

基于混沌映射和可逆神经网络的图像隐写算法研究^①

梁梦华^② 赵鸿图^③

(河南理工大学物理与电子信息学院 焦作 454001)

摘要 针对大容量隐写规模导致的图像质量下降问题,本文设计了一种全新的可逆大容量图像隐写网络,该网络结合了混沌映射与可逆神经网络,在保证大容量隐写的同时极大提高了隐写的安全性。首先,设计了一个具有线平衡的四维混沌系统,通过混沌映射加密算法对秘密图像的内容进行置乱,从而解决了秘密图像信息在传输过程中的泄露问题。其次,对载体图像进行预处理,在 Y 通道隐藏秘密图像,显著增强了隐写容量。在 DIV2K 和 COCO 数据集上进行了大量实验,载体图像与载密图像的峰值信噪比(peak signal-to-noise ratio, PSNR)高达 41.63 dB,秘密图像与恢复图像的 PSNR 为 43.29 dB。大量的实验结果表明,在大容量隐写条件下,本文方法不仅能保持优异的隐写性能,同时能高保真度显示隐写图像,其抗隐写能力远远优于现有的先进算法。

关键词 图像隐写;混沌映射;可逆神经网络;图像质量;抗隐写分析;鲁棒性

隐写术起源于 20 世纪 40 年代,最初的研究主要集中在对声音和文字信息的隐藏。随着互联网的普及和大数据时代的到来,数字图像作为信息的主要载体之一日益凸显。图像隐写作为一种有效的信息隐藏手段,在军事、商业、医疗等领域有着广泛的应用^[1]。

早期隐写术的安全要求主要是不可感知性,即携带秘密信息的图像在视觉上与正常图像无法区分。传统的信息隐藏技术主要分为空间域信息隐藏方法^[2-3]和变换域信息隐藏方法^[4-5]。最低有效位(least significant bits, LSB)^[6]是最早的图像隐写方法之一,也是空间域隐写方法最为经典的一种,通过修改图像像素值的最低有效位来隐藏信息。该方法简单易行,但鲁棒性较差。此外,现有空间域隐写方法还包括插值展开法^[7]、直方图移位法^[8]、差分展开法^[9]和自适应方法^[10]。空间域隐写方法在鲁棒性和安全性方面较弱。变换域隐写方法主要有离散余弦变换(discrete cosine transform, DCT)方法^[11],

DCT 首先将图像从空间域转换到频域,然后在频域隐藏信息。该方法具有较高的隐藏能力和较好的鲁棒性,但计算量大。此外还包括离散小波变换(discrete wavelet transform, DWT)^[12]等常见方法。变换域隐写嵌入容量小,图像视觉质量不佳。

随着大数据的发展,深度学习已成为一种趋势,并被广泛应用于信息隐藏。基于卷积神经网络(convolutional neural network, CNN)的图像隐写方法可分为编解码模型^[13-14]、无载波映射模型^[15-16]和边缘检测隐写模型^[17-18]。但以上图像隐写方法在提取过程中容易出现图像失真,并且需要耗费大量计算资源来训练模型。生成对抗网络(generate adversarial networks, GAN)可以生成逼真的图像来实现信息隐藏。基于 GAN 的图像隐写技术大致可分为非编解码网络的图像隐写^[19-20]和编解码网络的图像隐写^[21-22]。这些方法在信息提取的准确性、图像质量、隐写安全性和实用性等方面存在不足。

可逆神经网络(invertible neural network, INN)

① 河南省科技厅科技攻关和软科学项目(192102310446)和河南省高校基本科研业务费专项资金(NSFRF210406)资助项目。

② 女,1997年生,硕士生,研究方向:图像处理和智能信号处理;E-mail:liang_mmh@163.com。

③ 通信作者,E-mail:hongtuzhao@hpu.edu.cn。

(收稿日期:2024-09-15)

是一种具有双目标结构和高效可逆性的神经网络, 以其优异的性能受到了众多研究者的关注。Jing 等人^[23]将图像隐藏和图像恢复视为可逆神经网络的正向和反向过程, 设计了一种单色图像隐藏方法 HiNet。卞玉星等人^[24]将嵌入和提取过程分别与可逆神经网络的正向和逆向映射相关联, 提出了一种基于 INN 的多载体图像隐写模型。李泓萱等人^[25]设计了一种大容量、裁剪稳健的多级双向映射的可逆隐写网络, 在保持隐写不可感知性和大容量的同时, 有效提高了隐写的鲁棒性。与分别使用 2 个网络进行嵌入和提取的方法相比, 基于可逆神经网络的隐写方法提高了视觉效果和提取精度。

针对大容量隐写规模导致的图像质量下降问题, 本文提出一种基于混沌映射和可逆神经网络的图像隐写网络, 以实现隐写网络的高容量与鲁棒性。本文工作的主要贡献可以概括如下。

(1) 设计了一种新的可逆大容量图像隐写网络, 该网络结合了可逆神经网络与混沌映射, 在保证大容量隐写的同时大大提高了隐写的安全性。

(2) 设计了一个具有线平衡的四维混沌系统, 通过混沌映射加密算法对秘密图像的内容进行置乱, 解决了秘密图像信息在传输过程中的泄露问题。

(3) 对载体图像进行处理, 在 Y 通道隐藏彩色图像, 显著增强了隐写容量。

1 相关理论

1.1 可逆神经网络

自文献^[26]提出 INN 的概念以来, INN 以其优异的性能引起了众多研究者的关注。文献^[27]利用 INN 构建了潜在变量 z 与自然图像 x 之间的可逆映射。INN 通过一系列可逆变换将高维复杂分布 p_x 映射到简单潜在分布 p_z , 并利用神经网络学习 p_x 与 p_z 之间的映射关系。框架的前向过程以高维复杂数据 x 作为输入, 输出符合简单分布的数据 z 。逆过程是一个生成建模过程, 以采样数据 z 作为输入生成高维复杂数据 x 。INN 的基本网络架构是由 real NVP(real-valued non-volume preserving) 中的放射耦合层推广而来, 其工作原理是将输入数据分成 μ_1 和

μ_2 , 然后通过学习函数 s_i 和 t_i 进行转换并以交替的方式耦合, 输出耦合后的数据 v_1 和 v_2 。INN 正向过程如图 1 所示, 计算公式如式(1)和式(2)所示。

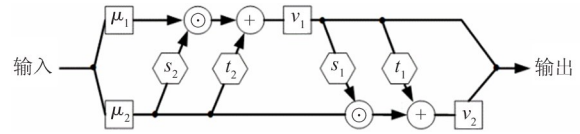


图 1 INN 的前向过程

$$v_1 = \mu_1 \odot \exp(s_1 \mu_2) + t_2(\mu_2) \quad (1)$$

$$v_2 = \mu_2 \odot \exp(s_1 v_2) + t_1(\mu_1) \quad (2)$$

式中: \exp 是指数函数; \odot 代表阿达玛乘积, 即将 2 个矩阵的对应元素相乘, 得到的矩阵大小与原矩阵相同。INN 的逆过程如图 2 所示, 计算如式(3)和式(4)所示。

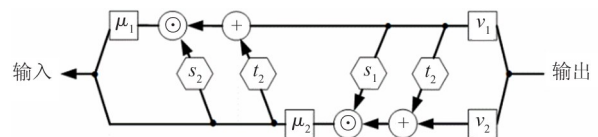


图 2 INN 的逆向过程

$$\mu_1 = (v_1 - t_2(\mu_2)) \odot \exp(-s_2(\mu_2)) \quad (3)$$

$$\mu_2 = (v_2 - t_1(v_1)) \odot \exp(-s_1(v_1)) \quad (4)$$

由于其强大的网络表示能力, INN 适用于图像隐藏、图像着色、图像缩放、图像压缩等各种图像任务。在本文中, 利用 INN 的可逆前向和后向过程分别对多个图像进行隐藏和恢复。

1.2 混沌映射系统

混沌系统是一种确定性系统, 其特征是看似随机、不规则的运动, 具有不确定性、不可约性和不可预测性。为了在混沌系统的复杂性和效率之间取得平衡, 气象学家 Lorenz 提出洛伦兹混沌系统(Lorenz chaotic system)^[28], 该系统由 3 个带表达式的耦合微分方程组成, 如式(5)所示。

$$\begin{cases} dx = \alpha(y - x) \\ dy = \beta x - y - xz \\ dz = xy - \sigma z \end{cases} \quad (5)$$

式中: x, y, z 为系统的状态变量, α, β, σ 为系统的参数。当参数设为 $\alpha = 10, \beta = 28, \sigma = 8/3$, 状态变量为 $(1, 1, 1)$ 时, 系统表现出经典混沌现象, 对初始条

件极其敏感,且具有随机性。

自 20 世纪 90 年代 Fridrich 首次将混沌系统与图像加密相结合以来,混沌系统因其遍历性、伪随机性、初值敏感性等特点在密码学中得到了广泛的应用。Jahanshahi 等人^[29]设计了一个无平衡的混沌系统,通过设计和实现模拟电路来研究系统的混沌行为。Khandelwal 等人^[30]通过采用离散小波变换和奇异值分解将秘密信息嵌入到封面图像中。刘帅等人^[31]将超 Lorenz 混沌系统置乱与量子随机图像结合,设计了基于 NEQR (novel enhanced quantum representation) 模型的 3 层量子图像加密算法。这些方法在提取的图像中存在一定的损失,难以保证秘密图像的质量。

为解决上述问题,本文设计了具有线平衡的四

维混沌系统,并将混沌映射与可逆神经网络结合,在保证大容量隐写的同时极大提升了隐写的安全性。

2 图像隐写方案

2.1 模型整体框架

本文在秘密图像前采用混沌加密技术,即通过混沌加密算法对秘密图像的内容进行置乱,然后将加密后的秘密图像与载体图像合并,从而解决了秘密图像信息在传输过程中的泄露问题。当接收者接收到加密图像时,首先通过提取网络对加密后的秘密图像进行提取,然后对加密后的图像进行恢复,得到秘密图像。采用双重加解密的方法,提高了隐写方案的安全性。本文所提模型框架分为前向隐藏过程和逆向揭示过程 2 部分,如图 3 所示。

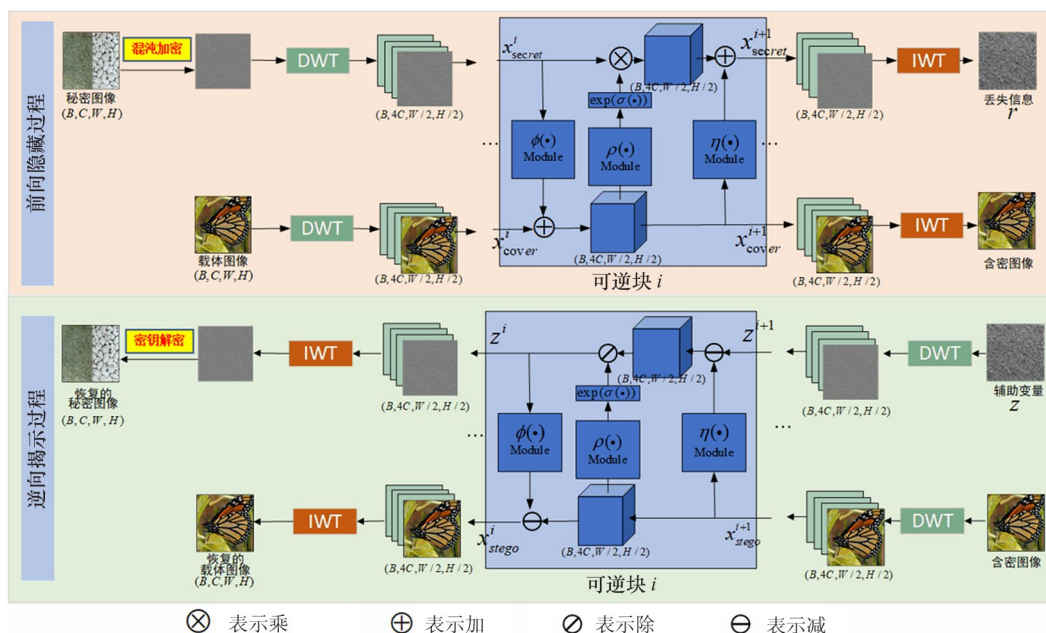


图 3 网络框架

在前向隐藏过程中,首先,载体图像由红绿蓝 (red, green, blue, RGB) 编码格式转换为亮度-蓝色色差-红色色差 (luminance, blue-difference chroma, red-difference chroma, YCrCb) 格式,选择 Y 通道作为实际的载体图像。随后, Y 通道载体图像和加密后的秘密图像使用 Haar 小波进行单独变换。Haar 小波是一种易于实现的离散小波变换变体。将信息隐藏在频域比隐藏在空间域更有利于保持图像的视觉质量。将变换后的 2 幅图像输入到网络

中。最终进行逆小波变换 (inverse wavelet transform, IWT),生成隐写图像和丢失信息。

在逆向揭示过程中,需要引入辅助随机变量 z , 变量 z 是从任意高斯分布中随机抽取的,与丢失信息 r 的分布相同。辅助变量和含密图像通过可逆块重构出载体图像和加密后的秘密图像。后者的逆置乱得到最终的秘密图像。在隐藏和重构块中, $\varphi(\cdot)$ 、 $\rho(\cdot)$ 和 $\eta(\cdot)$ 可以是任意函数。本文对 $\varphi(\cdot)$ 、 $\rho(\cdot)$ 和 $\eta(\cdot)$ 使用 DenseNet。

2.2 图像预处理

与将信息隐藏在 RGB 或灰色图像中的传统图像隐写方法相反,本文方法涉及将秘密信息隐藏在 YCrCb 图像的 Y 通道中。一些基于 RGB 的隐写技术往往会扭曲隐写图像的颜色,从而导致色彩准确性问题。该过程由式(6)和式(7)描述。

$$\begin{cases} Y = 0.299R + 0.587G + 0.114B \\ Cr = 0.713(R - Y) \\ Cb = 0.564(B - Y) \end{cases} \quad (6)$$

$$\begin{cases} R = Y + 1.402Cr \\ G = Y - 0.344Cb - 0.714Cr \\ B = Y + 1.772Cb \end{cases} \quad (7)$$

在 Y 通道隐藏彩色图像,显著增强了隐写容量,同时能够有效减少由于信息嵌入带来的视觉伪影。

2.3 混沌加密算法

本文提出的混沌算法是通过增加一个线性项从 Sprott b 型混沌流中推导出来的。如式(8)所示的常微分方程是 Sprott b 型混沌系统的原始模型。切换状态变量 x, y , 引入参数 α, β, σ , 常微分方程变为式(9)。加入一个参数为 γ 的线性项作为状态变量 w , 使所构造的混沌系统具有线均衡。式(10)中的最终方程是所提出的混沌系统。

$$\begin{cases} dx = yz \\ dy = x - y \\ dz = 1 - xy \end{cases} \quad (8)$$

$$\begin{cases} dx = \alpha(y - x) \\ dy = \beta xz \\ dz = \sigma - xy \end{cases} \quad (9)$$

$$\begin{cases} dx = \alpha(y - x) \\ dy = \beta xz \\ dz = \sigma w - xy \\ dw = \gamma y \end{cases} \quad (10)$$

通过求解 $dx = 0, dy = 0, dz = 0$ 和 $dw = 0$, 得到平衡点。

$$\begin{cases} \alpha(y - x) = 0 \\ \beta xz = 0 \\ \sigma w - xy = 0 \\ \gamma y = 0 \end{cases} \quad (11)$$

显然,对于非零参数, $(0, 0, z, 0)$ 为系统的平衡点。因此,该系统是一个具有线平衡的四维混沌系统(four-dimensional chaotic system with line equilibriums, 4DLECS)。从计算的角度来看,这些吸引子是隐藏的。将参数设为 $\alpha = 2, \beta = 6, \sigma = 2, \gamma = 0.1$, 初始值为 $[1, 1, 1, 1]$, 可得状态图,如图 4 所示。

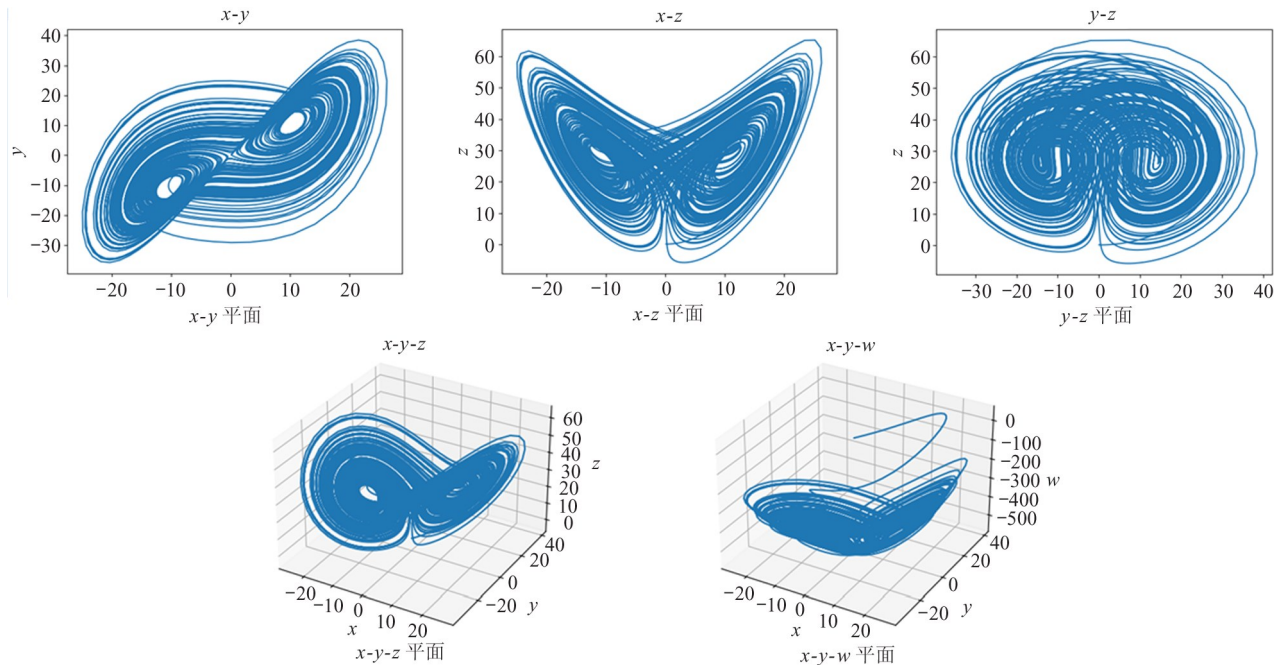


图 4 4DLECS 的隐藏吸引子

李雅普诺夫(Lyapunov)指数是衡量系统动态特性的重要定量指标。通过最大李雅普诺夫指数(maximum lyapunov index, MLE)是否大于0,可以直观地判断系统中是否存在动态混沌。李雅普诺夫指数如图5所示。计算结果为 $\lambda_1 = 0.341$, $\lambda_2 = 0.003$, $\lambda_3 = -0.017$, $\lambda_4 = -2.327$ 。由于MLE大于0,系统处于混沌状态。

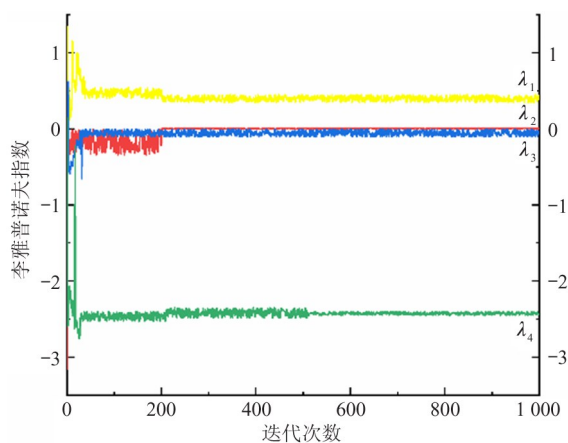


图5 4DLECS的混沌系统的Lyapunov指数

混沌加密算法如算法1所示。

算法1: 秘密图像加密算法

输入:秘密图像 x_{secret} , 混沌序列 X, Y, Z

输出:加密数据

function 加密函数 (x_{secret}, X, Y, Z)

$[height, width] = size(x_{secret});$

 for $i = 1 : height$

 对 x_{secret} 的第 i 列进行移位的步数 $X(i)$ 的循环移位;

 将移位后的结果存储到 $shuffle_row?$ 的第 i 行中;

 end

 for $i = 1 : width$

 对 x_{secret} 的第 i 行进行移位的步数 $Y(i)$ 的循环移位;

 将移位后的结果存储到新矩阵 $shuffle_col$ 的第 i 行中;

 end

 对 $shuffle_col?$ 和 Z 进行按位异或操作;

end function

其中, $shuffle_row$ 和 $shuffle_col$ 分别表示对数据的行和列进行随机洗牌,数据被重新存储。

本文在三维混沌映射的基础上,提出四维映射系统,相较于三维系统具有更加复杂和丰富的动力学行为,这使其在图像隐写方面具有更高的安全性和复杂性。同时,通过利用混沌系统的初值敏感性和长期不可预测性,可以生成难以破译的加密密钥和加密算法,从而增强图像隐写的安全性。此外,利用混沌系统的并行性和快速迭代特性,可以快速地生成加密图像和恢复原始图像信息,降低时间复杂度,减少训练时间。

2.4 损失函数

本文的损失函数由3部分组成,分别是隐藏损失、重构损失以及丢失信息损失。

总损失函数 L 表示为式(12)。

$$L = \lambda_1 L_{\text{隐藏}} + \lambda_2 L_{\text{重构}} + \lambda_3 L_{\text{丢失}} \quad (12)$$

(1) 隐藏损失 $L_{\text{隐藏}}$: 在前向隐藏过程中,网络将秘密信息隐藏在掩蔽图像中,生成隐写图像。目的是使隐写图像在视觉上接近封面图像。因此,隐藏损失定义如下:

$$L_{\text{隐藏}} = \sum_{n=1}^N l_{\text{隐藏}}(X_{\text{载体}}^{(n)}, X_{\text{含密}}^{(n)}) \quad (13)$$

式中: N 是训练样本的数量,本文为1000, $l_{\text{隐藏}}$ 为载体图像与含密图像之间的差值。

(2) 重构损失 $L_{\text{重构}}$: 通过反向重构过程重建的秘密图像需要与原始秘密图像保持一致,为此,重构损失定义为:

$$L_{\text{重构}} = \sum_{n=1}^N l_{\text{重构}}(X_{\text{秘密}}^{(n)}, X_{\text{恢复秘密}}^{(n)}) \quad (14)$$

式中: $l_{\text{重构}}$ 为秘密图像与恢复的秘密图像之间的差值。

(3) 丢失信息或丢失 $L_{\text{丢失}}$: 将损失信息 r 作为分布损失函数进行 l_2 正则化,约束损失信息分布更集中在接近零的值周围,从而降低了模型的复杂度,使模型更加平滑,训练更加稳定。

$$L_{\text{丢失}} = \|r\|_2^2 \quad (15)$$

式中: $\|r\|_2^2$ 表示 r 的平方和。

3 实验设置与结果分析

3.1 实验设置

实验环境:本实验在4块型号为 GeForce RTX 3090 的 GPU 和 24 GB RAM 的深度学习服务器上进行

训练,深度学习框架与版本号为 Pytorch1.7.1,Python解释器版本 3.7.16 和 CUDA 版本 11.0。

数据集:DIV2K 数据集是一个多样化、高分辨率、逼真的数据集,涵盖了自然景观、人物、动物、建筑等领域,非常适合信息隐藏的要求。本文使用 DIV2K 数据集进行实验。为了验证本文方法的泛化能力,还使用了 COCO 数据集。

实验设置:网络模型使用 Adam 优化器进行训练,学习率设置为 $1 \times 10^{-4.5}$,批量大小(batch size)设置为 32。整个网络模型的可逆块数量为 16 个,每个可逆块包含的卷积块数量为 7。 $\lambda_1 = 8, \lambda_2 = 1, \lambda_3 = 1$,隐藏损失采用 L2 正则化损失函数。

评价指标:使用峰值信噪比(peak signal-to-noise ratio,PSNR)和结构相似性指数(structural similarity,SSIM)值来衡量图像的质量,以相对有效载荷(relative payload,RP)作为评价指标来比较每个隐写图像的隐写能力。

RP 的数学表达式如式(16)所示。

$$RP = \frac{bits(\text{秘密信息})}{bits(\text{覆盖容量})} \times 100\% \quad (16)$$

此外,在第 3.2.3 节中通过实验进行了安全性分析,在 3.2.4 中进行了直方图分析,在 3.2.5 进行了鲁棒性分析。

3.2 实验与分析

3.2.1 对比实验

为了证明本文方法的优越性,本文方法与 5 种先进算法进行了对比:LSB-1bit 是一种传统的隐写方法;HiNet 是一种基于可逆神经网络的经典隐写算法;StegaLIN 是基于可逆网络的轻量化图像隐写方法;PRIS(practical robust invertible network for image steganography)是基于 HiNet 的实用的鲁棒可逆网络;DeepMIH 为基于可逆神经网络的多图像隐藏方法。对比结果如表 1 所示。由表 1 可以看出,本文方法载体和载密图像对和秘密和恢复的图像对的 PSNR 值都超过 40.00 dB,表明视觉质量非常好。与其他方法相比,本文方法获得了最好的精度值。与次优结果相比,在 DIV2K 数据集上,对载体和载密图像来说,PSNR 和 SSIM 分别提高了 0.24 dB 和 0.01;对秘密和恢复图像来说,PSNR 和 SSIM 分别提高了 2.52 dB 和 0.04。在 COCO 数据集上,对载体和载密图像来说,PSNR 和 SSIM 分别提高了 3.37 dB 和 0.02;对秘密和恢复图像来说,PSNR 提高了 2.29 dB,SSIM 高达 0.99。

此外,将采用深度学习方法进行图像隐写的处理时间进行评估,如表 2 所示,对比其他算法,本文方法复杂度最小,运行时间最少。

表 1 对比实验结果

方法	DIV2K 数据集						COCO 数据集					
	RP/%	训练时间	载体和载密图像 PSNR/dB SSIM	秘密和恢复图像 PSNR/dB SSIM	RP/%	训练时间	载体和载密图像 PSNR/dB SSIM	秘密和恢复图像 PSNR/dB SSIM				
LSB-1bit	50	-	33.19 0.95	30.82 0.90	50	-	33.25 0.94	30.77 0.91				
HiNet	300	15 h 23 min 46 s	34.32 0.87	40.30 0.98	300	15 h 58 min 19 s	33.95 0.88	40.45 0.99				
DeepMIH	300	17 h 51 min 17 s	35.43 0.95	40.77 0.95	300	18 h 47 min 51 s	35.00 0.95	40.05 0.99				
PRIS	300	12 h 19 min 47 s	41.39 0.98	40.71 0.99	300	13 h 28 min 16 s	37.98 0.96	39.16 0.98				
StegaLIN	300	12 h 53 min 16 s	34.24 0.94	32.70 0.91	300	14 h 01 min 48 s	34.91 0.94	32.89 0.93				
本文	300	11 h 52 min 10 s	41.63 0.99	43.29 0.99	300	12 h 32 min 11 s	41.35 0.98	42.74 0.99				

通过从载密图像中减去载体图像的像素得到残差图,如图 6 所示。

从视觉上看,LSB-1bit、HiNet 中载密放大图像显示出明显的纹理复制伪影,暴露了秘密信息的轮

廓,存在信息泄露的风险。DeepMIH 的隐写图像颜色失真较小,但放大后仍然肉眼可见。本文方法经过图像置乱后,与相应的封面图像非常相似,两者之间的残差图几乎是纯黑色的。这表明载体图像与载密图像之间的差异极小。本文方法在大容量和复杂的隐藏信息下也能显示出高保真的结果。

表 2 消融实验结果

图像 预处理	混沌 映射	载体和载密图像		秘密和恢复图像	
		PSNR/dB	SSIM	PSNR/dB	SSIM
×	×	34.32	0.87	40.30	0.98
×	√	39.95	0.96	42.26	0.97
√	×	35.95	0.92	41.25	0.97
√	√	41.63	0.99	43.29	0.99

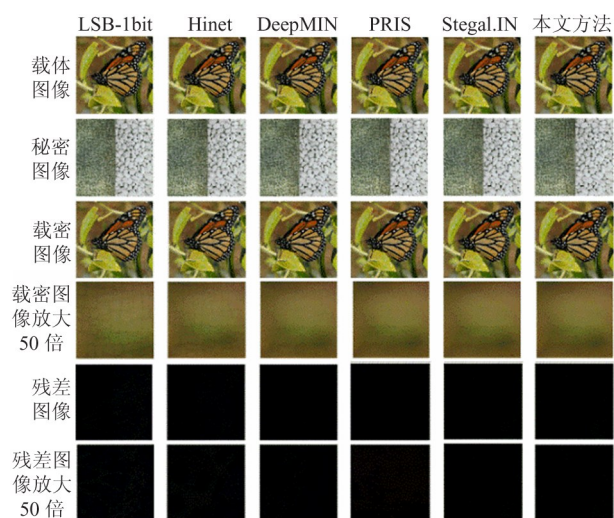


图 6 不同方法下载体图像与载密图像残差图的比较

3.2.2 消融实验

图像消融算法对于本文方法有 2 个至关重要的目的:(1)即使训练图像集泄露,也能防止通过相应的错误图像泄露机密信息;(2)确保攻击者在没有密钥和逆排列算法知识的情况下,即使有黑盒攻击尝试,也无法准确地重建正确的信息。

表 2 说明了混沌映射以及预处理的有效性。其中第一行为基准实验。从表 2 第 2 行数据可以看出,通过混沌映射生成的混沌序列,可以在保证图像质量的前提下,提高秘密信息的嵌入容量。这意味着可以在不增加载体图像失真的情况下,嵌入更多的秘密信息,且混沌映射在图像隐写中的应用还可

以通过对嵌入过程的精细控制来优化图像质量,能最大限度地减少载体图像的失真和降质。从表 2 第 3 行可以看出,通过预处理将秘密信息隐藏在 YCrCb 图像的 Y 通道中,能够有效减少由于信息嵌入带来的视觉伪影,从而提高图片质量。从表 2 第 4 行可以看出,混沌映射和图像预处理共同作用,能大幅提升图像质量。在秘密图像恢复过程中,也能保持提取图像良好的视觉效果。

尽管在封面图像中隐藏具有复杂纹理的秘密图像具有挑战性,但本文方法没有显示出与原始秘密图像相关的纹理复制伪影。如图 7 所示,本文方法不仅实现了相对较大的隐藏能力,而且保持了较好的视觉效果和安全性。

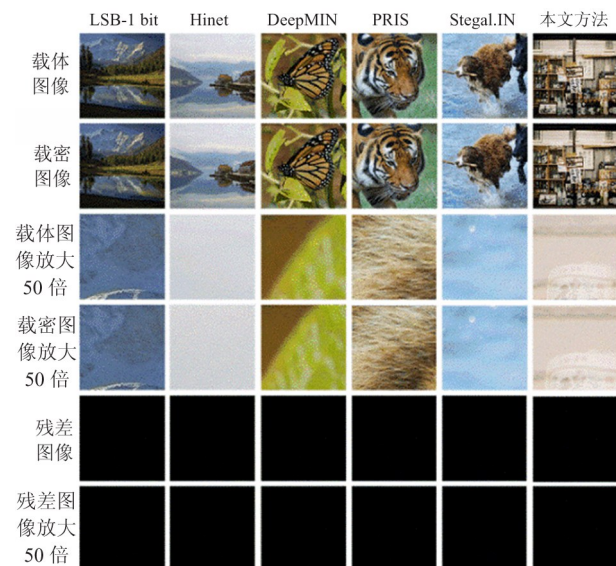


图 7 覆盖图像与隐去图像的残差图

3.2.3 安全分析

为了验证安全性,本文采用了 3 种隐写分析网络 XuNet^[32]、SRNet^[33] 和 WISERNet^[34] 来评估嵌入有效载荷为 1 bpp(bits per pixel)的不同隐写方法的隐写安全性。本实验为每种方法生成了 1 000 个载体/载密图像对,以重新训练隐写分析网络。表 3 列出了使用 3 种不同方法的隐写分析网络的检测误差。本文方法在 3 个隐写分析网络中的检测误差大大高于其他方法,这表明本文方法具有更好的隐写安全性。

3.2.4 直方图分析

直方图分析是一种指定图像像素值分布的方

表 3 3 种隐写分析网络检测误差比较 %

方法	检测误差		
	XuNet	SRNet	WISERNet
LSB-1bit	0.26	0.00	2.41
HiNet	2.34	0.11	1.27
DeepMIH	4.31	1.74	2.58
PRIS	10.15	2.67	4.68
StegaLIN	9.36	3.71	4.29
本文	10.57	4.25	8.51

法。不同的图像本身表现出不同的分布趋势,图像像素值的任何改变都会改变其直方图分布。因此,评估图像的直方图分布是否与其原始对应的图像保持相似或相同已成为识别图像是否隐藏任何秘密信息的经典方法。训练结束后,随机选择了 8 幅图像,比较它们的直方图。图 8 显示了本实验中载体图像和载密图像的直方图,揭示了视觉一致性,强调了本文方法的高安全性和不可检测性。

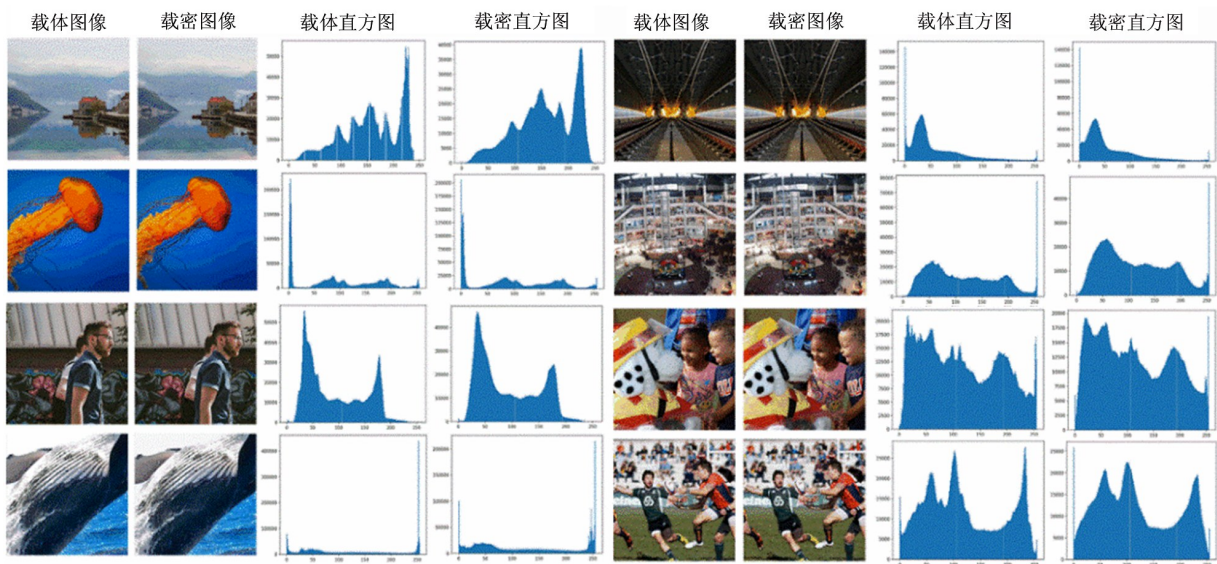


图 8 直方图分析结果对比

3.2.5 鲁棒性分析

在鲁棒性分析中,本文引入了多种攻击,包括 JPEG (Joint Photographic Experts Group) 压缩攻击、

高斯噪声、椒盐噪声。在 DIV2K 数据集的 1 000 张测试集上进行测试,得到的实验结果如表 4 所示。面对这些干扰,本文提出的隐写算法显示出优异的

表 4 鲁棒性分析实验结果

方法		无攻击		JPEG 攻击		高斯噪声		椒盐噪声	
		载体和载密图像	秘密和恢复图像	载体和载密图像	秘密和恢复图像	载体和载密图像	秘密和恢复图像	载体和载密图像	秘密和恢复图像
LSB-1bit	PSNR/dB	33.19	30.82	26.11	26.33	28.20	29.17	20.77	22.35
	SSIM	0.95	0.90	0.79	0.81	0.81	0.83	0.70	0.73
HiNet	PSNR/dB	34.32	40.30	29.36	29.80	30.81	0.92	25.19	24.70
	SSIM	0.87	0.98	0.88	0.88	0.90	31.16	0.74	0.73
DeepMIH	PSNR/dB	35.43	40.77	29.35	31.10	31.84	31.99	28.52	28.14
	SSIM	0.95	0.95	0.88	0.91	0.91	0.91	0.84	0.83
PRIS	PSNR/dB	41.39	40.71	30.17	32.73	33.53	33.40	29.46	30.15
	SSIM	0.98	0.99	0.90	0.92	0.92	0.93	0.88	0.90
StegaLIN	PSNR/dB	34.24	32.70	30.38	29.96	31.66	29.83	30.19	30.97
	SSIM	0.94	0.91	0.91	0.89	0.91	0.88	0.90	0.90
本文	PSNR/dB	41.63	43.29	30.97	0.94	0.94	0.94	0.90	0.92
	SSIM	0.99	0.99	0.91	33.04	34.16	34.91	30.14	31.33

稳健性,这进一步证明了该算法在复杂环境下的信息嵌入能力。如图4所示,在椒盐噪声攻击下,本文方法与StegaLIN方法的PSNR值仅差0.05 dB,其他情况下本文均为最优。这是由于本文所提混沌映射的敏感性和复杂性使得生成的混沌序列具有强大的抗攻击能力。即使载体图像受到一定程度的损坏或攻击(如JPEG压缩、高斯噪声、椒盐噪声等),通过混沌映射嵌入的秘密信息仍然能够保持较高的完整性和可提取性。

4 结论

本文提出了一种基于混沌映射和可逆神经网络的鲁棒性图像隐写方法,利用深度学习实现图像信息隐藏。该方法显著提高了隐写图像的质量、隐写网络的鲁棒性和安全性。首先,本文设计了一个具有线平衡的四维混沌系统。网络率先对秘密图像采用混沌映射加密技术,通过混沌映射加密算法对秘密图像的内容进行置乱,将加密后的秘密图像与载体图像合并,从而解决了秘密图像信息在传输过程中的泄露问题。此外,本文对载体图像进行预处理,在Y通道隐藏秘密图像,显著增强了隐写容量。在DIV2K和COCO数据集上进行了实验,载体图像与载密图像的PSNR高达41.63 dB,秘密图像与恢复图像的PSNR为43.29 dB。实验结果表明,在大容量隐写尺度下,本文方法能保持良好的隐写不可见性,同时能高保真度显示隐写图像,其抗隐写能力远远优于已有的隐写方法。

参考文献

- [1] 张笑. 图像加密技术综述[J]. 网络安全技术与应用, 2021(7):35-36.
- [2] Tamimi A A, Abdalla A M, Al-Allaf O. Hiding an image inside another image using variable-rate steganography[J]. International Journal of Advanced Computer Science and Applications, 2013,4:18-21.
- [3] Asad M, Gilani J, Khalid A. An enhanced least significant bit modification technique for audio steganography[C]// International Conference on Computer Networks and Information Technology. Abbottabad, Pakistan: IEEE, 2011: 143-147.
- [4] Kumar V, Kumar D. A modified DWT-based image steganography technique[J]. Multimedia Tools and Applications, 2018,77:13279-13308.
- [5] Hassaballah M, Hameed M A, Awad A I, et al. A novel image steganography method for industrial Internet of things security[J]. IEEE Transactions on Industrial Informatics, 2021,17(11):7743-7751.
- [6] Ehsan A U A, Ali E, Sohrawordi M, et al. A LSB based image steganography using random pixel and bit selection for high payload[J]. Mathematical Sciences and Computing, 2021,3:24-31.
- [7] Benhfid A, Ameer E B. Image data hiding scheme based on spline interpolation and OPAP[J]. International Journal of Knowledge Engineering and Soft Data Paradigms, 2019,6(2):139-150.
- [8] Huang L C, Tseng L Y, Hwang M S. A reversible data hiding method by histogram shifting in high quality medical images[J]. Journal of Systems and Software, 2013, 86(3):716-727.
- [9] Ge Y, Zhang M, Yang P. Reversible data hiding in encrypted domain based on color image channel correlation[J]. Proceedings of SPIE, 2023,12587:317-324.
- [10] Yu C, Zhang X, Li G, et al. Reversible data hiding with adaptive difference recovery for encrypted images[J]. Information Sciences, 2022,584:89-110.
- [11] 毛炳华,王子驰,张新鹏. 基于DCT域相关性的非对称JPEG隐写[J]. 计算机科学,2019,46(1):196-200.
- [12] Liu Q, Xiang X Y, Qin J H, et al. Coverless steganography based on image retrieval of DenseNet features and DWT sequence mapping[J]. Knowledge-Based Systems, 2020,192:105375.
- [13] Kich I, Taouil Y, Benhfid A. Image steganography scheme using dilated convolutional network[C]//2021 12th International Conference on Information and Communication System. Valencia, Spain: IEEE, 2021:305-309.
- [14] Liu L, Meng L, Peng Y, et al. A data hiding scheme based on U-Net and wavelet transform[J]. Knowledge-Based Systems, 2021,223:107022.
- [15] Ma S, Zhao X. Generating JPEG steganographic adversarial example via segmented adversarial embedding[C]// International Workshop on Digital Watermarking. Cham, Switzerland: Springer International Publishing, 2020: 68-79.
- [16] Liu Q, Xiang X Y, Qin J H, et al. A robust coverless steganography scheme using camouflage image[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2022,32(6):4038-4051.
- [17] Varma D, Mishra S, Meenpal A. An adaptive image steganographic scheme using convolutional neural network and dual-tree complex wavelet transform[C]//2020 11th International Conference on Computing, Communication and Networking Technologies. Kharagpur, India: IEEE, 2020:

- 1–7.
- [18] Ray B, Mukhopadhyay S, Hossain S, et al. Image steganography using deep learning based edge detection[J]. *Multimedia Tools and Applications*, 2021, 80(24):33475–33503.
- [19] Liu Q, Yang J, Jiang H, et al. When deep learning meets steganography: protecting inference privacy in the dark[C] // *IEEE INFOCOM 2022-IEEE Conference on Computer Communications*. London, UK: IEEE, 2022: 590–599.
- [20] Zhao J, Wang S. A stable GAN for image steganography with multi-order feature fusion[J]. *Neural Computing and Applications*, 2022, 34(18):16073–16088.
- [21] Yu C. Attention based data hiding with generative adversarial networks[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2020, 34(1):1120–1128.
- [22] Tancik M, Mildenhall B, Ng R. StegaStamp: invisible hyperlinks in physical photographs[C] // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE, 2020:2117–2126.
- [23] Jun P J, Deng X, Xu M, et al. HiNet: deep image hiding by invertible network[C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Montreal, Canada: IEEE, 2021:4733–4742.
- [24] 卞玉星, 黄荣, 周树波, 等. 基于可逆神经网络的多载体图像隐写模型[J]. *计算机工程*, 2024, 50(12): 213–223.
- [25] 李泓萱, 张松洋, 任博. 基于多级可逆神经网络的大容量裁剪稳健型图像隐写技术[J]. *图学学报*, 2023, 44(6):1149–1161.
- [26] Dinh L, Krueger D, Bengio Y. Nice: Non-linear independent components estimation[EB/OL]. (2014–10–30)[2024–09–15]. <https://arxiv.org/pdf/1410.8516>.
- [27] Kinga D, Adam J B. Adam: a method for stochastic optimization[EB/OL] // (2014–12–22)[2024–09–15]. <https://arxiv.org/pdf/1412.6980>.
- [28] Ye G, Wu H, Liu M, et al. Image encryption scheme based on blind signature and an improved Lorenz system[J]. *Expert Systems with Applications*, 2022, 205:117709.
- [29] Jahanshahih, Orozco-Lopez O, Munoz-Pacheco J M, et al. Simulation and experimental validation of a non-equilibrium chaotic system[J]. *Chaos, Solitons and Fractals*, 2021, 143:110539.
- [30] Khandelwal J, Kumar Sharma V, Singh D, et al. DWT-SVD based image steganography using threshold value encryption method[J]. *Computers, Materials Continua*, 2022, 72:3299–3312.
- [31] 刘帅, 邓文博, 刘福才. 基于超混沌的三层量子图像加密算法研究[J]. *高技术通讯*, 2023, 33(2):167–175.
- [32] Xu G, Wu H Z, Shi Y Q. Ensemble of CNNs for steganalysis: an empirical study[C] // *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security*. Vigo Galicia, Spain: ACM, 2016:103–107.
- [33] Boroumand M, Chen M, Fridrich J. Deep residual network for steganalysis of digital images[J]. *IEEE Transactions on Information Forensics and Security*, 2018, 14(5): 1181–1193.
- [34] Zeng J, Tan S, Liu G, et al. WISERNet: wider separate-then-reunion network for steganalysis of color images[J]. *IEEE Transactions on Information Forensics and Security*, 2019, 14(10):2735–2748.

Research on image steganography algorithm based on chaotic mapping and invertible neural network

Liang Menghua, Zhao Hongtu

(School of Physics and Electronic Information Engineering, Henan Polytechnic University, Jiaozuo 454001)

Abstract

To address the problem of image quality degradation caused by large-scale steganography, a new reversible steganography network is designed in this paper. This network combines chaotic mapping and invertible neural network, greatly improving the security of steganography while ensuring large-scale capacity. Firstly, a four-dimensional chaotic system with line balance is designed. The contents of the secret image are scrambled using a chaotic mapping encryption algorithm to prevent secret image information leakage during transmission. In addition, the cover image is preprocessed to hide the secret image in the Y channel, significantly enhancing the steganographic capacity. Extensive experiments have been conducted on the DIV2K and COCO datasets. The peak signal-to-noise ratio (PSNR) of the cover image and stego image is as high as 41.63 dB, and that of the secret image and recovered image is 43.29 dB. The experimental results demonstrate that the proposed method not only maintains excellent steganographic performance under large capacity steganographic conditions but also produces steganographic images with high fidelity. Furthermore, its anti-steganographic capability is far superior to existing advanced algorithms.

Key words: image steganography, chaotic mapping, invertible neural network, image quality, anti-steganography analysis, robustness