缩略图盲解密模型:利用改进生成对抗网络的免密钥图像重建

朱礼亚[1],孙雅娜[1],任 帅[2*],张刘坤[1],蒋东华[3]

(1. 长安大学电子与控制工程学院,陕西西安 710064;2. 长安大学信息工程学院,陕西西安 710064;
3. 中山大学计算机学院,广东广州 510006)

摘 要: 缩略图保持加密(Thumbnail-Preserving Encryption, TPE)是平衡云端图像隐私性和可用性的重要方法。为确保安全性,通常利用动态更新机制生成与原始图像信息相关联的密钥。针对图像信息非正常缺失或篡改,接收方无法获得正确的密钥进行解密的问题,设计了一种基于生成对抗网络的轻量化缩略图盲解密模型(Blind Decryption model based on Generative Adversarial Networks, BD-GAN)。通过对生成器、判别器和损失函数的改进和优化,提升解密图像的质量。生成器基于改进的U-Net网络,采用编码器-转码器-解码器的级联结构。在编码器和解码器中分别嵌入多尺度注意力融合模块(Multi-Scale Attention Fusion modules, MSAF),实现各尺度头部信息的跨层融合,在采样过程中最大限度保留图像细节,解决了多次采样和深层网络长距离依赖所导致的信息丢失问题。将U-Net网络的瓶颈层替换为由多个基础残差块级联堆叠构成的转码器,促进提取特征的高效传递与学习。为了解决局部判别器只输出最后一层局部特征,忽略中间层特征所含的多尺度语义与纹理信息的问题,改进并设计了具有多层特征反馈特性的局部判别器,返回多个中间层特征用于对抗训练。通过对不同尺度特征的逐层输出,提升判别器对多尺度纹理细节的感知能力。在损失函数优化方面,采用视觉感知损失、对抗损失和身份一致性损失的联合优化策略,将各损失项的线性组合作为优化目标,通过最小化目标函数进行训练,提升重建图像的视觉质量。实验结果表明,该模型能够对多种缩略图保持加密算法进行免密钥重建,解决了密钥有损情况下解密失败的问题。同时在不牺牲重建性能的前提下,具有更少的训练计算开销,有效地降低了部署成本。相比于采用U-Net网络和局部判别器的生成对抗网络模型,重建图像的峰值信噪比提升了0.632 8 dB,FID(Fréchet Inception Distance)性能提升了14.361 0。与去噪扩散概率解密模型相比,在保证重建效果的同时,参数量和计算量分别减少5.120×10^{7}和2.807×10^{10}以上。研究为通过身份验证和合法授权的用户提供了应急条件下的缩略图解密方案,提高了安全冗余度。

关键词: 盲解密模型;生成对抗网络;残差块;自注意力机制;局部判别器

基金项目: 国家自然科学基金(No.62372062)

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112(2026)03-1094-11

电子学报URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20250556

A Blind Decryption Model for Thumbnail-Preserving Encryption: Image Reconstruction without a Key via an Improved Generative Adversarial Network

ZHU Liya[1], SUN Yana[1], REN Shuai[2*], ZHANG Liukun[1], JIANG Donghua[3]

(1. *School of Electronics and Control Engineering, Chang'an University, Xi'an, Shaanxi* 710064, *China*;
2. *School of Information Engineering, Chang'an University, Xi'an, Shaanxi* 710064, *China*;
3. *School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, Guangdong* 510006, *China*)

Abstract: Thumbnail-preserving encryption (TPE) is an important method to balance the privacy and availability of images in cloud environments. In pursuit of superior security, secret keys are generally generated using the image-related information under the assistance of a dynamic update mechanism. However, receivers may fail to obtain the correct keys due to the information loss or tampering. To address this issue, a lightweight blind decryption model based on generative adversarial networks (BD-GAN) is proposed for TPE images. This work devotes to better reconstruction performance of the image through the improvement of the generator and discriminator, and optimization of the loss function. In our scheme, the generator is designed based on an improved U-Net architecture, and it employs a cascaded structure composed of encoder,

收稿日期:2025-06-26;录用日期:2026-03-06;责任编辑:覃怀银

*通信作者:任帅

transcoder, and decoder modules. Multi-scale attention fusion modules (MSAF) are embedded in the encoder and decoder, respectively, to achieve cross-layer fusion of multi-scale hierarchical information. By this means, image details are reserved to the maximum extent, and the problems of information loss caused by repeated sampling and long-range dependencies in deep networks are addressed. The bottleneck layer of the U-Net network is replaced with a transcoder composed of cascaded multiple residual blocks, so as to promote the efficient transmission and learning of extracted features. Moreover, to solve the problem that the local discriminator only utilizes the local features of the last layer and ignores the multi-scale semantic and texture information contained in the intermediate layer features, a local discriminator with multi-layer feature feedback is designed, which returns multiple intermediate layer features for adversarial training. Through the layer-by-layer output of features, the capability of perceiving the multi-scale textures is enhanced. In addition, a joint optimization strategy integrating visual perception loss, adversarial loss, and identity consistency loss is adopted to optimize the loss function. Specifically, the linear combination of these loss components serves as the optimization objective, and the model is trained by minimizing the overall loss to improve the visual quality of the reconstructed images. Experimental results show that the proposed model can perform for various TPE algorithms without any secret keys, and the failure of decryption can be avoided even if the secret keys are damaged. More importantly, on the premise of not degrading reconstruction performance, our model has lower training computational overhead and deployment cost. Compared with the traditional GAN employing a U-Net network and local discriminator, the peak signal to noise ratio (PSNR) values of the reconstructed images are improved by 0.632 8 dB, and the fréchet inception distance (FID) values are improved by 14.361 0. Notably, in comparison with the diffusion models, our model offers a satisfactory reconstructed image while reducing parameters and computational cost by more than 5.120×10^7 and 2.807×10^{10}, respectively. This study provides an effective emergency decryption scheme for users who have passed identity verification and legal authorization, which improves security redundancy.

Keywords: blind decryption model; generative adversarial network; residual blocks; self-attention mechanism; local discriminator

Foundation Item(s): National Natural Science Foundation of China (No.62372062)

0　引言

随着信息技术的快速发展和网络服务的日益完善,越来越多的公众利用iCloud、Google Drive等云存储平台进行个人数据存储与管理[1-3]。其中,记录个人生活的图像往往包含用户隐私信息,一旦这些信息因云端漏洞或恶意攻击而被非法获取,将会带来严重的信息安全问题。为了确保云端图像信息的机密性,研究者们利用加密技术将其转换为无视觉特征的类噪声图像[4-7],从而有效保护了用户隐私,但不可避免降低了图像的可用性。为了平衡云端图像的隐私性和可用性,Wright等人[8]于2015年提出了缩略图保持加密(Thumbnail-Preserving Encryption,TPE)方法。通过设置缩略块尺寸阈值对图像进行信息分级处理,保留大于缩略块尺寸的宏观结构作为粗略感知信息,使具备先验知识的用户能够获取一定的视觉可用性;同时隐藏低于阈值的微观细节,确保未授权方无法从粗略视觉内容中解析出隐私信息。

现有的TPE加密方法分为理想TPE[9-11]和近似TPE[12-15]两类。理想TPE是指密文图像和原始图像的缩略图完全一致,能够实现无损解密;近似TPE是指密文图像和原始图像的缩略图存在一定差异,不能无损解密[16]。Wright等人[8]仅对块内像素进行置乱操作,没有改变像素值,因而暴露了图像的部分特征,存在安全缺陷。Tajik等人[9]提出了一种基于替换—置乱框架的理想缩略图保持加密方法(Ideal-TPE, ITPE),但每次替换过程只能处理2个像素,连通性较差。为此,Zhao等人[10]将像素组长度增加到3,提出了一种基于排序加密的缩略图保持加密算法。进一步地,Zhang等人[11]提出了一种基于多像素和保持的缩略图保持加密算法,突破了向量长度的限制,显著提升了理想TPE方案的连通性。

在近似TPE方面,Marohn等人[12]采用动态改变像素值低有效位的方式,将像素值总和控制在预设的范围内。该方案生成密文图像的视觉保真度不高,且解密图像存在许多噪声点。Zhang等人[13]提出了一种高保真缩略图保持加密方法(High-Fidelity TPE, HF-TPE),有效地提升了加密图像的视觉保真度,但效率有待进一步优化。为解决JPEG图像的格式兼容问题,Chai等人[14]设计了一种基于自适应偏差嵌入的缩略图保持加密方法(TPE based on Adaptive Deviation Embedding, TPE-ADE),实现了解密与隐私可视化的平衡。Wang等人[15]利用可变中值边缘检测器和哈夫曼编码技术,提出了一种具有较高视觉自然度的缩略图保持加密方案(Thumbnail-Preserving images with High-Visual Naturalness, TP-HVN)。

为确保TPE方案的安全性,通常利用动态更新机制生成与原始图像信息(如拍摄时间、名称等)相关

联的密钥。在图像处理、传输和存储过程中，这些信息可能出现非正常缺失或错误修改等情况，导致接收方无法生成正确的密钥进行解密。针对这一问题，研究者们尝试利用盲解密和超分辨率模型从TPE图像中重建细节信息。Li等人[17]提出了一种基于Pix2Pix（Pixel-to-Pixel）网络的解密模型，但对于图像细节信息的还原度不足，存在模糊现象。部分超分辨率模型如SRGAN（Super-Resolution Generative Adversarial Network）[18]、STUNet（Swin Transformer U-Net）[19]的重建效果亦有限，原因在于SRGAN和STUNet属单一映射，对源域和目标域之间的分布差异有严格的要求，无法针对像素替换或置乱的TPE图像实现精细化重建。超分辨率模型SR3（Super-Resolution via Repeated Refinement）[20]和Jiang等人[21]提出的缩略图盲解密框架（Denoising Diffusion Cryptanalytic Model，DDCM）本质上均属于去噪扩散概率模型，通过迭代方法从随机噪声中逐步生成与原始图像一致的解密图像，实现高质量重建。但多次迭代去噪过程需要大量的训练计算开销，降低了模型在资源约束环境下的可用性和部署灵活性。

基于以上分析，本文从TPE图像解密质量与模型复杂度的均衡性出发，对生成对抗网络的生成器、判别器和损失函数进行改进和优化，设计了一种轻量化的缩略图盲解密模型（Blind Decryption model based on Generative Adversarial Networks，BD-GAN），旨在为通过身份验证和合法授权的用户提供密钥有损条件下的应急解决方案，提高了安全冗余度。本文主要贡献如下：

（1）为提升对图像高频信息的重建效果，将残差块和多尺度注意力融合模块集成到U-Net网络中，改进后的重建图像峰值信噪比提高了0.632 8 dB。

（2）为提高判别器对多尺度纹理细节的感知能力，改进了具有多层特征反馈特性的局部判别器，重建图像的FID（Fréchet Inception Distance）值提升了14.361 0。

（3）利用GAN具有的低复杂度特点，提出了一种轻量化的盲解密模型。相比于去噪扩散概率模型，在没有降低图像重建质量的前提下，模型的参数量和计算量分别减少5.120×10^{7}和2.807×10^{10}以上。

1 相关工作

1.1 缩略图保持加密

缩略图保持加密是格式保持加密的一种特殊形式。对于任意的密钥$K\in\{0,1\}^{\lambda}$和随机数$T\in\{0,1\}^{*}$，式（1）都成立[22]：

$$\begin{cases}\mathrm{Enc}_K(T,\boldsymbol{M})=\boldsymbol{C}\in\mathbb{M}\\ \mathrm{Dec}_K(T,\boldsymbol{C})=\boldsymbol{M}\\ \Phi(\boldsymbol{C})=\Phi(\boldsymbol{M})\end{cases}\tag{1}$$

其中，$\boldsymbol{M}$为原始图像，$\boldsymbol{C}$为密文图像，$\mathrm{Enc}_K(\cdot)$和$\mathrm{Dec}_K(\cdot)$分别表示加密和解密算法，Φ为格式保持函数，$\Phi(\boldsymbol{M})$表示原始图像$\boldsymbol{M}$的缩略图。

在经典的ITPE算法中，原始图像按照R、G、B通道划分成大小为$n\times n$的图像块，再对每个块进行像素组长度为2的多轮替换-置乱加密。替换操作在保持像素组和不变的同时，破坏了图像的局部统计特征，提高了安全性。具体过程如下：

$$\mathrm{rank}_s(a,b)=\begin{cases}a & \text{if } s<d\\ d-a & \text{otherwise}\end{cases}\tag{2}$$

$$\mathrm{rank}_s^{-1}(r)=\begin{cases}(r,s-r) & \text{if } s<d\\ (d-r,s-d+r) & \text{otherwise}\end{cases}\tag{3}$$

其中，a和b表示两个像素值，d表示像素取值范围（通常为255），函数$\mathrm{rank}_s(a,b)$用于确定向量(a,b)的序列号r，$\mathrm{rank}_s^{-1}(r)$将r映射为对应的像素组。

置乱则是对替换后的块内像素进行Fisher-Yates洗牌操作，从而打乱块内像素的位置。为提高安全性，TPE加密算法通常将拍摄时间、图像名称等相关信息充当随机数T，并在置乱和替换的过程中与用户密钥K进行拼接，生成密钥，实现“一图一密”。而当这些信息出现非正常缺失、错误修改或篡改等情况，会导致接收方无法生成正确的密钥进行解密，如图1所示。

(a) 原始图像 (b) 正确的解密图像 (c) 图像名称被篡改后的解密图像

(a) Original image (b) Correctly decrypted image (c) Decrypted image when the name is tampered with

图1 图像信息修改后重建人脸图像的对比结果

Figure 1 Comparison results of reconstructed facial images after image information modification

本文的解密实验、消融实验及对比实验均基于ITPE加密算法进行设计与实施，具体过程详见第3节。

1.2 生成对抗网络

Goodfellow等人[23]于2014年提出了生成对抗网络（Generative Adversarial Network，GAN），主要思想是

生成器和判别器进行“动态博弈”。一方面,生成器根据判别器的反馈对网络参数进行优化,使得生成的样本数据更加符合真实特征;另一方面,判别器试图将生成的数据和真实数据区分开来。GAN的训练过程如下:

$$\min_G \max_D V(D,G)=E_{x\sim P_{data}(x)}\left[\log\left(D(x)\right)\right]+E_{z\sim P_z(z)}\left[\log\left(1-D\left(G(z)\right)\right)\right] \tag{4}$$

其中,E表示期望,$D(x)$表示判别器D判断x为真实数据的概率,$D\left(G(z)\right)$则表示判别器D判断由生成器G生成的样本$G(z)$为真实数据的概率,$\left(1-D\left(G(z)\right)\right)$表示$D$判断生成样本为假的概率。在训练过程中,生成器$G$通过最小化$\log\left(1-D\left(G(z)\right)\right)$提升生成样本的真实度,而判别器$D$则需最大化$\log\left(D(x)\right)$与$\log\left(1-D\left(G(z)\right)\right)$的联合期望,增强对生成样本的鉴别能力。这种对抗性优化机制驱动双方在迭代中逼近纳什均衡。

本文提出的缩略图盲解密模型以生成对抗网络为核心架构,利用生成器与判别器的对抗式训练实现图像信息的免密钥重建。具体过程详见第2.1节。

1.3　自注意力机制

注意力机制的设计灵感来源于人类对外部信息的处理过程,即在面对海量信息时,人脑会有选择地关注其中某一部分,同时忽略其他信息。Vaswani等人[24]于2017年提出了自注意力机制并集成在Transformer模型中,自注意力机制可视为注意力机制的变体,通过在序列内部动态分配注意力权重,实现了更精细的长距离依赖建模,计算过程为

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_k}}\right)\boldsymbol{V} \tag{5}$$

其中,$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$分别为查询(Query)、键(Key)和值(Value)的权重矩阵,d_k为键向量的维度。在图像处理过程中,自注意力机制通过计算注意力权重矩阵衡量每个像素与其他像素之间的关联程度。通过这种方式,模型能够评估每个像素在全局中的重要性,从而构建出经过加权处理的图像。

本文将多尺度注意力融合模块集成到生成器中,利用自注意力机制在各尺度头部间实现跨层信息融合,有效地保留图像细节并增强网络的表征能力,具体过程详见第2.2节。

2　方法

2.1　解密模型

本研究构建了一个基于生成对抗网络的盲解密框架,适用于人脸、动物、建筑物、生活用品等多种类别加密图像,系统架构如图2所示。BD-GAN模型由生成器和判别器两个核心组件构成,其中生成器采用编码器-转码器-解码器的级联结构。潜在特征编码器通过多层卷积神经网络提取加密图像的特征,转码器通过深度残差学习实现加密特征到明文图像特征的映射与增强,解码器则通过反卷积操作将特征逐步重建为像素空间上可视化的图像。判别器通过空间条件约束与多层特征反馈,分析生成图像与真实图像的分布差异,引导生成器生成高质量的解密图像。

本文的改进思路如下:(1)为在采样过程中最大程度地保留图像细节,在编码器和解码器中分别嵌入多尺度注意力融合模块;(2)为促进特征的高效传递和深度学习,将传统U-Net网络中的瓶颈层替换成由多个基础残差块构建的转码器;(3)为提高判别器的判别能力,返回判别器的多个中间层特征用于对抗训练。

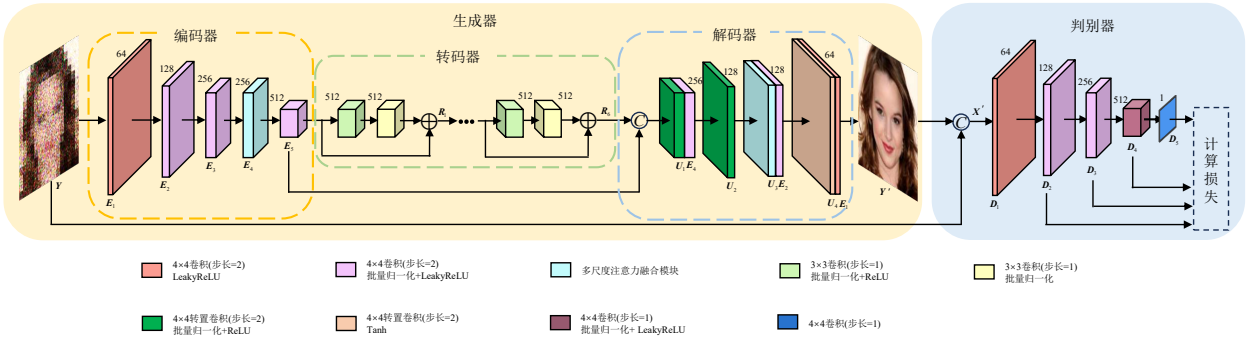


图2　BD-GAN网络结构

Figure 2　The network architecture of BD-GAN

2.2　生成器

生成器通过对加密图像的多阶段处理进行重建,其整体结构包括编码器、转码器和解码器,各模块协同作用,实现了从局部特征到全局特征的高效捕捉与恢复。

2.2.1　编码器

编码器对输入的加密图像进行逐层下采样和特征提取,如式(6)~(9)所示。加密图像经过下采样,

降低图像空间分辨率的同时扩大通道数，输出的特征再经过 LeakyReLU 激活函数以增强模型的非线性表达能力，随后经过批量归一化防止梯度消失和梯度爆炸。

$$\boldsymbol{E}_1=\mathrm{LeakyReLU}\left(\mathrm{conv}(\boldsymbol{Y})\right) \tag{6}$$

$$\boldsymbol{E}_i=\mathrm{LeakyReLU}\left(\mathrm{BN}\left(\mathrm{conv}(\boldsymbol{E}_{i-1})\right)\right) \tag{7}$$

其中，$\boldsymbol{E}_1$、$\boldsymbol{E}_i$、$\boldsymbol{E}_{i-1}$（$i=2,3$）表示解码器提取的图像特征，$\boldsymbol{Y}$表示加密图像。

将提取的特征$\boldsymbol{E}_3$输入多尺度注意力融合模块[25]（Multi-Scale Attention Fusion modules，MSAF），如图 3 所示。首先，$\boldsymbol{E}_3$通过层归一化处理后，将得到的特征$\boldsymbol{F}_1$沿通道维度划分为H个多尺度头部。其中，最高层头部保留原始空间分辨率，避免下采样导致图像的细节模糊。逐级向下头部的池化强度逐步增强，空间尺寸依次减半，感受野不断扩大，编码器能够捕捉更广泛的上下文信息。每个头部经过独立的自注意力机制处理后，输出特征经上采样至相邻上层头部，并与该头部特征逐元素相加，实现跨层特征融合，指导生成器对图像细节信息进行重建。融合后的上层头部特征再次执行自注意力计算，且向上传递直至最高层头部。完成所有头部处理后，将各头部的输出恢复至原始分辨率并在通道维度拼接，并通过 1 × 1 卷积和 GELU（Gaussian Error Linear Unit）激活函数实现通道融合，得到多尺度融合特征$\boldsymbol{F}_2$。$\boldsymbol{F}_2$与$\boldsymbol{F}_1$进行通道乘法后，得到的$\boldsymbol{F}_3$与$\boldsymbol{E}_3$进行融合得到初步增强特征$\boldsymbol{F}_4$。最后，$\boldsymbol{F}_4$经过层归一化后，通过由 3 × 3 卷积、GELU 激活函数和 1 × 1 卷积级联构成的细化模块，生成局部增强特征$\boldsymbol{F}_5$并与$\boldsymbol{F}_4$进行特征融合，完成一次缩放点积自注意力处理。重复上述计算过程N次后，再与$\boldsymbol{E}_3$相加，实现特征的保留与增强，如下式所示：

$$\boldsymbol{E}_4=\mathrm{MSAF}(\boldsymbol{E}_3) \tag{8}$$

$$\boldsymbol{E}_5=\mathrm{LeakyReLU}\left(\mathrm{BN}\left(\mathrm{conv}(\boldsymbol{E}_4)\right)\right) \tag{9}$$

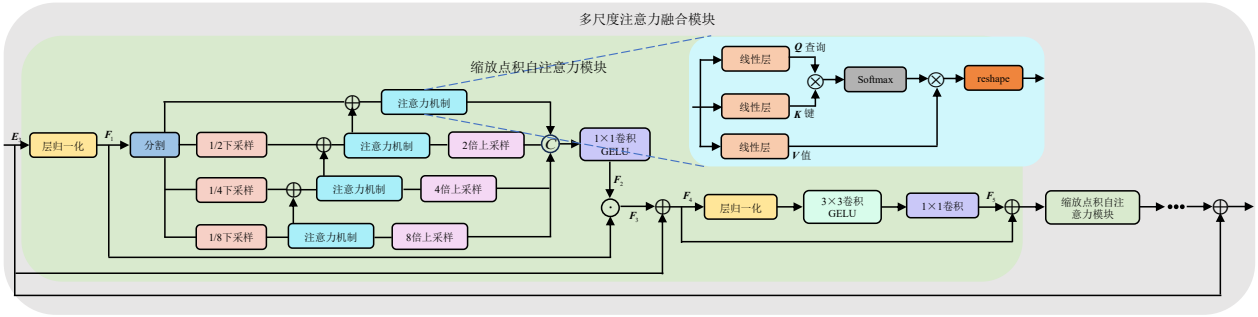


图 3 多尺度注意力融合模块结构

Figure 3 Structure of the multi-scale attention fusion modules

2.2.2 转码器

本文将传统 U-Net 网络[26]中的瓶颈层替换为由多个基础残差块（Residual Blocks，ResBlocks）级联堆叠构成的转码器，作为生成器的核心特征转换模块。通过非线性变换，将编码器提取的加密特征逐步解耦成与明文图像相关的语义特征。每个残差块由两个卷积层构成，并采用跨层跳跃连接，如式（10）所示：

$$\boldsymbol{R}_j=\mathrm{BN}\left(\mathrm{conv}\left(\mathrm{ReLU}\left(\mathrm{BN}\left(\mathrm{conv}(\boldsymbol{R}_{j-1})\right)\right)\right)\right)+\boldsymbol{R}_{j-1} \tag{10}$$

其中，$\boldsymbol{R}_j(j=1,2,\cdots,6)$为转码器提取的特征。

2.2.3 解码器

解码器通过上采样操作将转换器输出的特征逐步恢复到原始尺寸，并重构出细节丰富的图像，如式（11）~（15）所示：

$$\boldsymbol{U}_1=\mathrm{ReLU}\left(\mathrm{BN}\left(\mathrm{ConvTrans}\left(\mathrm{Cat}(\boldsymbol{R}_6,\boldsymbol{E}_5)\right)\right)\right) \tag{11}$$

$$\boldsymbol{U}_2=\mathrm{ReLU}\left(\mathrm{BN}\left(\mathrm{ConvTrans}\left(\mathrm{Cat}(\boldsymbol{U}_1,\boldsymbol{E}_4)\right)\right)\right) \tag{12}$$

$$\boldsymbol{U}_3=\mathrm{MSAF}(\boldsymbol{U}_2) \tag{13}$$

$$\boldsymbol{U}_4=\mathrm{ReLU}\left(\mathrm{BN}\left(\mathrm{ConvTrans}\left(\mathrm{Cat}(\boldsymbol{U}_3,\boldsymbol{E}_2)\right)\right)\right) \tag{14}$$

$$\boldsymbol{Y}'=\mathrm{Tanh}\left(\mathrm{Conv}\left(\mathrm{Cat}(\boldsymbol{U}_4,\boldsymbol{E}_1)\right)\right) \tag{15}$$

其中，Cat(·)为拼接函数，$\boldsymbol{U}_1$、$\boldsymbol{U}_2$、$\boldsymbol{U}_3$、$\boldsymbol{U}_4$为解码器恢复的特征，$\boldsymbol{Y}'$为重建图像。拼接函数将编码器提取的特征与解码器恢复的特征进行拼接，确保提取的多尺度特征能够直接参与上采样过程。

2.3 判别器

文献[27]提出的局部判别器将输入的特征图划分为多个重叠的局部区域进行训练，虽然提升了局部纹理的真实性，但会丢失中间层特征所含的多尺度语义与纹理信息。为此，本文将其改进为具有多层特征反馈特性的局部判别器，利用输出的多个中间层特征用于训练，从而提升判别器对多尺度纹理细节的感知能力。为确保生成器和判别器在训练时关注到一致的目标，将加密图像作为条件输入到判别器中，引导生成器生成符合条件的图像，从而减少生成图像和原始图像的结构差异。拼接后的图像$\boldsymbol{X}'$经过初始卷积

模块进行初级特征提取,完成空间下采样,如式(16)所示:

$$\boldsymbol{D}_1=\mathrm{LeakyReLU}\left(\mathrm{Conv}(\boldsymbol{X}')\right) \tag{16}$$

其中,$\boldsymbol{D}_1$为经过初始卷积模块提取的特征。在此基础上,$\boldsymbol{D}_1$通过四个递进式卷积模块,每个卷积核大小为4×4且均不使用偏置项,具体过程如下:

$$\boldsymbol{D}_k=\mathrm{LeakyReLU}\left(\mathrm{BN}\left(\mathrm{Conv}(\boldsymbol{D}_{k-1})\right)\right) \tag{17}$$

$$\boldsymbol{D}_5=\mathrm{Conv}(\boldsymbol{D}_4) \tag{18}$$

其中,$\boldsymbol{D}_k(k=2,3,4)$和$\boldsymbol{D}_5$为卷积模块提取的特征,用于判别器的对抗训练。与文献[27]仅输出最后一层特征$\boldsymbol{D}_5$的局部判别器相比,本文改进的多层特征反馈局部判别器返回多个中间层特征$\boldsymbol{D}_k$用于训练。前两个卷积模块采用步长为2的卷积操作,将通道数扩展至64和128;后两个模块的步长为1,在保持通道数为256和512的同时保留关键信息。

2.4　损失函数

本文使用视觉感知损失L_{per}、对抗损失L_{adv}和身份一致性损失L_{id}的联合优化策略设计损失函数。将各损失项的线性组合作为生成对抗网络的优化目标,并最小化目标函数来训练模型,公式如下:

$$\min_{\varphi} L_{\mathrm{total}}=L_{\mathrm{per}}+L_{\mathrm{adv}}+L_{\mathrm{id}} \tag{19}$$

其中,每项损失具体如下文所述。

2.4.1　视觉感知损失

视觉感知损失旨在鼓励生成器的输出在视觉上近似于原始图像。本文对文献[21]提出的视觉感知损失进行改进,引入基于VGG(Visual Geometry Group)网络的感知损失L_{VGG}。具体而言,该损失函数由矩阵范数、多尺度结构相似性损失和感知损失三部分构成。其中,矩阵范数确保重建图像与原始图像在像素级上保持一致性;多尺度结构相似性损失旨在增强模型对图像整体信息的捕捉能力;感知损失通过特征空间中的均方误差约束网络,以保留图像的语义信息与风格特征。公式如下:

$$L_{\mathrm{per}}=\lambda\left\|\boldsymbol{X}-\boldsymbol{Y}'\right\|_1+2\lambda\cdot\left(1-\mathrm{ms\_ssim}(\boldsymbol{X},\boldsymbol{Y}')\right)+2\lambda\cdot L_{\mathrm{VGG}} \tag{20}$$

$$L_{\mathrm{VGG}}=\mathrm{mse}\left(\mathrm{VGG}(\boldsymbol{X}),\mathrm{VGG}(\boldsymbol{Y}')\right) \tag{21}$$

其中:λ为预设加权系数,用于调节各项损失的权重;ms_ssim(·)为多尺度结构相似性指数函数;mse(·)为均方误差损失函数;VGG(·)为预训练网络[28]。

2.4.2　对抗损失

通过对抗损失建立动态博弈机制,引导生成器生成高质量图像。对抗损失L_{adv}由判别器损失L_D和生成器损失L_G两部分构成,如下式所示:

$$L_{\mathrm{adv}}=L_D+L_G \tag{22}$$

其中,

$$L_D=\frac{1}{2}\left\{\sum_i E_{(\boldsymbol{X},\boldsymbol{Y})}\left[\log D_i(\boldsymbol{X},\boldsymbol{Y})\right]+\sum_i E_{\boldsymbol{X}}\left[\log\left(1-D_i\left(\boldsymbol{X},G(\boldsymbol{X})\right)\right)\right]\right\} \tag{23}$$

$$L_G=\sum_i E_{\boldsymbol{X}}\left[\log D_i\left(\boldsymbol{X},G(\boldsymbol{X})\right)\right] \tag{24}$$

其中,$(\boldsymbol{X},\boldsymbol{Y})$表示密文-明文图像对。

2.4.3　身份一致性损失

身份一致性损失作为解密模型训练过程中的关键部分,指导生成器从给定的加密图像中恢复出原始身份信息。本文通过计算原始图像$\boldsymbol{X}$与生成图像$\boldsymbol{Y}'$的特征向量余弦相似度来定义L_{id},公式如下:

$$L_{\mathrm{id}}=1-\frac{f(\boldsymbol{X})\cdot f(\boldsymbol{Y}')}{\|f(\boldsymbol{X})\|_2\cdot\|f(\boldsymbol{Y}')\|_2} \tag{25}$$

其中,$f(\cdot)$表示FaceNet提取身份嵌入向量[29]。

3　实验

3.1　实验设置

本文使用公开的数据集CelebA-HQ、AFHQ、Places365和Kitchenware数据集进行模型训练。其中,CelebA-HQ包含30 000张分辨率为1 024×1 024的高清人脸图像;AFHQ包含10 727张分辨率为512×512的猫和狗的面部图像;Places365包含1.8×10^6张分辨率为256×256的多类别图像,本文选取风景和建筑物两类数据集;Kitchenware包含6类日常生活用品。为提高训练效率,从每类数据集中随机选取不重叠的8 000张用于训练,1 000张用于验证,1 000张用于测试。为加速模型收敛并统一输入尺寸,所有图像均通过三次样条插值法统一降采样至128×128。

在模型设计中,基础残差块的数量设置为6,缩放点积自注意力模块的数量N以及多尺度头部个数H都设置为4,加权系数λ为9。使用$\beta1$=0.5和$\beta2$=0.999的Adam优化器来训练模型,训练时生成器和判别器的学习率为2×10^{-4},批量大小设置为8以加快模型训练。实验在配备NVIDIA RTX 4060 GPU的笔记本电脑上进行,并基于PyTorch2.0.0平台完成。

3.2　性能分析

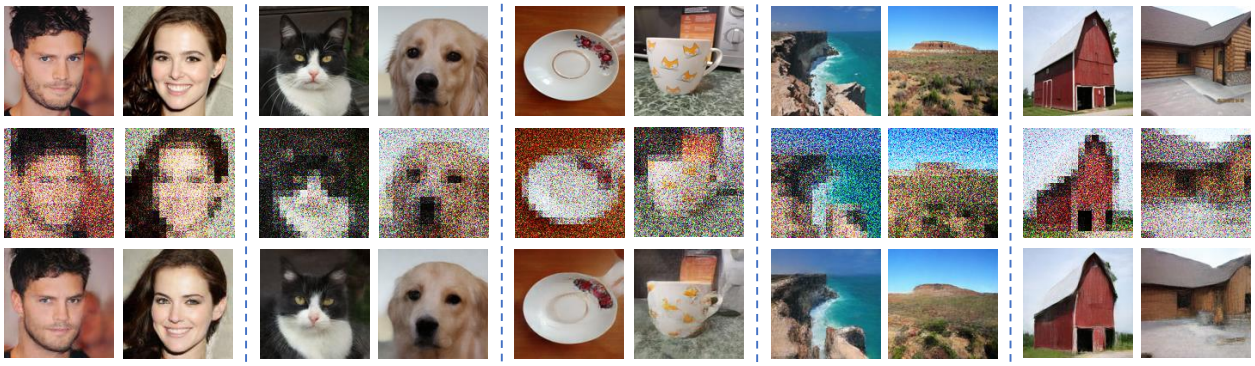
3.2.1　实验结果

本文以8×8的像素块为基本单位,使用ITPE[9]加密算法构建不同类别的数据集,利用峰值信噪比(Peak Signal to Noise Ratio,PSNR)和结构相似性指数(Structural Similarity Index,SSIM)两个指标进行定量评估。其中,PSNR用于衡量图像之间的像素级差异,SSIM用于衡量图像在亮度、对比度和结构上的相似性。PSNR值越大且SSIM值越接近于1,表明两幅图像越相似。

图4给出了不同类别的TPE图像可视化重建结

果，图中第一行为原始图像，第二行为ITPE图像，第三行为重建图像，在视觉效果上重建图像与原图相似。表1给出了不同类别测试集的重建性能，因数据本身的统计特征与结构复杂度的影响，不同类别数据的重建性能存在差异。其中，人脸图像具有高度一致的语义结构，且面部轮廓等特征分布相对规律，模型更易学习到稳定的映射；而动物的局部结构与人脸存在一定的相似性，但其高密度、不规则的毛发纹理，导致高频信息在解密过程中失真较为严重；生活用品因其形态和尺寸的多样性，导致模型难以学习到统一的映射规律；风景和建筑物通常包含复杂的全局布局，卷积核无法充分捕捉空间长程依赖，导致模型难以精确重建。综上分析，BD-GAN模型在解密结构复杂的图像时，重建性能略有下降，但仍能保留主体结构与语义信息。因此该模型可应用于不同类别TPE图像的盲解密。

考虑到人脸图像中包含大量的隐私信息，且对个人隐私保护的需求也更为迫切，本文主要围绕人脸图像开展实验，使用多种缩略图保持加密算法HF-TPE[13]、TPE-ADE[14]和TP-HVN[15]构建人脸图像数据集，评估模型在不同加密策略下的盲解密性能。图5给出了部分密文图像可视化重建结果，其中“X8”表示采用8×8像素块进行加密处理。在视觉效果上，重建的人脸图像与原始图像具有高度的相似性。



(a) 人脸 (b) 动物 (c) 生活用品 (d) 风景 (e) 建筑物

(a) Facial images (b) Animal images (c) Commodity images (d) Landscape images (e) Building images

图4 BD-GAN重建不同类别图像的可视化结果

Figure 4 Visualization results of reconstructed images from different categories via BD-GAN

表1 不同类别图像重建的性能结果

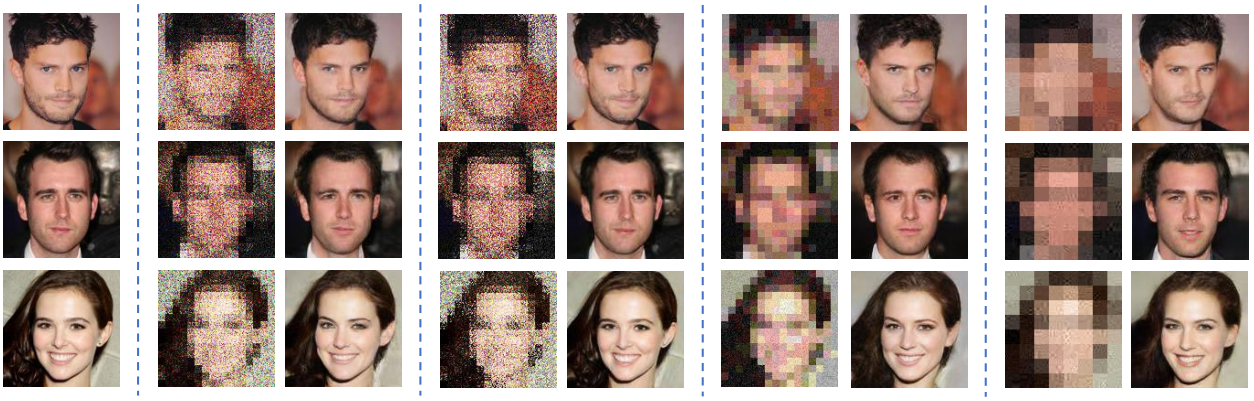
Table 1 Reconstruction performance for different image categories

数据集	PSNR/dB	SSIM
人脸	23.538 6	0.826 4
动物	22.425 1	0.722 2
生活用品	23.387 6	0.764 7
风景	22.326 2	0.703 1
建筑物	21.648 0	0.702 8

为进一步评估生成的人脸图像对原始身份信息的保留程度，本文还引入了平均FaceID指标：利用预先训练的人脸识别模型VGGFace提取人脸深度特征向量[30]，并计算余弦相似度，值越高表明重建人脸在身份层面的保真度越高。表2给出了测试集上各性能指标的平均值，在相同条件下，HF-TPE的解密效果在PSNR、SSIM和平均FaceID上均优于TPE-ADE和TP-HVN。原因在于：HF-TPE为提高解密的正确率和视觉质量，加密图像最大限度地保留了图像的结构信息与局部相关性，为BD-GAN网络提供了可解码的相关信息。相比之下，TPE-ADE在对图像进行加密时，将量化表中所有高频系数的量化步长设置为1，致使加密图像的频谱特征相似，导致BD-GAN网络无法依据不同频段的特征恢复图像细节。使用TP-HVN算法加密时，对图像块进行平均化处理并进行非线性比特嵌入，加密比特被均匀散布到所有位平面，扰乱了高频信息并掩盖了关键重建线索，模型难以区分真实纹理与嵌入噪声。综上分析，BD-GAN模型在多种TPE算法下均表现出良好的重建性能。尽管对部分算法的重建效果稍差，但依然能够恢复出包含原始身份信息的关键特征。实验结果验证了该模型适用于多种人脸TPE算法的盲解密。

3.2.2 参数设置

本小节探讨加权系数λ对重建图像质量的影响，针对ITPE图像，图6展示了λ值与重建图像性能指标之间的关系。分析可得，当加权系数取值为9时，重建图像的PSNR和SSIM值达到最大，随着取值的增大，重建图像的PSNR值不断下降。由此可见，选取合适的加权系数(λ=9)可有效提升重建图像的PSNR值，过大或过小的权重都会降低重建图像的质量。



(a) 原始图像　(b) ITPE(X8)图像(左)和重建图像(右)　(c) HF-TPE(X8)图像(左)和重建图像(右)　(d) TP-HVN(X8)图像(左)和重建图像(右)　(e) TPE-ADE(X16)图像(左)和重建图像(右)

(a) Original images　(b) The encrypted images by ITPE algorithm with 8×8 pixel blocks (left) and the corresponding reconstructed images (right)　(c) The encrypted images by HF-TPE algorithm with 8×8 pixel blocks (left) and the corresponding reconstructed images (right)　(d) The encrypted images by TP-HVN algorithm with 8×8 pixel blocks (left) and the corresponding reconstructed images (right)　(e) The encrypted images by TPE-ADE algorithm with 16×16 pixel blocks (left) and the corresponding reconstructed images (right)

图5　BD-GAN重建人脸图像的可视化结果

Figure 5　Visualization results of reconstructed facial images via BD-GAN

表2　重建人脸图像的性能结果

Table 2　Reconstruction performance of facial images

方法	PSNR/dB	SSIM	平均FaceID/%
ITPE(X8)	23.538 6	0.826 4	60.81
HF-TPE(X8)	30.439 8	0.927 5	95.48
HF-TPE(X16)	29.170 6	0.908 0	93.55
TP-HVN(X8)	22.306 7	0.777 3	51.92
TPE-ADE(X16)	20.274 4	0.761 9	40.62

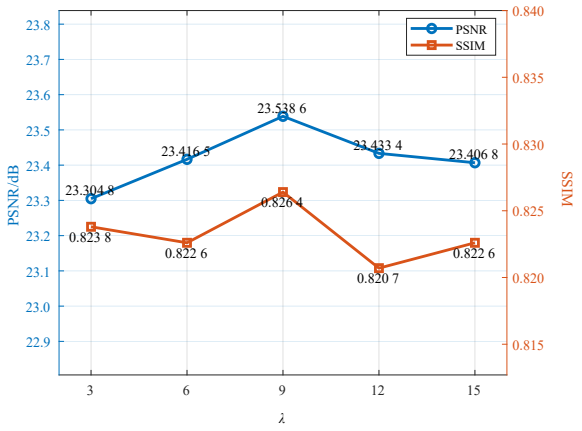


图6　加权系数与重建图像的实验结果

Figure 6　PSNR and SSIM values of reconstructed images with different weighting coefficients

3.3　消融实验

在本节中,通过消融实验评估各模块对模型性能的贡献。实验所用训练集与测试集均基于CelebA-HQ构建,并采用ITPE算法在8×8像素块下生成加密图像。

3.3.1　残差块数量

本节在未引入多尺度注意力融合模块的条件下,通过调整转码器中残差块的数量进行了定量分析。图7从残差块数量及其对应的生成器模型参数两个维度,对人脸图像重建任务中的性能指标进行了对比。当残差块的数量从0逐步增加到12时,生成器的参数量由1.063×10^7增加到6.492×10^7,同时重建图像的PSNR和SSIM值分别提升了0.509 0 dB和0.018。值得注意的是,当残差块数量增至15时,参数量为7.909×10^7,PSNR和SSIM两项指标均有所下降。表明适量的残差块能够有效增强网络的非线性表达能力,从而提高图像重建质量。但是当网络结构复杂度超出最优阈值后会导致过拟合,模型泛化能力也会随之下降。

通过对图像生成质量和模型复杂度指标的综合评估,最终选取6个基础残差块构建转码器。在保证图像重建质量的前提下,可满足实际应用场景中对计算资源与图像生成质量的平衡需求。

3.3.2　多尺度注意力融合模块

为了在下采样与上采样过程中最大程度地保留高频细节,本文在生成器中嵌入多尺度注意力融合模块,在高分辨率特征与低分辨率特征之间传递信息,从而增强模型的重建能力。表3给出了生成器中各模块消融实验结果,与采用U-Net网络的原始模型相比,仅在解密模型中加入多尺度注意力融合模块时,重建图像的PSNR值提高了0.460 9 dB,SSIM提高了

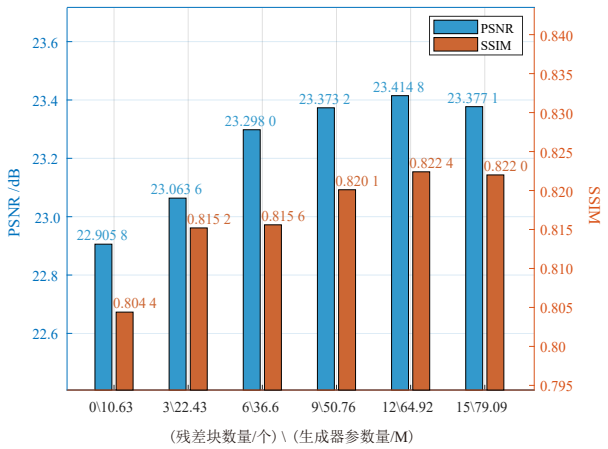


图7 残差块数量及生成器参数量与重建图像的实验结果

Figure 7 Quality of reconstructed images versus the number of residual blocks and generator parameters

0.011 5；当同时加入残差块和多尺度注意力融合模块时，PSNR 和 SSIM 值分别提升了 0.632 8 dB 和 0.022。上述结果表明，多尺度注意力融合模块有效地解决了多次采样和深层网络长距离依赖所导致的信息丢失问题，提高了重建图像的质量。

3.3.3 判别器

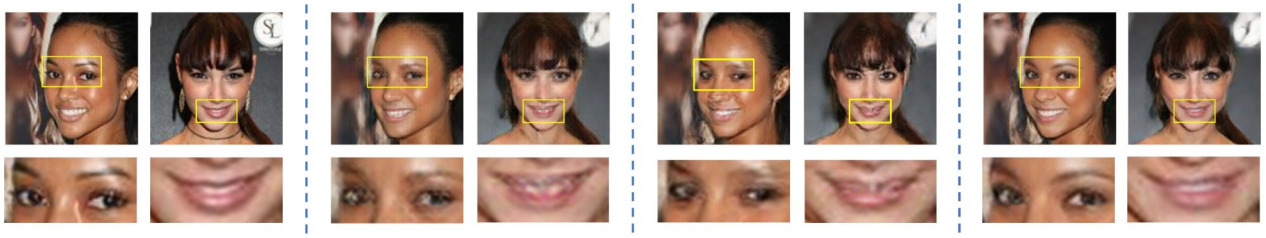
在保持生成器结构不变的条件下，通过对比多层特征反馈局部判别器与传统判别器[23]、局部判别器[27]的重建性能，验证了多层特征反馈特性在图像重建任务中的优势，对比结果如图 8 所示。从视觉效果来看，传统判别器重建的人脸图像存在模糊现象，局部判别器重建的细节信息仍显不足，而多层特征反馈局部判别器重建的图像在纹理和边缘细节上更为清晰。考虑到 PSNR 和 SSIM 难以捕捉人类感知上的质量差异，因此，本小节引入 FID 衡量生成图像与真实图像之间的相似度[31]，FID 越小，表明生成图像的质量越接近于真实图像。不同判别器重建图像性能对比如表 4 所示。

表3 生成器中各模块消融实验结果

Table 3 Ablation experiment results of each module in the generator

模型	PSNR/dB	SSIM
U-Net 模型	22.905 8	0.804 4
U-Net+ResBlocks 模型	23.298 0	0.815 6
U-Net+MSAF 模型	23.366 7	0.815 9
U-Net+ResBlocks+MSAF 模型	23.538 6	0.826 4

由表 4 数据分析可得：与传统判别器相比，使用多层特征反馈局部判别器重建的图像 PSNR 和 SSIM 相近，但 FID 提升了 32.163 3；相较于局部判别器，PSNR 提高了 0.130 8 dB，FID 提升了 14.361 0。原因在于传统判别器主要关注图像的全局特征，在训练过程中能够有效捕捉图像的整体信息，但难以精准重建局部细节，导致图像细节模糊。而局部判别器未能将多尺度语义信息反馈给生成器，导致重建图像的细节信息不足。相比之下，多层特征反馈局部判别器通过输出多个中间层特征，实现高质量重建。



(a) 原始图像 (b) 使用传统判别器重建的图像 (c) 使用局部判别器重建的图像 (d) 使用多层特征反馈局部判别器重建的图像

(a) Original images (b) The reconstructed images via the traditional discriminator (c) The reconstructed images via the local discriminator (d) The reconstructed images via the local discriminator with multi-layer feature feedback

图8 不同判别器重建图像的可视化结果

Figure 8 Visualization results of reconstructed images with different discriminators

3.4 对比分析

为了验证 BD-GAN 模型在图像重建任务上的优势，将其与 Pix2Pix[32]、STUNet[19]、SR3[20]和 DDCM[21]模型分别对 ITPE 生成的密文图像进行解密，并从以下两个方面进行对比：在重建性能方面，使用 SSIM 作为量化指标；在模型复杂度方面，利用参数量和计算量进行评估。考虑到扩散模型包含前向加噪与反向去噪的双向过程，其推理依赖于多步迭代采样。为评估网络本身复杂度，SR3 和 DDCM 仅统计反向去噪过程中的单步计算量。不同方案性能对比如表 5 所示。

由表 5 数据分析可得，Pix2Pix 的计算开销最小，但其重建图像质量损失较为明显。STUNet 由于堆叠大量

表4　不同判别器重建图像性能对比

Table 4　Comparison of reconstruction performance metrics for different discriminators

判别器	PSNR/dB	SSIM	FID
传统判别器	23.618 9	0.821 9	65.036 0
局部判别器	23.407 8	0.820 8	47.233 7
多层特征反馈局部判别器	23.538 6	0.826 4	32.872 7

的自注意力模块,计算量较大,且由于其无条件生成特性,对目标域分布要求严格,不利于图像的精细重建。SR3和DDCM均利用扩散概率模型实现了高质量重建,然而大量的计算与存储需求限制了实时应用及移动端部署。与扩散概率模型SR3和DDCM相比,BD-GAN模型在保证重建性能的前提下,参数量和计算量分别减少了5.120×10^{7}、6.484×10^{7}和2.807×10^{10}、5.041×10^{10}。

表5　不同方案性能对比

Table 5　Performance comparison of different schemes

模型	SSIM	参数量	计算量
Pix2Pix[32]	0.695	4.460×10^{7}	1.057×10^{10}
STUNet[19]	0.791	2.481×10^{7}	4.178×10^{10}
SR3[20]	0.822	9.781×10^{7}	4.593×10^{10}
DDCM[21]	0.826	1.115×10^{8}	6.827×10^{10}
BD-GAN	0.826	4.661×10^{7}	1.786×10^{10}

4　结论

本文提出了一种基于生成对抗网络的轻量化缩略图盲解密模型,能够在无需原始密钥的条件下,对TPE图像进行重建,解决了数据异常情况下解密失败的问题。通过在传统U-Net网络中加入残差块和多尺度注意力融合模块,增强对图像高频信息的重建能力,解决了多次采样和深层网络长距离依赖所导致的信息丢失问题;改进了具有多层特征反馈特性的局部判别器,提高对多尺度纹理细节的感知能力。实验结果表明,该模型适用于多种TPE算法的盲解密,对图像细节信息呈现出较好的重建效果。同时在不牺牲重建性能的前提下,具有更少的参数量和计算量,有效地降低了部署成本,为应急条件下的图像隐私保护与高保真重建提供了切实可行的解决方案。

参考文献

[1] Kagan D M, Alpert G F, Fire M. Zooming into video conferencing privacy[J]. IEEE Transactions on Computational Social Systems, 2024, 11(1): 933-944.

[2] Fainmesser I P, Galeotti A, Momot R. Digital privacy[J]. Management Science, 2023, 69(6): 3157-3173.

[3] Wen W Y, Yuan Z Y, Qi S R, et al. PPM-SEM: A privacy-preserving mechanism for sharing electronic patient records and medical images in telemedicine[J]. IEEE Transactions on Multimedia, 2024, 26: 5795-5806.

[4] 宋昭阳, 王一诺, 王浩文, 等. 基于Hopfield网络“伪吸引子”与交替量子随机行走的抗攻击彩色图像加密方案[J]. 电子学报, 2023, 51(8): 2030-2042.
Song Zhaoyang, Wang Yinuo, Wang Haowen, et al. Anti-attack color image encryption scheme based on Hopfield network “pseudo attractor” and alternating quantum random walk[J]. Acta Electronica Sinica, 2023, 51(8): 2030-2042. (in Chinese)

[5] Song J C, Chen B, Zhang J. Dynamic path-controllable deep unfolding network for compressive sensing[J]. IEEE Transactions on Image Processing, 2023, 32: 2202-2214.

[6] 陈云, 罗成, 许璐. 基于两种置乱方式的混合混沌图像加密算法[J]. 通信技术, 2024, 57(9): 925-933.
Chen Yun, Luo Cheng, Xu Lu. A hybrid chaotic image encryption algorithm based on two scrambling methods[J]. Communications Technology, 2024, 57(9): 925-933. (in Chinese)

[7] Jiang D H, Njitacke Z T, Long G Q, et al. Novel Tabu learning neuron model with variable activation gradient and its application to secure healthcare[J]. Chaos, Solitons & Fractals, 2024, 189: 115632.

[8] Wright C V, Feng W C, Liu F. Thumbnail-preserving encryption for JPEG[C]//Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security. New York: ACM, 2015: 141-146.

[9] Tajik K, Gunasekaran A, Dutta R, et al. Balancing image privacy and usability with thumbnail-preserving encryption[C]// Proceedings 2019 Network and Distributed System Security Symposium. Internet Society, 2019.

[10] Zhao R Y, Zhang Y S, Xiao X L, et al. TPE2: Three-pixel exact thumbnail-preserving image encryption[J]. Signal Processing, 2021, 183: 108019.

[11] Zhang Y S, Zhou W T, Zhao R Y, et al. F-TPE: Flexible thumbnail-preserving encryption based on multi-pixel sum-preserving encryption[J]. IEEE Transactions on Multimedia, 2023, 25: 5877-5891.

[12] Marohn B, Wright C V, Feng W C, et al. Approximate thumbnail preserving encryption[C]//Proceedings of the 2017 on Multimedia Privacy and Security. New York: ACM, 2017: 33-43.

[13] Zhang Y S, Zhao R Y, Xiao X L, et al. HF-TPE: High-fidelity thumbnail- preserving encryption[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022, 32(3): 947-961.

[14] Chai X L, Ma Y K, Wang Y J, et al. TPE-ADE: Thumbnail-preserving encryption based on adaptive deviation embedding for JPEG images[J]. IEEE Transactions on Multimedia, 2024, 26: 6102-6116.

[15] Wang X, Qu L F, Wu H T, et al. Reversible image thumbnail preservation with high-visual naturalness[J]. IEEE In-

ternet of Things Journal, 2024, 11(10): 18739-18752.
[16] 赵若宇, 叶茜, 周文韬, 等. 云存储图像缩略图保持的加密研究进展[J]. 中国图象图形学报, 2023, 28(3): 645-665. ZHAO Ruoyu, YE Xi, ZHOU Wentao, et al. Cloud-stored image thumbnail-preserving encryption[J]. Journal of Image and Graphics, 2023, 28(3): 645-665. (in Chinese)
[17] Li Z Y, Xie D, Liu S Q, et al. Known-plaintext attacks to thumbnail-preservation encryption using Pix2pix generative adversarial network[C]//ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE, 2025: 10890779.
[18] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 105-114.
[19] Zhang P Y, Zhang K H, Luo W H, et al. Blind face restoration: Benchmark datasets and a baseline model[J]. Neurocomputing, 2024, 574: 127271.
[20] Saharia C, Ho J, Chan W, et al. Image super-resolution via iterative refinement[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(4): 4713-4726.
[21] Jiang D H, Ni J Q, Alasbali N, et al. DDCM: Cracking anonymized facial images using denoising diffusion cryptanalytic model[J]. IEEE Transactions on Consumer Electronics, 2025, 71(2): 4464-4474.
[22] Bellare M, Ristenpart T, Rogaway P, et al. Format-preserving encryption[M]//Selected areas in cryptography. Berlin, Heidelberg: Springer, 2009: 295-312.
[23] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2. New York: ACM, 2014: 2672-2680.
[24] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[J]. Advances in Neural Information Processing Systems, 2017, 30: 5998-6008.
[25] Kim J, Nang J, Choe J. LMLT: Low-to-high multi-level vision transformer for lightweight image super- resolution[C]//2025 IEEE/CVF International Conference on Computer Vision Workshops. Piscataway: IEEE, 2025: 5568-5578.
[26] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[M]//Medical image computing and computer-assisted intervention - MICCAI 2015. Cham: Springer International Publishing, 2015: 234-241.
[27] Isola P, Zhu J Y, Zhou T H, et al. Image-to-image translation with conditional adversarial networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 5967-5976.
[28] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[M]//Computer vision - ECCV 2016. Cham: Springer International Publishing, 2016: 694-711.
[29] Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 815-823.
[30] Parkhi O M, Vedaldi A, Zisserman A. Deep face recognition[C]//Proceedings ofthe British Machine Vision Conference 2015. British Machine Vision Association, 2015: C.29.41.
[31] Heusel M, Ramsauer H, Unterthiner T, et al. GANs trained by a two time-scale update rule converge to a local Nash equilibrium[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 6629-6640.
[32] Liu X R, Meng X F, Wang Y R, et al. Known-plaintext cryptanalysis for a computational-ghost-imaging cryptosystem via the Pix2Pix generative adversarial network[J]. Optics Express, 2021, 29(26): 43860.

作者简介

朱礼亚 男,1980年7月出生于安徽省淮北市。现为长安大学电子与控制工程学院高级工程师、硕士生导师。主要研究方向为多媒体信息安全。
E-mail: lyzhu@chd.edu.cn

孙雅娜 女,2002年2月出生于陕西省商洛市。现为长安大学电子与控制工程学院硕士研究生。主要研究方向为多媒体信息安全。
E-mail: ynsun@chd.edu.cn

任 帅 男,1982年9月出生于山西省太原市。现为长安大学信息工程学院教授、博士生导师。主要研究方向为信息隐藏与数字水印技术、数字取证与篡改技术、信息隐藏分析技术。
E-mail: shuairen@chd.edu.cn