dataset is a key factor that affects the effectiveness of the above method. A correlation ciphertext generation method that generates three kinds of ciphertext sets——scrambled ciphertext, diffused ciphertext, and encrypted ciphertext—is proposed to address this issue. This generation method makes full use of the characteristics of chaotic image encryption systems and plaintext sensitive keys to ensure the authenticity and effectiveness of the generated ciphertext and the proposed evaluation method. When cryptanalyzing different chaotic image encryption schemes, changing only the generation scheme of ciphertext in each encryption stage based on the encryption algorithm is necessary. Without changing the training stage, testing stage, and model, the cryptanalysis and security evaluation of different chaotic encryption schemes can be completed. **Result** This paper takes Arnold scrambling, 2D-SCL image encryption scheme, and quantum image encryption scheme based on 2D Sine2-Logistic chaotic map as examples to verify the proposed ciphertext evaluation method. The datasets used in the experiment are MNIST and Fashion-MNIST. Experimental results show that the proposed cryptanalysis model has a different analysis ability for the ciphertexts generated by the above encryption scheme and their various stages. For the ciphertexts of Arnold scrambling, 2D-SCL's diffusion, and bite scrambling in quantum encryption, the SSIM values between the decrypted images and the real plain-images are all greater than 0.6. The cryptanalysis model can learn low-dimensional structural features, same as the equivalent keys, to restore the ciphertext image. Although the effect of cryptanalysis in other stages is lower, it can also decipher some key plaintext information, showing a high degree of structural similarity. This finding also indicates that, for an encryption scheme, a high plaintext sensitivity of the secret key corresponds to a high security of the chaotic sequence, and a strong plaintext sensitivity of its equivalent key corresponds to a reduced likelihood that it can be cracked. **Conclusion** The proposed ciphertext image analysis method can evaluate the security of encryption schemes comprehensively and effectively by using objective data as the evaluation index, which provides an intuitive and effective quantitative basis for improving the security of chaotic image encryption methods, and has high guiding significance.

Key words: image security; ciphertext analysis; chaotic image encryption system; plaintext sensitivity; deep learning; denoising autoencoder

0 引 言

随着信息技术的飞速发展,数字图像作为信息的主要载体,每天产生海量的数据。为了保护图像中的隐私数据,研究人员认为图像加密技术是解决这一问题的关键。与传统的分组加密算法(AES(advanced encryption standard)、DES(data encryption standard)、RC(rivest code)、SPECK(set partitioned embedded block)等)相比,基于混沌系统的图像加密算法在安全性和加密速度方面存在着巨大的发展潜力,目前仍处于快速发展阶段。自从Fridrich(1998)提出著名的扩散—置乱混沌加密框架以来,涌现了众多加密方法。首先,以关注置乱的加密方法为例,李玉珍等人(2016)提出了基于zigzag变换的彩色图像加密方案;Sun和Chen(2021)提出了基于Arnold变换的混沌加密系统;梁颖和张绍武(2018)提出一种基于位级同步置乱扩散和像素级环形扩散的图像加密算法以提高图像加密效率和安全性。其次,以借鉴DNA编码规则的加密方法为例,Rehman等人(2018)提出基于SHA-2(secure Hash algorithm 2)和DNA(deoxyribonucleic acid)互补规则的彩色图像加密技术;周辉等人(2021)提出混沌系统和DNA编码并行的遥感图像加密算法,解决了遥感图像加密速度差和安全性不足的问题。最后,以使用高维混沌系统的加密方法为例,Joshi等人(2020)提出基于三维混沌映射和二维离散小波变换的图像加密方案。

安全性的高低是评价混沌图像加密系统成败的关键,因此对加密系统的安全性进行评估是非常有必要的。密文分析方法作为混沌密码的重要组成部分,在衡量和增强混沌图像加密系统的安全性方面发挥着至关重要的作用。现有针对混沌图像加密的密文分析方法一般分为两类:基于数值统计的密文分析方法和密码学分析(攻击)方法(温贺平,2019)。

基于数值统计的密文分析方法主要包括:以密钥空间分析为代表的穷举攻击;以直方图分析、相邻像素相关性等为代表的统计分析;以像素个数变化率(numbers of pixels change rate, NPCR)、统一平均变化强度(unified average changing intensity, UACI)为代表的抗差分分析;以剪裁攻击、噪声攻击为代表的鲁棒性分析等。利用数值统计分析方法对混沌图像加密算法进行密文分析是一种常见且通用的分析方法,但实际上经过数值统计分析方法评估的加密算法,却有可能被蛮力攻击、选择明文攻击和选择密文攻击等分析方法破解。如Solak等人(2010)采用选择密文攻击对基于混沌理论和置乱—扩散结构的图像加密算法(Fridrich,1998)实现了破译;Li等人(2018)采用选择明文攻击对一种基于三维比特矩阵置乱和扩散的加密算法(Zhang等,2016)实现了破译。Dou和Li(2020)提出了一种针对参数未知情况的一维混沌加密算法(Pak和Huang,2017)的有效攻击方案;Li等人(2019)针对改进的一维比特级混沌图像加密算法(Pak等,2019)提出了有效的选择明文攻击策略。

密码学分析方法攻击强度由弱到强分别是:唯密文攻击、已知明文攻击、选择明文攻击、选择密文攻击。另外,还有两种常用的密码学分析方法:线性攻击和差分攻击。线性攻击属于已知明文攻击,利用已知的明密文对去获取相关密钥;差分攻击则属于选择明文攻击,通过比较不同的明文差分和对应的密文差分获取相关密钥。密码学分析方法具有极强的密文分析能力,能有效地评估混沌加密方法安全性的高低,但需要根据不同的加密方法设计不同的分析方案,缺乏通用性。

究其原因,主要是因为现有混沌图像加密系统的评估方法主要关注密文的随机性统计特征,难以有效评估密钥的明文敏感性及扩散和置乱混沌序列的明文敏感性,忽略了等效密钥的存在,致使现有基于密文的评估方法存在严重缺陷,难以有效评估混沌加密系统的安全性。因此,设计能有效度量扩散和置乱混沌序列的明文敏感性的加密系统评估方法是一个亟待解决的问题。

深度学习迅猛发展,成为多个领域的研究热点,其在密码学领域也有许多应用。Wu等人(2021)引入了生成式对抗网络与SHA-256、Logistic映射结合实现图像加密;Melicher等人(2016)构建人工神经网络模拟文本密码对猜测攻击的抵抗力,通过密码破译器有效快速地对密码进行猜测攻击,完成了破译;Hou等人(2020)基于深度残差网络提出线性攻击架构,将适量完整和部分线性表达式作为输入,经过训练的网络,能有效地对减轮DES进行密钥恢复,成功利用神经网络实现了对多轮DES的已知明文攻击。然而,目前还没有基于深度学习的混沌图像加密系统分析方法方面的研究,因此,有必要开展面向

混沌加密系统的基于深度学习的密文分析与评估方法研究。

鉴于现有评估方法的缺陷和深度学习在密文分析中的优势，本文针对混沌密码系统提出了一种基于降噪自编码器的密文图像分析方法。该方法通过神经网络模型从密文图像中提取图像的结构化特征，之后由特征重新生成对应的解密图像，可训练出针对某种已知加密方案的破解模型，进而达到密文分析的目的。该方法可以在密钥未知的情况下直接攻击混沌加密方法生成的密文，在训练解密模型后，可以快速、自动地从加密图像中高效地重建明文图像。与传统评估方法相比，不但直观有效，而且灵活并具有通用性和可扩展性。此密文分析方法的研究，为提升混沌加密方法的安全性提供了直观有效的量化依据，也为密文分析方法的研究指引了新的发展方向。

1　基于降噪自编码器的混沌图像加密系统分析方法

1.1　降噪自编码器

降噪自编码器（denoising autoencoder，DAE）是深度学习中一种具有对称结构的深度网络模型（Vincent 等，2008），在训练时给原始数据添加随机的噪声扰动作为输入，其主要功能是将输入的受污染数据转换为未被损坏的目标数据。降噪自编码器的核心思路是首先采用编码器将输入的高维数据投影到低维空间的特征信息，然后借助解码器将低维空间的特征信息投影为高维空间的目标信息并作为降噪自编码器的输出数据，最后通过计算目标数据（即原始数据）和输出数据间的差值对自编码器模型进行约束优化，进而得到合适的模型变换系数。

降噪自编码器的结构如图 1 所示，通过一定概率的节点置零或给原始数据 $\boldsymbol{x}$ 添加采样自某种类型分布 N 的噪声 ε 生成编码器的输入 $\boldsymbol{x}'$，即 $\boldsymbol{x}'=\boldsymbol{x}+\varepsilon,\varepsilon\sim\mathrm{N}$。添加噪声后，前面的子网络需要学习映射关系 $f_{\theta_1}:\boldsymbol{x}'\rightarrow z$，即从 $\boldsymbol{x}'$ 中学习到数据的真实隐藏变量 $z=f_{\theta_1}(\boldsymbol{x}')$；后面的解码器网络学习映射关系 $g_{\theta_2}:z\rightarrow\bar{\boldsymbol{x}}$，即重构出与原始输入尽可能接近的输出 $\bar{\boldsymbol{x}}=g_{\theta_2}(z)$。网络模型参数 θ 的优化目标可描述为

$$\theta=\underbrace{\arg\min}_{\theta}L\left(\boldsymbol{x},g_{\theta_2}\left(f_{\theta_1}(\boldsymbol{x}')\right)\right)\tag{1}$$

式中，$L(\boldsymbol{x},\bar{\boldsymbol{x}})$ 表示 $\boldsymbol{x}$ 和 $\bar{\boldsymbol{x}}$ 的距离度量，称为误差函数。网络通过选择合适的优化函数使参数朝目标的方向进行更新。借助于深层神经网络的非线性特征提取能力，降噪自编码器可以学习到具有较高鲁棒性的输入数据的特征。

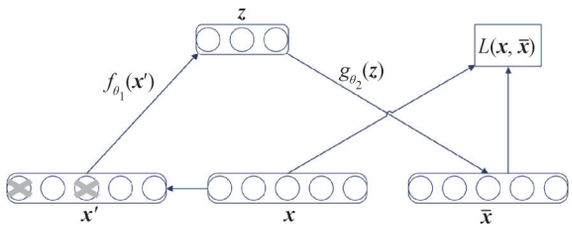


图 1　降噪自编码器模型图

Fig. 1　The diagram of denoising autoencoder model

1.2　基于降噪自编码器的密文图像分析模型

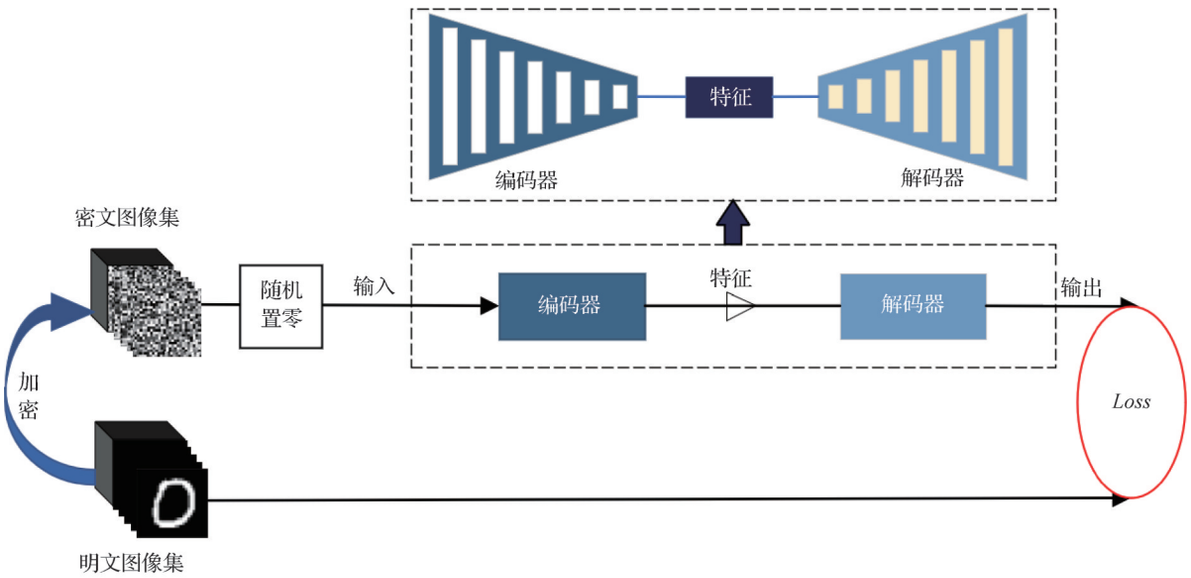
密文分析是一种将密文转换为明文的技术，其思路和降噪自编码器非常相似，而且降噪自编码器在生成目标数据方面具有优异的表现。鉴于此，基于降噪自编码器提出了一种新的密文分析方法。为了提升上述方法的有效性，本文结合密文分析的特点对常规降噪自编码器进行了 3 项针对性改进：

1）采用卷积层和池化层代替常规降噪自编码器的全连接层，卷积层用于提取特征，池化层用于压缩特征。卷积层能够较好地保留二维数据的空间信息，相较全连接网络更适用于图像的密文分析；卷积神经网络能够提升网络结构的稳定性和泛化能力，避免过拟合现象的产生。

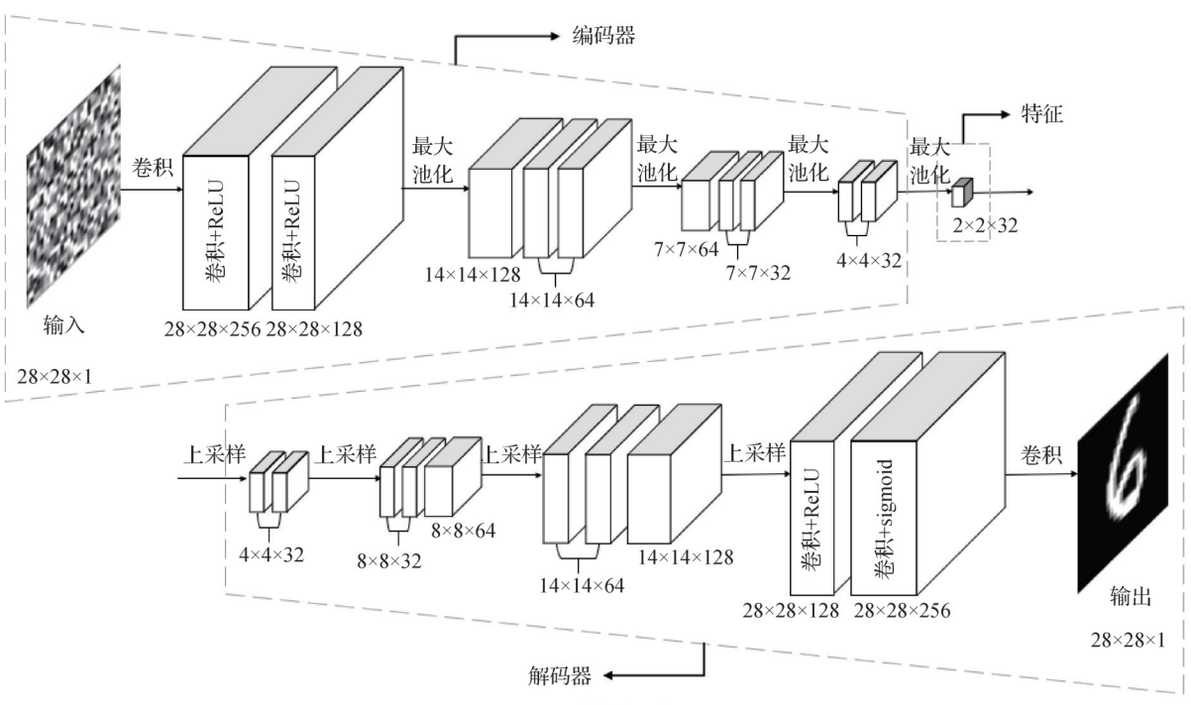
2）以椒盐噪音代替高斯噪音，根据均匀分布模型随机选取输入的某些值并将其随机置为 0 或 255。这是由于在传输过程中产生的误差主要由脉冲噪音引起，其密度分布符合均匀分布，而非原模型假定的高斯噪音。

3）为更好地对混沌加密系统进行分析，模型的输入是密文，输出是相应的明文，而非输入和输出均是明文，用于完成密文到明文的转换。

基于以上 3 点，本文提出了基于降噪自编码器的混沌加密系统分析模型，具体信息如图 2 所示。基于 DAE 的密文分析模型的输入是密文，输出则是作为目标的明文，结构如图 2（a）所示。该模型的编码器模块将加噪密文 $\boldsymbol{X}'$ 映射到低维特征空间，得到加噪密文 $\boldsymbol{X}'$ 的深度表示 $\boldsymbol{Z}$，具体过程可描述为 $\boldsymbol{Z}=$



(a) 模型结构图



(b) 网络架构图

图2　混沌加密系统分析模型

Fig. 2　The cryptanalysis method of chaotic encryption system ((a) structure of cryptanalysis method; (b) network architecture of cryptanalysis method)

$f_{\theta_1}(\boldsymbol{X}')$;该模型的解码器以深度表示$\boldsymbol{Z}$为输入,输出则是生成的明文图像$\overline{\boldsymbol{X}}$,这一过程可用公式描述为$\overline{\boldsymbol{X}}=g_{\theta_2}(\boldsymbol{Z})$,$\theta_1$表示编码器的权重$\boldsymbol{W}$和偏置$\boldsymbol{b}$;$\theta_2$表示解码器的权重$\boldsymbol{W}$和偏置$\boldsymbol{b}$。

编码器和解码器通过前向传播生成破译图像,而密文分析模型的目的就是将输入的密文图像转化为明文,因此需要借助损失函数计算破译图像与明文图像间的误差,该模型选用均方误差函数计算二者间的误差。对于原始明文$\boldsymbol{X}$和破译明文$\overline{\boldsymbol{X}}$,基于均方误差的损失函数定义为

$$Loss\left(\boldsymbol{X},\overline{\boldsymbol{X}}\right)=\frac{\sum_{j=1}^{J}\left(\overline{\boldsymbol{X}}_j-\boldsymbol{X}_j\right)^2}{J} \tag{2}$$

式中,J为像素点个数。模型通过反向传播对编码器和解码器的权重参数进行优化,为了学习得到最

佳的权重 $\boldsymbol{W}$ 和偏置 $\boldsymbol{b}$，需要在训练阶段逐步最小化损失函数的值。因此，该密文分析模型的优化目标表示为

$$\theta_{\text{opti}} = \underbrace{\arg\min}_{\theta} \left\| g_{\theta_2}\left(f_{\theta_1}\left(\boldsymbol{X}' \right) \right) - \boldsymbol{X} \right\|^2 \tag{3}$$

式中，θ_{opti} 代表参数 θ 最优值。

借助于前向传播和后向传播算法，通过对式(3)的迭代更新，即可求得密文分析模型的权重参数 $\boldsymbol{W}$ 和偏置参数 $\boldsymbol{b}$ 的值。在此基础上就可对密文进行分析，进而生成相应的明文信息。

1.3 密文分析模型的网络结构

本文提出的密文分析模型主要由编码器和解码器构成，具体结构如图 2(b)所示。编码器由输入层、隐藏层以及输出层组成。输入层以密文图像(本文以尺寸为 28 × 28 像素的单通道图像为例，若分析对象为彩色图像，则模型输入的密文图像与最终输出的明文图像均为三通道矩阵)作为编码器的输入，为了使模型在训练过程中平稳收敛，密文图像在输入前需要将像数值归一化为[0,1]区间的值。隐藏层主要由卷积层和池化层组成。卷积层的目的是提取图像中特定区域的特征映射；池化层的目的是通过减少特征的数量以降低模型复杂度。隐藏层中每个卷积层的卷积核大小均为 3 × 3，步长为 1 × 1，为了使卷积的输出和输入大小一致，图像的顶和底分别填充 1 行 0 值、左右分别填充 1 列 0 值，每个卷积层都添加 ReLU 非线性激活函数；此外每层卷积核的数量分别是 256、128、64 和 32；另外，为了提升模型的训练速度和性能，隐藏层中首个卷积层的输出借助 Batch Normalization 进行了归一化。隐藏层中每个池化层均采用最大池化技术对卷积层的输出降采样，池化层的感受野和步长均为 2 × 2。编码器的输出是输入密文的低维表示，并作为后续解码器的输入。

解码器以密文的低维表示为输入，后连接隐藏层和输出层。其中隐藏层由卷积层和上采样层组成。上采样的目的是增大感受野，将图像由小分辨率映射到大分辨率，使之恢复到原来的尺寸以便于和目标数据进行误差计算。隐藏层中每个卷积层的卷积核大小和步长、填充值、激活函数与编码器一致；卷积核数量分别为 32、64、128 和 256。上采样层的功能是对前面紧邻的卷积层的输出进行上采样，步长为 2 × 2。输出层采用 sigmoid 激活函数，通过卷积重构输入的明文图像。

在训练阶段，将抓取的批量图像送入模型进行前向计算获得重建图像向量，使用均方误差函数计算重建图像与目标明文图像的误差损失；获得损失值后，利用深度学习框架的自动求导机制完成参数的梯度计算，并使用 Adam 方法(Kingma 和 Ba，2015)优化目标，更新网络权值参数。

在测试阶段，将测试密文集输入到训练好的模型中进行测试，输出为破译图像，通过破译效果判断密文分析的结果，并使用相应的评价指标进行评估。

1.4 密文数据集生成方法

本文提出的密文分析模型在训练阶段需要使用大量明文及其对应的密文。密文数据集的生成直接影响密文评估方法的有效性。密文分析方法通常假定加密方法已知但密钥未知。由于密文攻击本身是针对不同加密算法进行破解，破解对象不同，涉及到的密钥分配也不同。此外，基于混沌系统的图像加密算法普遍采用置乱—扩散结构，为了有效评估该加密算法的安全性，不仅需要对加密系统进行整体评估，而且还需要对加密算法中的扩散、置乱等不同步骤单独评估，便于发现加密系统中的安全性缺陷所在。加密系统必须满足扩散和置乱两类密文中至少一项的绝对安全以及完整加密过程绝对安全，否则其安全性存在问题。因此，密文数据集的生成需要遵循以下准则：

1)如果加密系统的密钥是明文敏感的，那么为每个明文分配一个对应的密钥。

2)如果加密系统的部分密钥值是明文敏感的，则仅为每个明文分配一个明文相关的部分密钥值。

3)如果加密系统的密钥是明文不敏感的，那么直接假设密钥是固定但未知的。

4)针对同一个加密系统，需要分别生成置乱密文、扩散密文和加密密文 3 类密文数据集(针对单个加密步骤的扩散、置乱密文集数量由具体加密方案而定)。

根据上述密文生成准则，对于某个混沌加密系统，该密文集生成方法的步骤如下：

1)将明文图像数据集按一定比例分为训练集和测试集，比例为 n:1。

2)设置密钥。将训练集和测试集都分为 t 组，在密钥明文不敏感的部分，每组设置相同密钥值。

3)针对同一加密方法的置乱阶段、扩散阶段和

完整加密过程，分别执行步骤1）和步骤2）即可生成置乱密文、扩散密文和加密密文3类密文图像集。

密文数据集是影响密文分析方法有效性的关键因素，本方法充分利用了明文敏感性密钥的特性，确保了生成的密文和密文分析方法的真实性和有效性。

1.5 评估指标

评估指标的主要作用是计算密文分析方法破译出的明文图像与原始明文图像间的相似度，是评估加密系统有效性的客观度量指标。峰值信噪比（peak signal to noise ratio，PSNR）和结构相似性指数（structural similarity，SSIM）是两个广泛应用于评估重构图像与原始图像间相似度的指标，本文也将它们作为密文分析效果的基本度量指标。但是，密文分析与常规图像处理不同，况且本文的密文分析模型在对加密方法破译时需要统计整个测试集的破译情况来衡量一个加密算法抵抗攻击的能力，鉴于其特殊性，为了客观全面地度量密文分析的效果，在PSNR和SSIM基础上，提出了新的度量指标，分别是最大峰值信噪比（max PSNR，MPSNR）、平均峰值信噪比（average of PSNR，APSNR）、峰值信噪比的累积分布（cumulative distribution of PSNR，CDPSNR）、最大结构相似性指数（max SSIM，MSSIM）、平均结构相似性指数（average of SSIM，ASSIM）、结构相似性指数的累积分布（cumulative distribution of SSIM，CDSSIM）。

PSNR是度量图像间相似性的经典指标，PSNR值越大，图像的差异越小；但这种基于均方误差的评价指标难以体现人类视觉感受到的图像相似性。SSIM基于样本之间的亮度、对比度和结构3个维度比较衡量其相似度，取值范围为[0, 1]，值越大，结构相似度越高。这种基于人眼会提取图像中结构化信息的假设，更符合人眼视觉感知。

MPSNR、MSSIM、APSNR、ASSIM是计算明文图像集$\boldsymbol{X}$与破译图像集$\overline{\boldsymbol{X}}$之间的PSNR值和SSIM值，分别取其最大值和平均值，假定图像集样本数为H，具体为

$$\begin{cases} f_{\mathrm{MPSNR}} = \max\left(f_{\mathrm{PSNR}}(\boldsymbol{X}_h, \overline{\boldsymbol{X}}_h)\right) \\ f_{\mathrm{MSSIM}} = \max\left(f_{\mathrm{SSIM}}(\boldsymbol{X}_h, \overline{\boldsymbol{X}}_h)\right) \\ f_{\mathrm{APSNR}} = \dfrac{1}{H}\sum_{h=1}^{H} f_{\mathrm{PSNR}}(\boldsymbol{X}_h, \overline{\boldsymbol{X}}_h) \\ f_{\mathrm{ASSIM}} = \dfrac{1}{H}\sum_{h=1}^{H} f_{\mathrm{SSIM}}(\boldsymbol{X}_h, \overline{\boldsymbol{X}}_h) \end{cases} \tag{4}$$

式中，$h \in [1, H]$，$\boldsymbol{X}_h$和$\overline{\boldsymbol{X}}_h$分别是图像集中某一明文图像与其对应的破译图像。

将这4项数据作为衡量加密算法安全性的评价指标，究其根本是密文分析与常规图像处理的关注点不同，MPSNR和MSSIM代表所有破译图像中单幅图像的最优破译结果，最优结果超过阈值，表示有密文图像可以被成功破译，则证明此加密方案不安全；APSNR和ASSIM代表整体破译水平，平均值越高，表示加密方案越不安全。

累积分布函数（cumulative distribution function，CDF）是概率密度函数的积分，代表实数随机变量Y的取值小于等于y的概率。对于所有可能取值y，累积分布函数定义为

$$F_Y(y) = P(Y \leqslant y) \tag{5}$$

取值范围为[0，1]。求得Y处于$(u, v]$之间的概率为

$$P(u \leqslant Y \leqslant v) = F_Y(v) - F_Y(u) \tag{6}$$

CDPSNR是由统计明文图像与破译图像间PSNR值的累积分布得来，CDSSIM同理。累积分布图能清晰地展现小于等于某个值的概率分布，便于观察PSNR和SSIM值在某个区间的占比，统计解码率。CDPSNR和CDSSIM与横坐标间的面积越小，则破译效果越好，安全性越低。

综上所述，本文提出的以上3类评估指标具有实用性，且对所有基于深度学习模型的密文分析方法，在衡量破译图像与明文图像差异时，除人眼直观感受外，能够用真实数据直观地展示出明文图像与解密图像之间的相似度和差异性，完成该加密方法安全性的评估。

1.6 基于DAE的密文分析方法

基于前述密文分析模型，本节主要对相应的密文分析方法进行介绍，密文分析流程如图3所示。

1）对于明文图像集$\boldsymbol{X}$，使用待分析的混沌加密系统按1.4节所述分别对其进行扩散、置乱、完整加密得到3类密文图像集。

2）分别给3种密文集添加椒盐噪音，得到加噪扩散密文集$\boldsymbol{X}_1'$、加噪置乱密文集$\boldsymbol{X}_2'$和完整加密的加噪密文集$\boldsymbol{X}_3'$。

3）将明文图像集与加噪扩散密文集中所有图像进行预处理，即归一化。

4）将明文图像集作为目标数据，以加噪扩散密文集为输入构建1.2节所提出的密文分析模型，通

过神经网络的训练不断减小重构明文图像与真实明文之间的误差，达到密文分析的目的。

5）将测试密文集送至训练好的模型中测试。

6）分别针对加噪置乱密文集、完整加密的加噪密文集重复步骤 3）—步骤 5），即可完成对扩散阶段、置乱阶段和完整加密过程的密文分析。

7）通过各个阶段预测的破译明文和 1.5 节提出的评估指标判断破解效果，以此衡量该混沌加密系统的安全性。不仅完全加密图像不可破译，而且置乱和扩散两步骤中也必须有一项是不可破译的，否则表明加密方法存在严重的安全缺陷。

8）在对不同的混沌图像加密方法进行分析时，只需在步骤 1）中根据加密算法改变各类密文的具体生成方法即可，而密文分析的训练阶段、测试阶段和模型本身并不需要改变，重复步骤 1）—步骤 7）可完成对不同混沌加密算法的分析和评估。

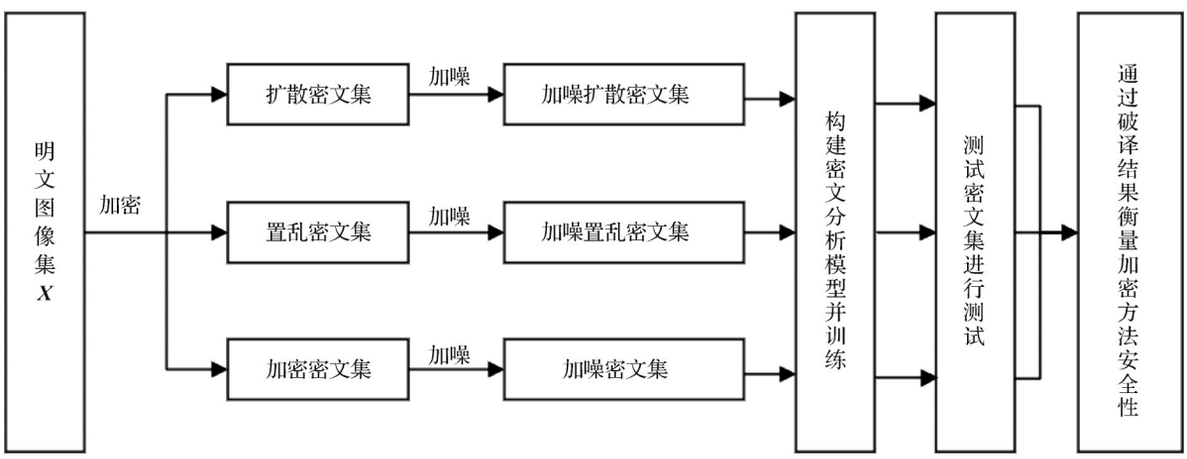


图 3　基于 DAE 的密文分析方法流程图

Fig. 3　The flow chart of cryptanalysis method based DAE

2　仿真实验

仿真实验中计算机显卡配置为 NVIDIA GeForce RTX2060，显存为 6 GB；实验使用的深度学习框架为 TensorFlow。MNIST 数据集（Lecun 等，1998）是手写体图像集，Fashion-MNIST 数据集（Xiao 等，2017）是服饰类图像集，这两个数据集中的图像尺寸均为 28 × 28 像素，图像数量均是 70 000 幅，其中 60 000 幅作为训练数据集，其他作为测试数据集。这两个数据集具有结构简单、图像小、训练速度快等特点，因此常广泛应用于深度学习和图像处理领域。由于这两个数据集中的图像也比较适合作为密文分析的明文，因此本文选用它们作为模型训练和测试阶段的明文数据集。选用相同个数的密钥初始值加密训练集和测试集，生成密文图像集。

2.1　加密方法介绍

本文选用 3 个具有代表性的加密方法作为实验对象：密钥明文不敏感的 Arnold 置乱方法（Farwa 等，2017）、密钥明文敏感的 2D-SCL 加密方法（Chen 等，2020）和密钥明文高度敏感的量子加密方法（Hu 等，2023）。

1）Arnold 置乱。借助有密钥生成的变换矩阵，经过多次变换明文像素的坐标值转换为变换空间的新坐标值，进而达到置乱的目的。对于一个为 $M \times M$ 的灰度图像 $\boldsymbol{X}$ 来说，Arnold 置乱的运算式为

$$\begin{pmatrix} x_c' \\ y_c' \end{pmatrix} = \begin{bmatrix} 1 & p \\ q & pq+1 \end{bmatrix} \begin{pmatrix} x_c \\ y_c \end{pmatrix} \bmod(M) \tag{7}$$

$$\Leftrightarrow \begin{cases} x_c' = (x_c + py_c) \bmod(M) \\ y_c' = (qx_c + (pq+1)y_c) \bmod(M) \end{cases} \tag{8}$$

式中，(x_c, y_c) 为明文图像坐标值，(x_c', y_c') 为经过一次 Arnold 置乱后的坐标值，p、q 取正整数。通过改变置乱的迭代次数可增加其置乱度，由于置乱密钥 p、q 与原像素值无关，即属于明文不敏感的置乱方法。

2）2D-SCL 加密方法。主要借助混沌序列对明文进行置乱和扩散，进而生成密文（Chen 等，2020），整体加密思路如下：

（1）给定密钥 $x_o, y_o, x_1, y_1, k, a, c, d$，其中，$x_0, y_0, x_1, y_1 \in [0, 1]$，$k \in (0, 1)$，$a, c \in (0, +\infty)$，$d \in (-\infty, +\infty)$。

并根据明文图像的SHA-1值生成最后一位密钥λ，计算式为

$$\lambda = \frac{K_1 + K_2 + \cdots + K_{40}}{40 \times 16} \tag{9}$$

式中，$K_1, K_2, \cdots, K_{40}$为160位哈希值每4位为一组生成的40个十进制数，由式(9)可知，λ共有600种不同取值。

(2)假定图像$\boldsymbol{X}$的尺寸为$M \times N$，根据以上密钥生成随机序列$D_{rx}, D_{ry} \in [1, 255], D_{cx}, D_{cy} \in [1, N]$。

(3)D_{rx}, D_{ry}分别参与行扩散与行置乱，D_{cx}, D_{cy}分别参与列扩散与列置乱。扩散、置乱操作后加密完成。

图像哈希值通过密钥λ作用到扩散和置乱过程中，即2D-SCL加密是一种明文敏感的加密方法，其置乱和扩散过程也都是明文敏感的。

3)基于二维Sine2-Logistic交叉混沌映射的量子图像加密方法。整体加密思路可概括如下：

(1)密钥设置。给定密钥α, β，获取明文图像$\boldsymbol{X}$的SHA-256哈希值，得到一个十六进制数组表示的256位哈希值$H(\boldsymbol{X})$，密钥x_{q0}和y_{q0}计算为

$$\begin{cases} x_{q0} = \text{hex2dec}\big(H(\boldsymbol{X})(\alpha : \alpha + 7)\big) \times 10^{-10} \\ y_{q0} = \text{hex2dec}\big(H(\boldsymbol{X})(\beta : \beta + 7)\big) \times 10^{-10} \end{cases} \tag{10}$$

式中，hex2dec(·)表示十六进制数到十进制数的转换函数。

(2)主要加密步骤包括：量子比特级选择性置乱；像素级行列循环移位置乱；重叠反馈扩散。

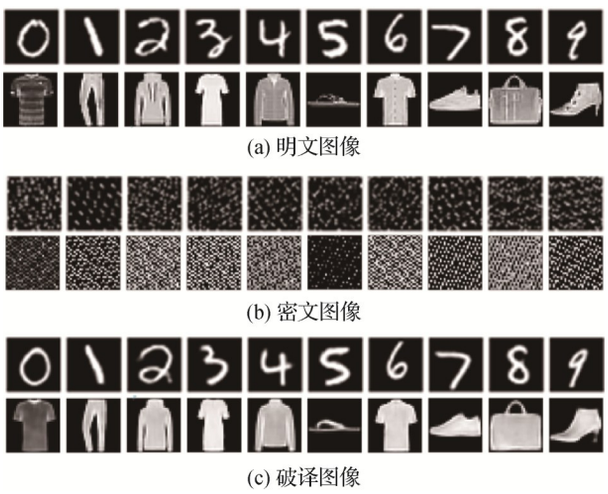
(3)用于扩散的混沌序列$X1$，用于循环行/列移位的控制序列$CRS1$、$CRS2$、$CCS1$、$CCS2$，用于比特级置乱的混沌序列SC、CS，都是由x_{q0}和y_{q0}以公式计算而来。

图像哈希值通过明文高度敏感的密钥x_{q0}和y_{q0}控制着整个加密过程，即本文提到的量子加密方法与明文高度敏感。

2.2 实验验证与分析

本节分别对2.1节所述的3种加密方法及其扩散、置乱、完整加密3类密文进行了密文分析，使用训练集完成训练，测试集上破译结果如下。

1)Arnold置乱分析。使用本文的密文分析方法对Arnold置乱方法进行攻击，生成密文时，设置密钥初始值$p = 1, q = 1$，共变换100次，即p, q取值范围为[1, 100]。两图像集分析结果如图4所示。由图4可知，经Arnold置乱加密的手写体图像被轻易破解，破译图像与明文差别较小，服饰图整体结构与原图也基本一致，破译效果较好。



(a) 明文图像

(b) 密文图像

(c) 破译图像

图4 Arnold分析效果

Fig. 4 Cryptanalysis results of Arnold cipher images ((a) plain images; (b) ciphertext images; (c) cryptanalysis results)

2)2D-SCL扩散分析。生成密文时设置密钥初始值$x_0 = 0.12342, y_0 = 0.56782, x_1 = 0.32324, y_1 = 0.33491, k = 1, a = 2\pi, c = \pi, d = 8$，数据集加密过程中密钥变换6次，即每1/6的图像集使用同一固定密钥，λ由不同图像分别得出。分析结果如图5所示，其中第1行为MNIST图像，第2行为Fashion-MNIST图像。从图中可以看出，对只进行2D-SCL扩散的手写体图像分析后，破译图像与明文无明显差别，人眼可以清楚地分辨每个再生图像所代表的数字；服饰图虽然与原图有细节上的不同，但整体结构和外部轮廓也几乎全部还原，结构相似度很高，破译效果较好。

3)2D-SCL置乱分析。密文集生成时密钥初始值设置与2D-SCL扩散相同，分析结果如图5所示。从图中结果来看，破译图像有25%左右与原图一致，余下也破译出了部分结构，与明文不敏感的Arnold置乱相比，分析效果下降。

4)2D-SCL加密分析。密文集生成时密钥初始值设置与2D-SCL扩散相同，两图像集上的分析结果如图6所示。图6表明，进行完整2D-SCL加密的图像，40%左右的破译图像能恢复其对应明文的内容信息，其他的也存在某部分结构与明文相似。

5)量子bite置乱分析。使用本文的密文分析方法对量子加密方法的bite置乱阶段进行攻击,两图像集分析结果如图7(b)所示。其中第1行为MNIST图像,第2行为Fashion-MNIST图像。由破译结果可知,经bite置乱的两类图像均被轻易破解,破译图像与明文差别较小,破译效果较好,此阶段安全性低。

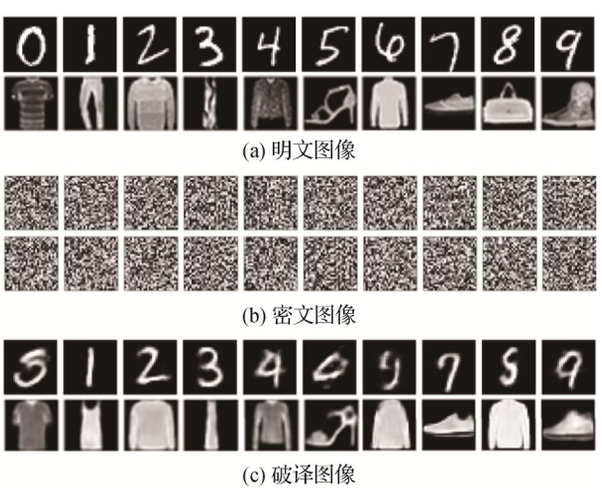
(a) 明文图像

(b) 扩散密文的破译结果

(c) 置乱密文的破译结果

图5　2D-SCL扩散—置乱分析效果

Fig. 5　Cryptanalysis results of diffusion and confusion cipher images of 2D-SCL encryption scheme ((a) plain images; (b) cryptanalysis results of diffusion cipher images; (c) cryptanalysis results of confusion cipher images)



(a) 明文图像

(b) 密文图像

(c) 破译图像

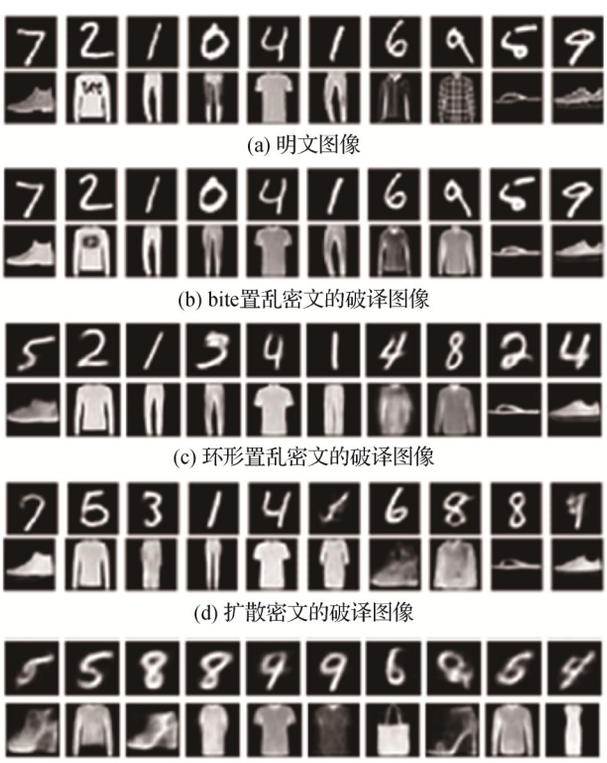
图6　2D-SCL加密分析效果

Fig. 6　Cryptanalysis results of cipher images of 2D-SCL encryption method ((a) plain images; (b) ciphertext images; (c) cryptanalysis results)

6)量子环形置乱分析。对该量子加密方法的像素循环移位即环形置乱阶段进行攻击,两图像集分析结果如图7(c)所示。由图像可知,破译明文有将近50%与原图一致,或破译出部分结构,相对bite置乱来说破译效果明显下降,但依旧存在安全问题。

7)量子扩散分析。使用本文的密文分析方法对量子加密方法的反馈扩散阶段进行攻击,分析结果如图7(d)所示。MNIST数据集上扩散密文的破译比例约为25%,余下图像也有部分结构能够还原;Fashion-MNIST数据集的破译效果相对前者有所上升,解码率大约在50%左右,此阶段依仍然是不安全的。

8)量子加密分析。使用本文的密文分析方法对量子加密方法进行攻击,分析结果如图7(e)所示。由破译图像可知,完整的量子加密密文破译效果较差,只有极少图像完成破解,此加密方法安全性相对较高,但仍存在安全缺陷。



(a) 明文图像

(b) bite置乱密文的破译图像

(c) 环形置乱密文的破译图像

(d) 扩散密文的破译图像

(e) 完整加密密文的破译图像

图7　量子加密各阶段密文的破译图像

Fig. 7　Cryptanalysis results of cipher images in various stages of quantum encryption ((a) plain images; (b) cryptanalysis results of bite scrambling cipher images; (c) cryptanalysis results of cycle scrambling cipher images; (d) cryptanalysis results of diffusion cipher images; (e) cryptanalysis results of quantum encryption cipher images)

2.3　PSNR和SSIM分析

计算测试集上明文图像与由密文分析方法生成的不同加密方法的破译明文之间的PSNR值和SSIM

值，2D-SCL加密和Arnold置乱的MPSNR、MSSIM、APSNR、ASSIM如表1所示，量子图像加密结果如表2所示，CDPSNR和CDSSIM曲线分别如图8和图9所示。

置乱密文的分析实验呈现不同的结果。对于明文不敏感的Arnold置乱来说，两个数据集的MPSNR都在30 dB左右，MSSIM也极接近1，APSNR、ASSIM都处于整体较高的水平，破译效果较好，表明Arnold置乱完全是不安全的。对于2D-SCL置乱方法的分析，虽然其APSNR和ASSIM与Arnold置乱有较大差距，平均水平偏低，但MPSNR和MSSIM仍旧较高，这表明部分图像仍然可以被成功破译，该置乱方法的混沌序列仍然存在安全性问题。CDPSNR和CDSSIM图也能够说明，明文敏感的置乱方法破译效果相对较差，抗攻击能力更强。

2D-SCL扩散密文的分析效果较好，如在两个测试集上的MPSNR均较高，相应的值分别高达32 dB和28 dB；而且相应的MSSIM值也高达0.99和0.98。除此之外，它们的APSNR和ASSIM值也相对较高，特别是在MNIST数据集上的ASSIM超过了0.9。这些数据表明本文提出的分析方法对2D-SCL方法的扩散密文具有较高分析能力，这一结果也从反面表明2D-SCL方法的扩散混沌序列的安全性较低。

关于2D-SCL加密方法的密文分析结果显示，两个测试集对应的MPSNR和MSSIM均较高，但是APSNR和ASSIM的值较低。这表明该方法整体破译效果较低，但少量密文依然能够被有效破译。此外，它们对应的累积分布曲线与2D-SCL置乱方法对应的累积分布曲线非常接近，表明二者的抗攻击能力高度近似。

表1　不同加密方法的分析结果

Table 1　Analysis results of different encryption methods

分析对象	MNIST数据集				Fashion-MNIST数据集			
	MPSNR/dB	APSNR/dB	MSSIM	ASSIM	MPSNR/dB	APSNR/dB	MSSIM	ASSIM
Arnold置乱	31.412 92	17.094 08	0.989 464	0.786 541	28.738 07	18.102 72	0.990 103	0.708 993
2D-SCL扩散	32.080 39	20.582 99	0.991 148	0.912 158	26.327 37	17.446 46	0.984 34	0.716 734
2D-SCL置乱	27.597 28	11.666 6	0.977 635	0.421 834	26.042 71	15.243 12	0.977 237	0.576 568
2D-SCL加密	25.629 21	11.375 26	0.964 604	0.388 644	23.760 23	12.130 48	0.964 478	0. 372 614

表2　量子加密不同阶段的分析结果

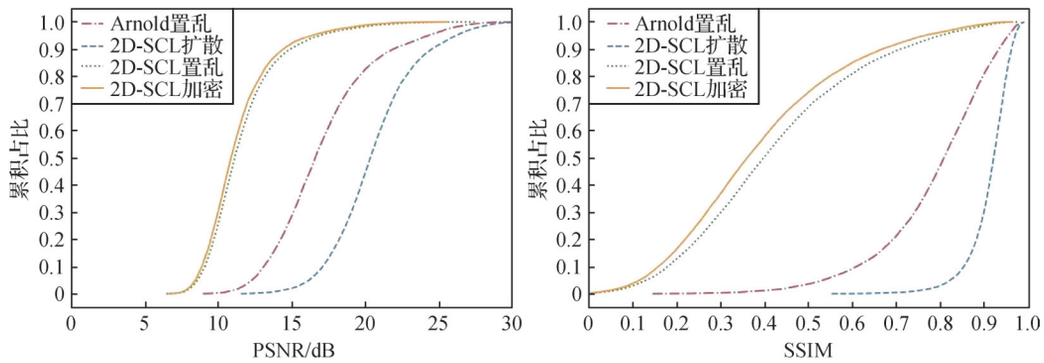
Table 2　Analysis results of different stage for quantum image encryption

分析对象	MNIST数据集				Fashion-MNIST数据集			
	MPSNR/dB	APSNR/dB	MSSIM	ASSIM	MPSNR/dB	APSNR/dB	MSSIM	ASSIM
bite置乱	32.244 5	23.232 5	0.993 4	0.950 7	29.741 4	18.866 7	0.992 9	0.769 3
环形置乱	25.694 6	11.270 4	0.972 5	0.383 5	25.612 2	14.933 7	0.973 5	0.557 8
反馈扩散	25.588 1	10.734 4	0.971 5	0.344 5	24.253 9	13.467 3	0.948 9	0.477 4
完整加密	22.840 6	10.196 2	0.891 4	0.233 8	21.496 2	9.444 1	0.933 1	0.196 1

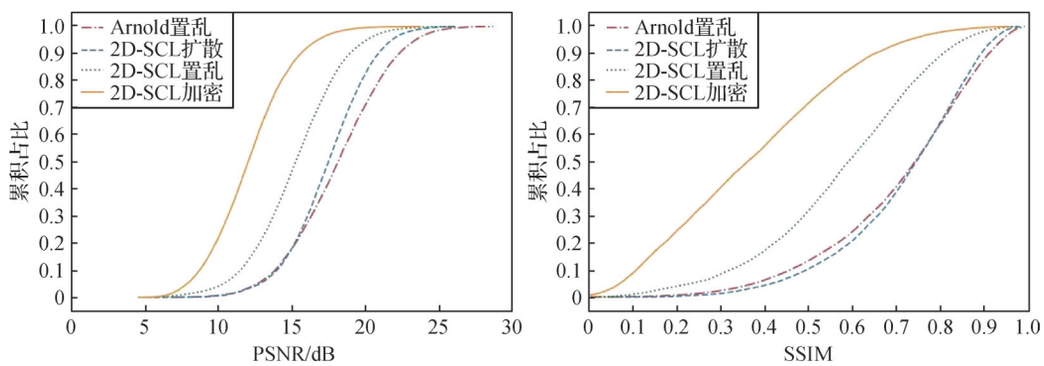
图8清晰地展现了以上4种不同密文的分析结果，累积分布曲线与横坐标轴间的面积越小，则破译效果越好，同时也侧面反映了其安全性越低。

相对而言，明文高度敏感的量子加密方法与前二者相比破译效果明显下降，安全性更高。其扩散阶段与2D-SCL扩散破译效果明显较低，如手写体集APSNR只有10.7 dB，ASSIM只有0.35左右，但是由于MPANR和MSSIM均高达25.7 dB和0.97，说明仍有部分密文能够被破译，服饰集测试数据也存在类似问题。

对于量子加密中两个不同的置乱阶段而言，在不同数据集上bite置乱的APSNR和ASSIM都较高，特别在MNIST数据集ASSIM高达0.95，说明基本所有图像都能被成功破解，安全性极低。像素环形置乱的破译结果相对较差，APSNR小于15 dB、ASSIM在0.5左右，由图9可知，两数据集SSIM在0.6以上



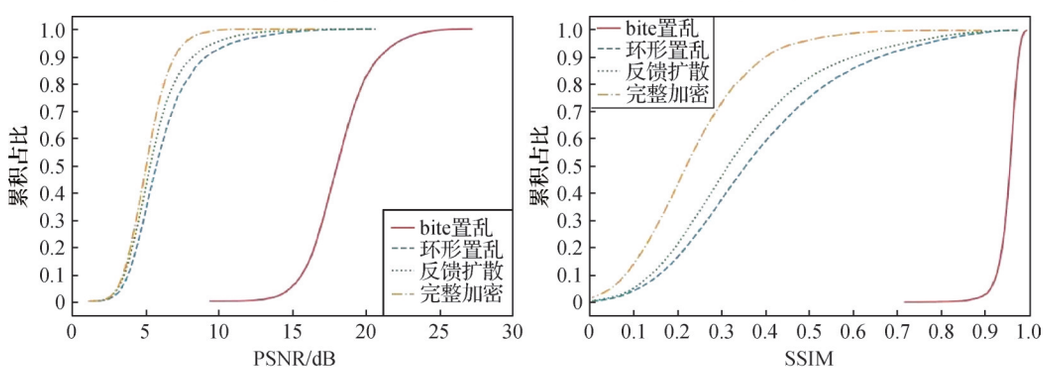
(a) MNIST数据集上结果



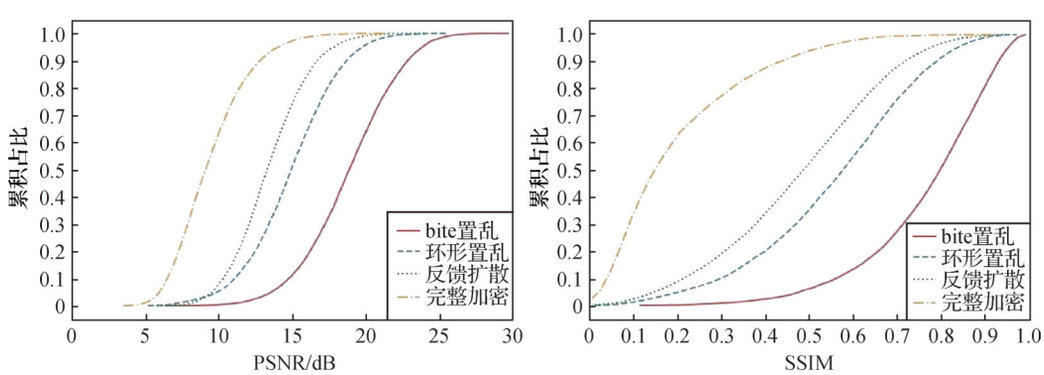
(b) Fashion-MNIST数据集上结果

图 8　CDPSNR 和 CDSSIM 曲线图

Fig. 8　Plot of CDPSNR and CDSSIM ((a) cryptanalysis results on MNIST dataset; (b) cryptanalysis results on Fashion-MNIST dataset)



(a) MNIST数据集上结果



(b) Fashion-MNIST数据集上结果

图 9　量子加密破译的 CDPSNR 和 CDSSIM 曲线图

Fig. 9　Plot of CDPSNR and CDSSIM for quantum image encryption

((a) cryptanalysis results on MNIST dataset; (b) cryptanalysis results on Fashion-MNIST dataset)

的只有20%和40%左右，说明大部分图像未被破解；但其MSSIM仍高达0.97左右，部分图像被还原为明文，其安全性仍存在问题。

完整的量子加密方法是所有密文中破译效果最差的，APSNR在10 dB左右，ASSIM在0.2左右，表明大部分密文图像不能被有效破译。但是其MSSIM在两数据集上分别是0.9与0.93，表明仍存在部分图像可以还原为明文或与明文高度相似，意味着该量子加密方法仍然存在安全缺陷。图9的CDPSNR和CDSSIM同样说明了上述问题的存在。

2.4 小结

综上所述，本文提出的密文图像分析模型能够破解明文不敏感的Arnold置乱和明文敏感的2D-SCL扩散，能学习到低维的结构特征即等价密钥，从而还原密文图像；而对于明文敏感的2D-SCL置乱和加密，破解率较低，但也能分析出部分关键信息，完成部分图像的破译。2D-SCL加密方法的扩散阶段被完全破解，严重影响了整体加密的有效性，置乱阶段也明显存在一定漏洞，这些都导致其无法抵抗攻击，整个加密算法存在明显的安全漏洞。

对于量子加密而言，其各个步骤都有图像能够被破解，存在一定的安全问题，特别是bite置乱部分被完全破解，加密性能低，也因此导致了经过完整量子加密的密文仍有少数被破译；但鉴于其密钥对明文高度敏感，其混沌序列的安全性较2D-SCL加密有很大上升，等价密钥的明文敏感性较强，模型较难学习到量子加密的等效密钥进而完成破译。

3 结 论

本文针对现有密文分析方法无法统一有效地评估加密算法的安全性问题，基于降噪自编码器提出了一种无需破解密钥，能综合、通用地评估混沌图像密码系统安全性的密文分析方法模型。使用已知具体加密方案的密码系统加密明文图像数据集，将密文图像集作为输入，原始图像集作为目标数据构建模型，首先通过编码器学习到低维抽象特征，然后通过解码器对此特征进行高维重建生成解密图像，实现密文分析和破译。在对不同的混沌图像加密算法进行分析时，只需根据加密算法改变扩散阶段、置乱阶段和完整加密阶段密文的生成方法即可，无需改变密文分析的训练阶段、测试阶段和模型本身，就可以完成对不同混沌加密算法的分析和评估。

实验结果表明，该方法通过综合考量直观的破译图像和客观的评价指标数据，能够有效评估加密方法的安全性；对于明文不敏感的加密方法都能成功破译且效果显著，然而对明文敏感的密文图像，模型只能学习到一定的特征信息即等效密钥，完成部分图像的破译，且破译效果很大程度上取决于密钥的明文敏感性。而神经网络模型多种多样，功能也不尽相同，如何调整模型结构使其能适应于复杂加密算法密文图像的分析破译，将是后续研究要关注的主要问题，这也将进一步推动混沌密码系统的发展。

参考文献（References）

Chen C, Sun K H and He S B. 2020. An improved image encryption algorithm with finite computing precision. Signal Processing, 168: #107340 [DOI: 10.1016/j.sigpro.2019.107340]

Dou Y Q and Li M. 2020. Cryptanalysis of a new color image encryption using combination of the 1D chaotic map. Applied Sciences, 10(6): #2187 [DOI: 10.3390/app10062187]

Farwa S, Muhammad N, Shah T and Ahmad S. 2017. A novel image encryption based on algebraic S-box and Arnold transform. 3D Research, 8(3): #26 [DOI: 10.1007/s13319-017-0135-x]

Fridrich J. 1998. Symmetric ciphers based on two-dimensional chaotic maps. International Journal of Bifurcation and Chaos, 8(6): 1259-1284 [DOI: 10.1142/S021812749800098X]

Hou B T, Li Y Q, Zhao H Y and Wu B. 2020. Linear attack on round-reduced DES using deep learning//Proceedings of the 25th European Symposium on Research in Computer Security. Guildford, UK: Springer: 131-145 [DOI: 10.1007/978-3-030-59013-0_7]

Hu M T, Li J Q and Di X Q. 2023. Quantum image encryption scheme based on 2D Sine^2-Logistic chaotic map. Nonlinear Dynamics, 111(3): 2815-2839 [DOI: 10.1007/s11071-022-07942-1]

Joshi A B, Kumar D, Mishra D C and Guleria V. 2020. Colour-image encryption based on 2D discrete wavelet transform and 3D logistic chaotic map. Journal of Modern Optics, 67(10): 933-949 [DOI: 10.1080/09500340.2020.1789233]

Kingma D P and Ba J. 2015. Adam: a method for stochastic optimization//Proceedings of the 3rd International Conference on Learning Representations. San Diego, USA: [s.n.]

Lecun Y, Bottou L, Bengio Y and Haffner P. 1998. Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11): 2278-2324 [DOI: 10.1109/5.726791]

Li M, Guo Y Z, Huang J and Li Y. 2018. Cryptanalysis of a chaotic image encryption scheme based on permutation-diffusion structure.

Signal Processing: Image Communication, 62: 164-172 [DOI: 10.1016/j.image.2018.01.002]

Li M, Wang P C, LIU Y F and Fan H J. 2019. Cryptanalysis of a novel bit-level color image encryption using improved 1D chaotic map. IEEE Access, 7: 145798-145806 [DOI: 10.1109/ACCESS.2019.2945578]

Li Y Z, Jin X, Zhao G, Li X D, Tian Y L and Wang Z Y. 2016. Color image encryption scheme based on Zigzag transformation and chaotic map. Computer Engineering and Design, 37(8): 2002-2006 (李玉珍, 金鑫, 赵耿, 李晓东, 田玉露, 王子亦. 2016. 基于Zigzag变换与混沌的彩色图像加密方案. 计算机工程与设计, 37(8): 2022-2006) [DOI: 10.16208/j.issn1000-7024.2016.08.005]

Liang Y and Zhang S W. 2018. Image encryption algorithm based on bit-level synchronous permutation diffusion and pixel-level annular diffusion. Journal of Image and Graphics, 23(6): 814-826 (梁颖, 张绍武. 2018. 位级同步置乱扩散和像素级环形扩散图像加密算法. 中国图象图形学报, 23(6): 814-826) [DOI: 10.11834/jig.170433]

Melicher W, Ur B, Segreti S M, Komanduri S, Bauer L, Christin N and Cranor L F. 2016. Fast, lean, and accurate: modeling password guessability using neural networks//The 25th USENIX Conference on Security Symposium. Austin, USA: USENIX Association: 175-191

Pak C, An K, Jang P, Kim J and Kim S. 2019. A novel bit-level color image encryption using improved 1D chaotic map. Multimedia Tools and Applications, 78(9): 12027-12042 [DOI: 10.1007/s11042-018-6739-1]

Pak C and Huang L L. 2017. A new color image encryption using combination of the 1D chaotic map. Signal Processing, 138: 129-137 [DOI: 10.1016/j.sigpro.2017.03.011]

Rehman A U, Liao X F, Ashraf R, Ullah S and Wang H W. 2018. A color image encryption technique using exclusive-OR with DNA complementary rules based on chaos theory and SHA-2. Optik, 159: 348-367 [DOI: 10.1016/j.ijleo.2018.01.064]

Solak E, Cokal C, Yildiz O T and Biyikoğlu T. 2010. Cryptanalysis of Fridrich's chaotic image encryption. International Journal of Bifurcation and Chaos, 20(5): 1405-1413 [DOI: 10.1142/S0218127410026563]

Sun X Y and Chen Z. 2021. A new image encryption strategy based on Arnold transformation and logistic map//Liu Q, Liu X D, Chen B, Zhang Y M and Peng J S, eds. Proceedings of the 11th International Conference on Computer Engineering and Networks. Singapore, Singapore: Springer: 721-729 [DOI: 10.1007/978-981-16-6554-7_77]

Vincent P, Larochelle H, Bengio Y and Manzagol P A. 2008. Extracting and composing robust features with denoising autoencoders//Proceedings of the 25th International Conference on Machine Learning. Helsinki, Finland: Association for Computing Machinery: 1096-1103 [DOI: 10.1145/1390156.1390294]

Wen H P. 2019. Cryptanalysis of Some Chaotic Ciphers. Guangzhou: Guangdong University of Technology (温贺平. 2019. 若干混沌加密算法的密码分析. 广州: 广东工业大学)

Wu J H, Xia W X, Zhu G L, Liu H, Ma L J and Xiong J P. 2021. Image encryption based on adversarial neural cryptography and SHA controlled chaos. Journal of Modern Optics, 68(8): 409-418 [DOI: 10.1080/09500340.2021.1900440]

Xiao H, Rasul K and Vollgraf R. 2017. Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms [EB/OL]. [2023-04-07]. https://arxiv.org/pdf/1708.07747.pdf

Zhang W, Yu H, Zhao Y L and Zhu Z L. 2016. Image encryption based on three-dimensional bit matrix permutation. Signal Processing, 118: 36-50 [DOI: 10.1016/j.sigpro.2015.06.008]

Zhou H, Xie H W, Zhang H and Zhang H T. 2021. Parallel remote sensing image encryption algorithm based on chaotic map and DNA encoding. Journal of Image and Graphics, 26(5): 1081-1094 (周辉, 谢红薇, 张昊, 张慧婷. 2021. 混沌系统和DNA编码的并行遥感图像加密算法. 中国图象图形学报, 26(5): 1081-1094) [DOI: 10.11834/jig.200344]

作者简介

常晓琦，女，硕士研究生，主要研究方向为图像信息安全和深度学习。E-mail：waile68@qq.com

游大涛，通信作者，男，副教授，主要研究方向为人工智能、医学图像处理和信息安全。E-mail：youdatao@163.com

王明合，男，硕士研究生，主要研究方向为图像安全和密文分析。E-mail：104754211944@henu.edu.cn

武相军，男，教授，主要研究方向为信息安全和复杂网络。E-mail：wuxj@henu.edu.cn