

基于深度残差网络的医学图像鲁棒可逆水印算法

李 智^{1,2*}, 周旭阳^{1,2}, 殷昕旺^{1,2}, 张 丽^{1,2}

(1.贵州省智能医学影像分析与精确诊断重点实验室,贵州 贵阳 550025; 2.贵州大学 计算机科学与技术学院,贵州 贵阳 550025)

摘要:医学图像作为医生诊断的重要依据,其版权保护一直是研究重点。在对医学图像进行版权保护时,必须保证图像修改后能无损恢复。但是,当前大多数应用于医学图像的可逆水印算法并未考虑鲁棒性。因此,本文针对具有一定鲁棒性的可逆水印算法进行研究,提出一种基于深度残差网络 ResNet 的鲁棒可逆医学图像水印算法。首先,算法利用深度残差模型提取医学图像的深度特征信息,自适应确定最优嵌入强度,均衡水印的不可见性和鲁棒性;此外,结合遗传算法和模糊 C-均值的聚类算法对水印区域动态划分,根据聚类结果提取水印信息,有效克服信号攻击对含水印图像的影响,提高可逆水印算法的鲁棒性。实验结果表明:在嵌入水印后,含水印医学图像的 PSNR 值都可以达到 36 dB 以上,具有较好的图像质量;在提取水印后,算法可以完全无损恢复医学图像;在受到高斯噪声、椒盐噪声、JPEG 压缩等常见信号攻击后,算法仍然可以准确提取水印信息,表现出较强的鲁棒性。

关键词:医学图像;版权保护;深度残差网络;聚类算法;鲁棒可逆水印

中图分类号:TP309 **文献标识码:**A

随着网络技术的快速发展,智能医学和远程诊断技术日趋成熟。大量医学图像经常在网络上进行传输和使用,未经授权者可轻易通过网络获取、存储、使用和篡改网络上的医学图像^[1]。因而,保护医学图像的版权信息显得十分重要。数字水印算法是一种常用的信息隐藏技术,可用于医学图像的版权保护^[2]。

在对医学图像进行保护时,为了不影响医生的诊断,不可破坏原始医学图像的信息,因此,基于医学图像的可逆水印算法成为研究者们关注的重点。郑洪英等^[3]提出了基于位平面的可逆信息隐藏算法。首先,将医学图像分解为八位平面,通过压缩最高的四位平面获得对空间进行像素填充后的重建图像;其次,分别对重建图像的头、中间、尾部进行加密;再次,利用直方图移位的方法将水印信息嵌入图像中。DENG 等^[4]针对医学图像的分区域典型特征,提出一种基于直方图平移的高容量无损信息隐藏水印算法。利用最大类间距方法确定前景区域和背景区域,通过使用聚合多边形和图像

拟合算法确定前景嵌入区域,最终在前景和背景区域分别嵌入不同的水印。李智等^[5]提出基于实质区域的精确分割算法获取医学图像中的实质区域为嵌入区域,以及基于隶属度的不规则实质区域拟合方法,并将多比特的基于编码的直方图平移(Code based Histogram Shifting, CHS)算法应用于整数小波变换中高频子带,实现可逆水印嵌入。同时,使用增强奇异值分解在整数小波变换低频子带构建零水印,实现医学图像的版权保护和篡改定位。现有的医学图像可逆水印算法大多在无损环境中运行,即载体在嵌入信息后不能受到任何攻击和修改。而在现实场景中,时常出现诸如图像压缩和几何变换等图像操作和攻击^[6-7]。

在图像水印算法的研究中,通常是利用人工设定水印的嵌入强度参数,但是人工设定的参数具有较强的随机性,且得到的参数无法较好地均衡水印不可见性和鲁棒性^[8]。水印的嵌入强度越大,则水印的鲁棒性越强,但不可见性越差;水印嵌入强度越小,则水印的不可见性越好,但鲁棒性就越弱^[9]。

收稿日期:2019-12-25

基金项目:国家自然科学基金项目资助(61462013)

作者简介:李 智(1977-),女,副教授,博士,研究方向:医学影像分析、计算机视觉、信息隐藏,Email: zhili@gzu.edu.cn.

* 通讯作者:李 智, Email: zhili@gzu.edu.cn.

文献[10]对载体图像进行 Contourlet 变换后,对低频部分做块奇异值分解,其水印的主成分是通过修改块的最大奇异值的方式进行嵌入,虽然在数值上达到了不可见性的标准,但具有明显的块效应。文献[11]提出了一种基于奇异值分解和蜂群优化的鲁棒水印算法。嵌入强度的参数采用蜂群优化算法来选取,自适应均衡水印算法的鲁棒性与透明性,但蜂群优化算法收敛速度慢,寻找最优解时间较长。文献[12-13]基于群智能算法对水印嵌入强度进行优化,能够根据不同的图像确定最优的嵌入强度,但这些算法普遍不能较好地抵抗信号攻击。人类视觉系统(Human Visual System, HVS)具有纹理掩蔽、频率掩蔽、亮度掩蔽等特性^[14]。文献[15]基于 HVS 特性确定图像掩蔽因子,并将其作为水印嵌入的强度,此类算法能降低载体图像的视觉失真,但算法较为复杂。ResNet 是一种在图像处理领域应用的深度学习方法,它不仅提取图像的高维复杂特征,而且可以解决网络层数增多引起的精度退化问题,提升深度网络的性能^[16-17]。

在平衡水印的不可见性和鲁棒性的同时,以上文献有一个共同特征为鲁棒性并不能满足需求。文献[18]在提取水印信息过程中引入 K -means 聚类方法,实现了水印区域的动态划分,但是初始聚类中心的选择对聚类结果影响很大,这就会造成水印算法性能不稳定。遗传模糊 C-均值作为另一种聚类算法^[19],先应用遗传算法确定最优初始化聚类中心,再使用模糊 C-均值(Fuzzy C-Means, FCM)方法得到最终的聚类结果,聚类结果不受聚类中心选择的影响,可解决水印算法性能不稳定的问题。

结合以上问题,本文提出一种基于 ResNet 的医学图像鲁棒可逆水印算法,利用改进的 ResNet 计算水印嵌入强度以平衡水印的不可见性和鲁棒性。在提取水印信息过程中,利用遗传模糊 C-均值方法,有效提高了水印提取算法的鲁棒性。从嵌入和提取两个角度使基于医学图像的鲁棒可逆水印具有一定的实用性。

1 相关技术

1.1 残差学习

深度学习的基础网络从 AlexNet(5 个卷积层)、VGG(19 个卷积层)到 GoogLeNet(22 个卷积层),网络的结构在不断变深,更深的网络可以提取更复杂的特征^[20-21]。但是,随着网络的加深,出现

了训练集准确率下降的现象。HE 等^[22]提出 ResNet,该网络结构可以避免简单堆叠的卷积神经网络梯度消失或爆炸以及精度退化问题,模型更容易优化,性能提升明显。

ResNet 引入了残差学习,令 \mathbf{x} 表示输入, $H(\mathbf{x})$ 表示残差单元的输出。一般情况下,卷积神经网络直接通过训练来学习 $H(\mathbf{x})$,而残差学习则是使用多个含有参数的网络层来学习输入和输出之间的残差, $F(\mathbf{x}) := H(\mathbf{x}) - \mathbf{x}$,那么残差单元的输出最终变为 $F(\mathbf{x}) + \mathbf{x}$ 。实验证明,残差函数 $F(\mathbf{x})$ 比 $H(\mathbf{x})$ 更容易优化和学习。

一个完整的残差单元结构如图 1 所示。

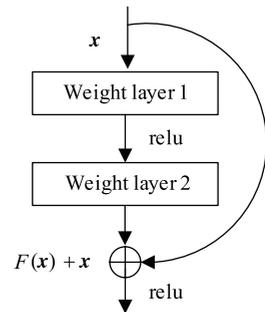


图 1 残差学习单元

Fig.1 Residual learning unit

在本文,残差单元定义为

$$\mathbf{y} = F(\mathbf{x}, \{W_i\}) + \mathbf{x}. \quad (1)$$

式中: \mathbf{x} 和 \mathbf{y} 分别为残差单元的输入和输出; $F(\mathbf{x}, \{W_i\})$ 为当前网络想要学习的残差函数。

定义:

$$F(\mathbf{x}) = W_2 \sigma(W_1 \mathbf{x}). \quad (2)$$

式中: σ 为 relu 激活函数, W_1 和 W_2 分别为 layer1 和 layer2 的权重。为了简化表示,省略了偏差。

$F(\mathbf{x}) + \mathbf{x}$ 的操作是通过 1 个“跳跃连接”将对应元素进行相加来执行的。这种计算方式既没有引入额外的参数,也不增加计算复杂度,可解决层数增加之后出现的性能退化问题。式(1)中的输入向量 \mathbf{x} 和函数 F 的维度应保持一致,否则,对输入向量 \mathbf{x} 执行线性投影 W_s 来实现维度匹配,即

$$\mathbf{y} = F(\mathbf{x}, \{W_i\}) + W_s \mathbf{x}. \quad (3)$$

1.2 遗传模糊 C-均值算法

FCM 算法是一种有效的聚类算法。其基本思想:将 n 个样本数据 $N = \{n_1, \dots, n_i, \dots, n_n\}$ 分为 c 类,并求得聚类中心 $V = \{v_1, \dots, v_j, \dots, v_c\}$, N 中任意样本 n_i 对 j 类的隶属度为 u_{ij} ,分类结果可以表

示为模糊隶属度矩阵 $U = \{u_{11}, \dots, u_{ij}, \dots, u_{nc}\}$ 。FCM 是通过最小化隶属度矩阵 U 和聚类中心 V 的目标函数 $J_{FCM}(U, V)$ 来实现:

$$J_{FCM}(U, V) = \sum_{i=1}^n \sum_{j=1}^c (\mu_{ij})^m d_{ij}^2, \quad (4)$$

约束条件:

$$\sum_{j=1}^c u_{ij} = 1, u_{ij} \in [0, 1] \quad 1 \leq i \leq n, 1 \leq j \leq c. \quad (5)$$

式中:参数 $m > 1$ 为模糊系数,用来控制隶属矩阵 U 的模糊程度, m 越大越模糊,通常 $m = 2$ 是比较理想的取值; $d_{ij} = \|n_i - v_j\|$ 为 n_i 到聚类中心 v_j 之间的欧式距离,应用拉格朗日乘数法并结合约束条件,使目标函数 $J_{FCM}(U, V)$ 得到最小值的必要条件如式(6)(7)所示。对式(6)和(7)进行迭代运算,直至算法收敛。

$$\mu_{ij} = 1 / \sum_{k=1}^c \left(\frac{d_{ij}}{d_{ik}} \right)^{\frac{2}{m-1}}, \quad (6)$$

$$v_j = \sum_{i=1}^n (\mu_{ij})^m x_i / \sum_{i=1}^n (\mu_{ij})^m. \quad (7)$$

FCM 算法对初始聚类中心敏感,容易收敛于局部最优解。遗传算法(Genetic Algorithm, GA)具有领域无关的群体性全局搜索能力,将遗传算法与模糊聚类算法进行组合,可以有效地解决 FCM 算法局部寻优的缺点。由遗传算法生成最优初始聚类中心,再使用 FCM 算法得到最终的分类结果。遗传 FCM 算法的处理流程如图 2 所示。

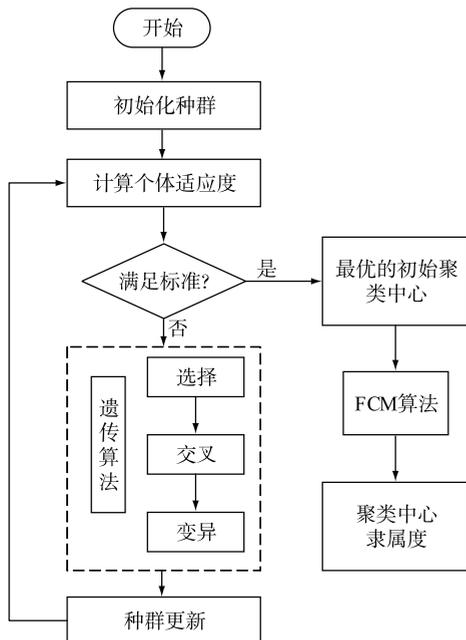


图 2 遗传模糊 C-均值流程图

Fig.2 Flow chart of genetic fuzzy C-mean

2 算法实现

2.1 深度残差网络模型

在现有算法中,大多是依靠经验手动设置嵌入强度,不仅没有理论依据,也很难获取最优嵌入强度平衡不可感知性和鲁棒性。为了实现不可感知性与鲁棒性的有效折衷,本文提出一种深度残差网络模型。它通过残差学习可以提高提取图像尺度、亮度、纹理等复杂特征的能力,更为精确地评估图像对噪声的局部敏感性,自适应获得水印嵌入强度 λ 。

此深度残差网络模型以医学图像作为输入,由医学图像对应的嵌入强度作为标签,学习医学图像和图像嵌入强度之间的映射关系。模型训练完成后,可以直接通过医学图像预测嵌入强度。由于医学图像的大小一般较大,为了提高网络的拟合速度和增加样本数量,医学图像被分成 32×32 的图像块作为输入。

2.1.1 网络结构

如图 3 所示,网络模型一共有 19 层,包括卷积层、残差单元、全局平均池化层和全连接层。网络输入大小为 32×32 的医学图像。首先,通过的 1 个卷积层(使用大小为 $3 \times 3 \times 16$ 的卷积核);其次,经过 9 个残差单元(分别使用大小为 $3 \times 3 \times 16$ 、 $3 \times 3 \times 32$ 、 $3 \times 3 \times 64$ 的卷积核),为了获取更多、更丰富的特征信息,卷积核的数目随着网络的深入不断增加;再次,网络以全局平均池化层和全连接层结束,输出大小为 1×1 的嵌入强度。网络通过步长为 2 的卷积层直接进行下采样。为了保持特征图的大小与输入一致,将步长和边缘填充都设置为 1。

残差单元由 2 个卷积层和 1 个“跳跃连接”组成,卷积层主要有大小为 3×3 的卷积核,并遵循 2 个简单的设计规则^[22]:(i)当输入和输出特征图的大小相同时,该卷积层和上一层具有相同数量的卷积核;(ii)如果特征图的大小减半,则卷积核的数量加倍,以便保持每层的时间复杂度。在跳跃连接中,当输入和输出的维度是相同时,可以直接使用式(1)连接(如图 3 实曲线);当维度增加一倍时,用式(3)中的线性投影匹配维数(如图 3 虚曲线)。线性投影通过卷积核大小为 1×1 、步长为 2 的卷积实现。

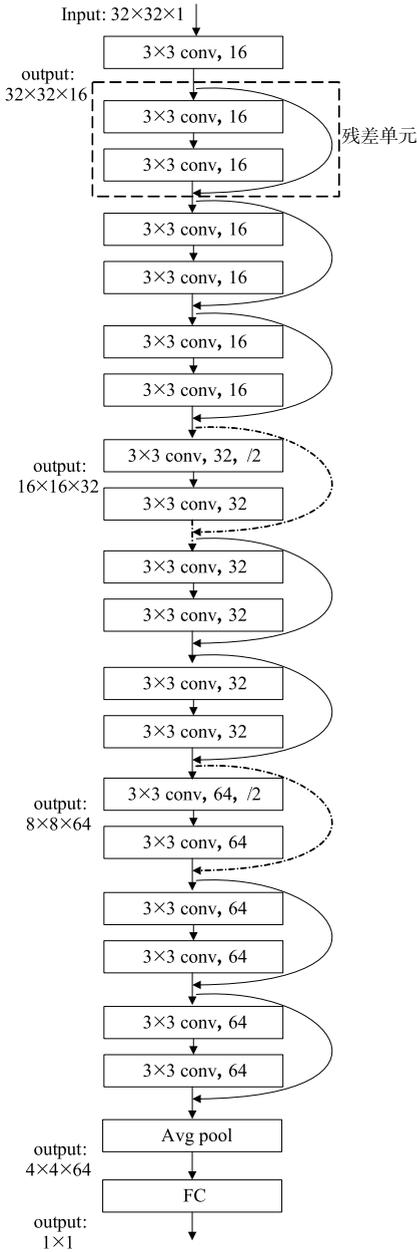


图3 深度残差网络结构

Fig.3 Deep residual network structure

2.1.2 网络训练

卷积神经网络的训练过程就是损失函数的最小化过程,本文的深度残差网络模型将预测嵌入强度和标签嵌入强度的均方误差作为损失函数:

$$l(W) = \frac{1}{n} \sum_{i=1}^n \|f(W, x_i) - y_i\|^2. \quad (8)$$

式中: n 为批处理量大小, W 为权重参数集合, x_i 为批处理图像中第 i 幅医学图像, y_i 为相对应 x_i 的嵌入强度, $f(W, x_i)$ 为预测第 i 幅医学图像的嵌入强度。

本文采用随机梯度下降法最小化损失函数。在每个卷积之后和激活函数之前,我们采用批量归一化(Batch Normalization, BN)。BN层有助于网络的收敛,批量大小设置为64,即在样本图像中随机选取64个图像和与之对应的标签嵌入强度作为一个批量进行网络训练。网络一共训练了200 epoch,学习率从0.1开始,当误差停滞时,将学习率除以10。

2.2 水印嵌入算法

步骤1 预处理医学图像。医学图像 I 的大小为 $2M \times 2N$, 深度为 t bit。为了避免像素溢出的问题,实现算法的可逆性,在嵌入水印之前,对医学图像 I 的像素进行调整:

$$I'(i, j) = \begin{cases} I(i, j) - \eta, & \text{if } I(i, j) > 2^t - 1 - \eta \\ I(i, j) + \eta, & \text{if } I(i, j) < \eta \end{cases}. \quad (9)$$

式中: $I(i, j)$ 为医学图像 I 在 (i, j) 处的像素值, $I'(i, j)$ 为调整后的像素值, i, j 为像素坐标,且 $1 \leq i \leq 2M, 1 \leq j \leq 2N$; η 是调整尺度,且 $\eta \geq \lambda$ 。

步骤2 构造小波系数均值(Mean of Wavelet Coefficients, MWC)直方图。算法选择MWC直方图作为嵌入区域,有助于实现水印图像的不可感知性。利用整数小波变换(Integer Wavelet Transform, IWT)将预处理后的医学图像 I' 进行分解,然后选取小波子带中的低高频子带(LH)和高低频子带(HL),将其互不重叠地分割成大小为 $h \times w$ 的子带块,计算每个子带块的MWC,并构造MWC直方图,定义:

$$S_k = \frac{1}{(h-2) \times (w-2)} \sum_{u=2}^{h-1} \sum_{v=2}^{w-1} P_k^{(u,v)}. \quad (10)$$

式中: S_k 为第 k 块MWC, $P_k^{(u,v)}$ 为第 k 块中 (u, v) 位置的小波系数。

步骤3 利用阈值约束选取感兴趣块。MWC直方图服从零均值的类拉普拉斯分布,如图4所示。选取MWC直方图峰值及其邻域作为感兴趣块,用于嵌入水印信息。图中的阴影部分表示选取的感兴趣块。采用阈值约束选取感兴趣块 S^{ROI} , 定义:

$$d(x, S_k) = |x - S_k|, \quad (11)$$

$$d(x, S_k) \leq \delta, 1 \leq k \leq n. \quad (12)$$

式中: $d(\cdot)$ 为Euclidean距离函数, $x \in \{x_l, x_r\}$ 为MWC直方图的两个峰值点, δ 为预先定义的阈值,并可以通过调整阈值 δ 来灵活控制水印容量。

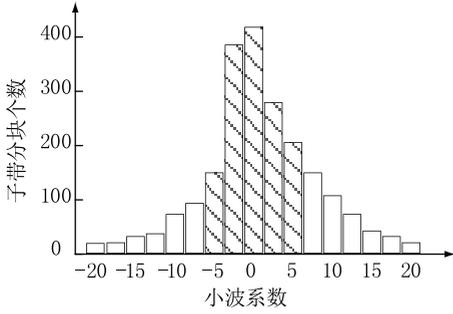


图 4 MWC 直方图

Fig.4 MWC histogram

步骤 4 基于深度残差网络计算水印嵌入强度 λ 。利用深度残差网络可以更好地提取医学图像尺度、亮度和纹理敏感度等复杂特征,训练医学图像和嵌入强度的关系模型,如图 3 所示。使用训练好的深度残差网络模型对医学图像进行计算得到嵌入强度 λ 。

步骤 5 水印嵌入模型。对步骤 3 所得的感兴趣块嵌入水印信息,嵌入过程定义:

$$S_k^w = S_k^{\text{ROI}} + \beta \lambda b_k, \quad (13)$$

$$\beta = \frac{S_k^{\text{ROI}} - S^*}{\text{abs}(S_k^{\text{ROI}} - S^*)}, \quad (14)$$

$$S^* = \arg \min_{x \in \{x_l, x_r\}} \{d(x, S_k^{\text{ROI}})\}. \quad (15)$$

式中: S_k^w 为第 k 个嵌入水印后感兴趣块的 MWC, S_k^{ROI} 为第 k 个感兴趣块的 MWC, λ 为基于深度残差网络计算得到的嵌入强度, b_k 为第 k 位水印信息。

步骤 6 重构医学图像。对嵌入水印后的小波子带进行 IWT 重构,即可得到嵌入水印后的医学图像 I^w 。需要说明的是,边信息需要作为密钥传送到接收方,包括子带分块大小 $h \times w$ 、水印嵌入强度 λ 和调整的像素位置等。

2.3 水印提取算法

步骤 1 获取感兴趣块。首先,对嵌入水印后的医学图像 I^w 进行 IWT 分解,选取小波子带中的低高频子带(LH)和高低频子带(HL)互不重叠地分成 n 块;其次,计算每个块 MWC,并构造 MWC 直方图;再次,利用阈值约束选取嵌入水印后的感兴趣块 $S^w = [S_1^w, \dots, S_k^w, \dots, S_n^w]$ 。

步骤 2 使用遗传模糊 C-均值聚类算法对 S^w 聚类。设定聚类个数为 3,类集合为 $\text{class} = \{\text{class I}, \text{class II}, \text{class III}\}$,采用遗传模糊 C-均值聚类算法对嵌入区域进行动态划分,嵌入区域划分的具体步骤:

输入:样本集合 $S^w = [S_1^w, \dots, S_k^w, \dots, S_n^w]$, 聚类数目 $c = 3$, 迭代次数 $T = 20$

输出:隶属度矩阵 U , 聚类中心 V

GA 算法(部分):

Step1 设置初始迭代次数 $t = 0$,并随机初始化种群(种群规模 $N = 50$);

Step2 计算个体适应度,其适应度函数 $f = 1/J_{\text{FCM}}$, J_{FCM} 如式(4)所示;

Step3 选择操作,采用转盘式选择策略^[23];

Step4 交叉操作,采用均匀杂交策略^[23],杂交概率 $p_c = 0.6$;

Step5 变异操作,以变异概率 $p_m = 0.2$ 将所选个体取反,获得新一代的种群, $t = t + 1$;

Step6 如果 t 大于最大迭代次数 T ,则停止迭代,得到最优初始聚类中心,否则返回 Step2 继续迭代;

FCM 算法(部分):

Step7 将最优初始聚类中心代入输入到 FCM 算法;

Step8 利用式(7)更新聚类中心 V ;

Step9 利用式(6)计算隶属度矩阵 U ;

Step10 计算式(5),如果目标函数收敛,则算法停止,输出隶属度矩阵 U 和聚类中心 V ;否则返回 Step8。

基于聚类结果,MWC 直方图被分成了三个区域,依次记为 class I, class II 和 class III,如图 5 所示。

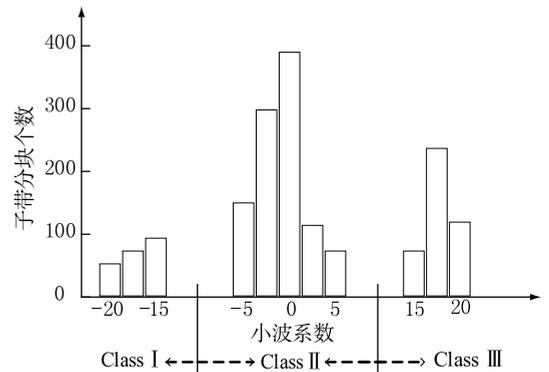


图 5 嵌入水印图像 MWC 直方图

Fig.5 MWC histogram of embedded watermark image

步骤 3 根据聚类结果提取水印信息。提取水印模型为

$$b_k^r = \begin{cases} 0, & \text{if } S_k^w \in \text{class II}, \\ 1, & \text{if } S_k^w \in \text{class I or class III}. \end{cases} \quad (16)$$

式中: S_k^w 表示第 k 个嵌入水印后感兴趣块的 MWC, b_k^r 为提取出的第 k 位水印信息。

步骤4 恢复感兴趣块。恢复公式为

$$S_k^r = S_k^w - \beta \lambda b_k^r. \quad (17)$$

式中 S_k^r 为第 k 个恢复后感兴趣块的 MWC。

步骤5 复原医学图像 I 。恢复后的小波子带经过 IWT 重构之后,然后采用式(9)的逆操作来恢复嵌入过程中调整的像素,进而得到复原医学图像 I 。

3 实验结果分析

为了验证本文水印算法的高效性,首先,分析参数对性能的影响,为参数的优化选择提出建议;其次,给出了鲁棒性测试、可逆性测试、不可感知性测试的实验仿真结果;再次,分别与基于传统方法和基于深度学习方法的鲁棒可逆水印算法^[18,24]进行性能对比。实验采用医学图像数据库(Medical Image Database, MID)的 DICOM 样本图像集中的 300 幅磁共振成像(Magnetic Resonance Imaging, MRI)图像,本文选取大小为 512×512 的头、肺、腹腔、肝脏等不同部位的 MRI 图像作为载体,如图 6 所示。水印信息采用伪随机二进制序列。

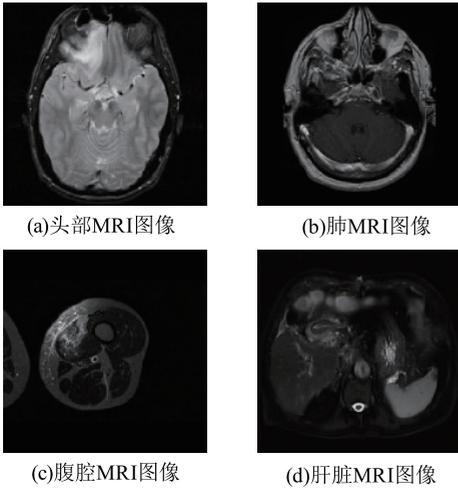


图6 测试图像

Fig.6 Test image

3.1 性能评测指标

本文的关键问题是要同时实现水印算法的可逆性和鲁棒性,即:在无损环境下可以恢复宿主图像与水印,以及当受到攻击后可以正确恢复水印。一般来说,鲁棒可逆算法的评测指标主要包括:

(1) 可逆性

图像错误率(Image Error Rate, IER)来评估可逆性,其值越低,表明算法的可逆性越好。定义:

$$R_{IE} = \frac{N_{Err_img}}{N_{img}} \times 100\%. \quad (18)$$

式中: N_{Err_img} 为恢复图像错误的像素点数, N_{img} 为载体图像的总像素点数。

(2) 不可感知性

不可感知性通常用宿主图像与水印图像间的峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)来衡量,定义:

$$R_{PSN} = 10 \times \lg \frac{2M \times 2N \times 255^2}{\sum_{i=1}^{2M} \sum_{j=1}^{2N} [I(i,j) - I^w(i,j)]^2}. \quad (19)$$

式中: $I(i,j)$ 和 $I^w(i,j)$ 分别为嵌入水印前后图像中 (i,j) 位置的像素值, $2M \times 2N$ 为图像大小。PSNR 值越高,说明水印图像的视觉质量越好,隐秘信息越不容易被察觉,反之亦然。通常认为当 PSNR 值大于 35 dB 时,图像的差异通过肉眼已经无法进行判别。

(3) 鲁棒性

在鲁棒性方面,主要考虑应对高斯噪声(方差为 0.01)、椒盐噪声(方差为 0.005)、JPEG 压缩(质量因子为 25)和 JPEG 2000 压缩。采用比特误差率(Bit Error Rate, BER)来衡量提取水印的正确性,定义:

$$R_{BE} = \frac{N_{Err}}{N_{bits}} \times 100\%. \quad (20)$$

式中: N_{Err} 为提取水印信息的错误比特数, N_{bits} 为嵌入水印信息的总比特数。BER 值越低,表明提取水印的正确性越高,水印抗攻击的鲁棒性越好。

(4) 容量

容量反映在嵌入过程中宿主图像能嵌入的最大信息数量。一般在嵌入过程中,水印常常会和边信息一起进行嵌入,这时称宿主图像中实际嵌入的最大水印位数为纯容量。

3.2 参数分析

3.2.1 嵌入强度

嵌入强度代表着嵌入水印的强弱,影响着可逆性、鲁棒性和不可感知性。本文实验选取小波子带中的低高频子带(LH)和高低频子带(HL),子带分块大小为 8×8 ,阈值为 10,当嵌入强度选取 5、10、15、20、25 时,不同嵌入强度相对应的可逆性、鲁棒性和不可感知性测试如图 7、8、9 所示。由图可知,随着水印的嵌入强度越大,算法的鲁棒性和可逆性越来越强,但是不可感知性越来越差。

3.2.2 阈值

实验选取小波子带中的低高频子带(LH)和高低频子带(HL),子带分块大小为 8×8、嵌入强度为 10,当

阈值选取 2、5、10 时,相应嵌入水印信息后的 MWC 直方图如图 10 所示。由图可知,当阈值越小时,直方图聚类越明显,则可以更准确地提取水印信息。

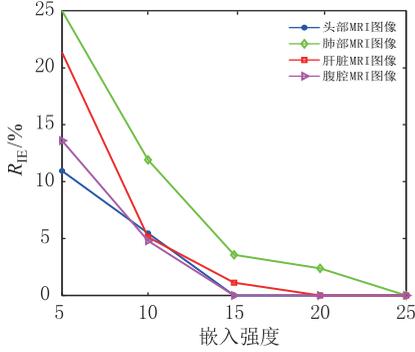


图 7 嵌入强度和可逆性的关系

Fig.7 Relationship between embedding strength and reversibility

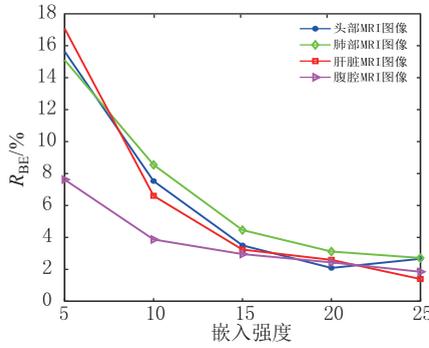


图 8 不同嵌入强度下抗 JPEG 的鲁棒性

Fig.8 Robustness against JPEG at different embedding strengths

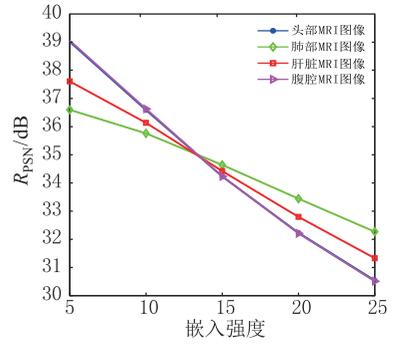


图 9 嵌入强度和 PSNR 的关系

Fig.9 Relationship between embedding strength and PSNR

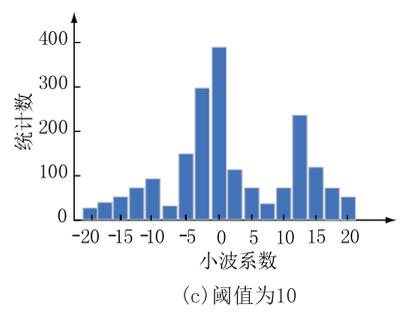
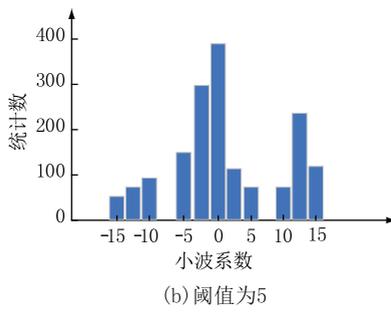
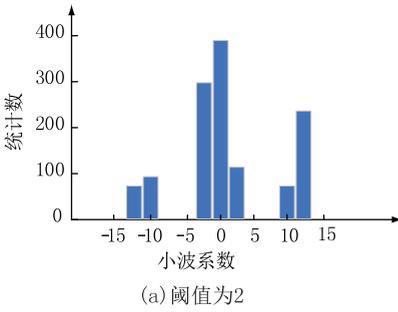


图 10 MWC 直方图

Fig.10 MWC histogram

图 11 反映阈值选取 2、5、10、15 时,不同阈值下算法的鲁棒性。随着阈值的减小,算法的鲁棒性越好,反之亦然。根据多次实验得出,当阈值 $\delta = \lambda/2$ 时,算法有较好的鲁棒性,且可保证较大容量。

3.2.3 子带分块大小

子带分块大小也是影响容量、不可感知性和鲁棒性的重要因素。实验选取小波子带中的低高频子带(LH)和高低频子带(HL)、嵌入强度为 10,当子带分块大小选取 4×4、8×8、16×16、32×32 时,研究子带分块大小对各性能的影响。

子带分块越大,子带分块个数越少,感兴趣块的个数就越少,水印嵌入容量随着块大小的增加而减少。由图 12 可以看出,随着块大小的减少,水印嵌入容量增加,反之亦然。由图 13 可以看出,随着子带分块大小的增加,PSNR 呈下降趋势。由图 14 可以看出,从整体上来说,随着子带分块大小的增加,抵抗攻击能力越强,鲁棒性呈增加趋势。

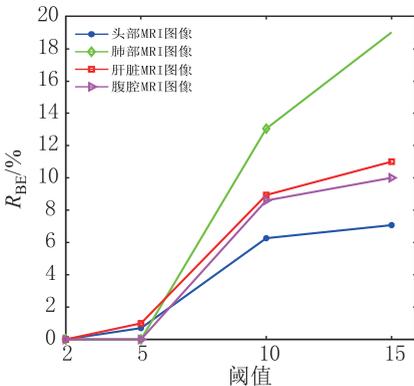


图 11 阈值和鲁棒性的关系

Fig.11 Relationship between threshold and robustness

3.3 实验仿真

为了验证本文算法的有效性,给出以不同部位医学图像为载体的水印嵌入和提取测试,实验结果如图 15 所示。图 15 中,(a1)、(b1)、(c1)、(d1)为

原始图像, (a2)、(b2)、(c2)、(d2) 为含水印的图像。从视觉效果来看, 含水印图像与原始图像相比没有明显变化, 具有良好的不可感知性; 采用 PSNR 客观评价含水印图像与原始图像的质量差别, 可以得出含水印图像的 PSNR 值分别为 38.9、36.5、37.5、39.0 dB, 都可以达到 36 dB 以上, 图像的质量都表现良好。图 15 中, (a3)、(b3)、(c3)、(d3) 为含水印图像与原始图像的差值图。可以明

显看出水印嵌入前后的图像差别, 以证明水印信息已嵌入载体图像。图 15 中, (a4)、(b4)、(c4)、(d4) 为提取水印信息后的恢复图像, (a5)、(b5)、(c5)、(d5) 为恢复图像与原始图像的差值图。差值图为全黑则表明恢复图像和原始图像完全一致, 在没有受到攻击的情况下, 算法没有更改原始图像的像素, IER 值为 0, 说明本文算法实现了完全可逆的效果。

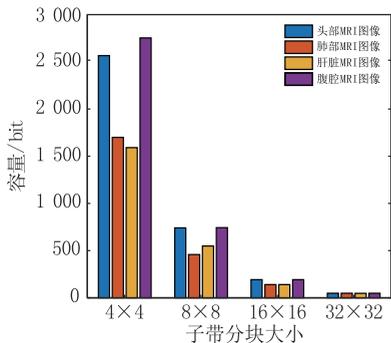


图 12 子带分块大小和容量的关系
Fig.12 Relationship between subband block size and capacity

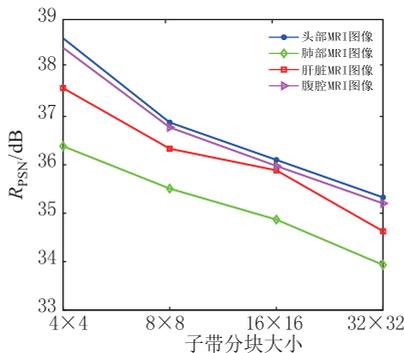


图 13 子带分块大小和 PSNR 的关系
Fig.13 Relationship between subband block size and PSNR

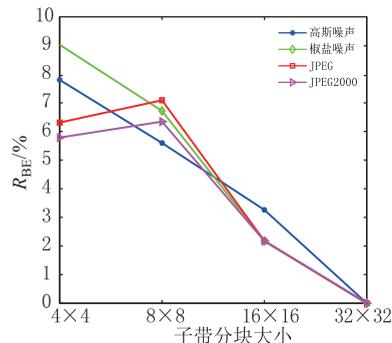


图 14 子带分块大小和鲁棒性的关系
Fig.14 Relationship between sub-band block size and robustness

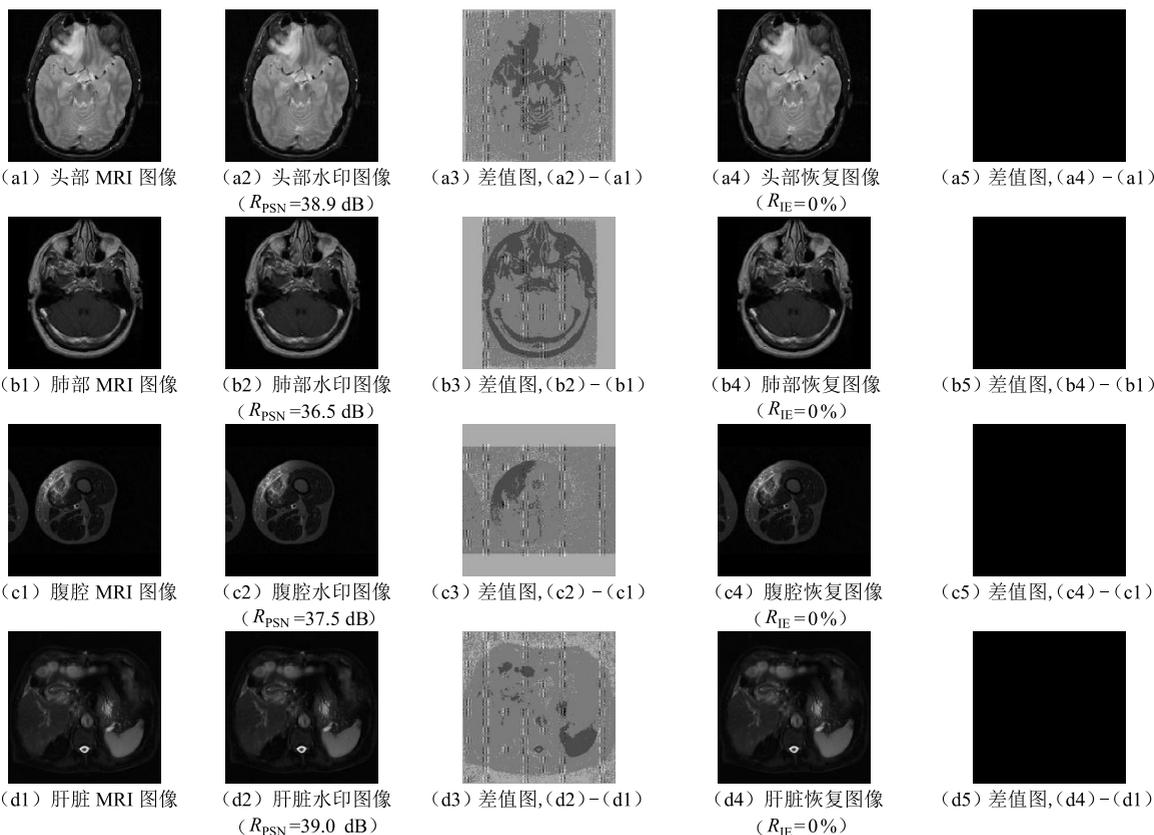


图 15 实验结果
Fig.15 Experimental results

为了检测本算法具有较好的鲁棒性,对含水印图像分别进行高斯噪声(方差为 0.01)、椒盐噪声(方差为 0.005)、JPEG 压缩(质量因子为 25)和 JPEG 2000 压缩等常规攻击。表 1 列出了含水印图像在受到不同攻击后所提取水印的 BER 值。从表可以看出:攻击对提取出的水印 BER 值有一定的影响,但 BER 值都低于 9%,特别是腹腔和肝脏受到压缩攻击时,提取出来的水印 BER 值大部分都能低于 1%,说明本算法可以有效抵抗各种常规攻击,具有较强的鲁棒性。

表 1 鲁棒性测试

Tab.1 Robustness test %

攻击测试	头部	肺部	腹腔	肝脏
未受攻击	0	0	0	0
高斯噪声	3.9	7.2	1.2	2.2
椒盐噪声	5.6	8.5	2.7	3.3
JPEG 压缩	2.6	6.7	0.9	0.7
JPEG 2000 压缩	2.6	7.2	0.9	0.6

3.4 实验对比

为了进一步说明本文算法的优越性,从可逆性、不可感知性、鲁棒性和容量 4 个方面与文献[18][24]进行性能对比。文献[18]是基于聚类和小波变换的鲁棒可逆水印算法,文献[24]是基于卷积神经网络的鲁棒可逆水印算法。

(1) 可逆性对比

我们采用 IER 来评估不同算法的可逆性,它在无损环境下衡量是否能实现图像与水印的无失真恢复。实验结果如表 2 所示,前 2 种算法接近实现可逆,相比之下,本文算法可以完全实现可逆效果。

表 2 可逆性对比

Tab.2 Comparison of reversibility %

测试图像	文献[18]	文献[24]	本文
头部 MRI 图像	0.4	0	0
肺部 MRI 图像	0.7	0.3	0
腹腔 MRI 图像	1.1	0.2	0
肝脏 MRI 图像	0.3	0	0

(2) 不可感知性对比

不可感知性是对水印图像的失真情况进行评估,表 3 对比了不同算法的 PSNR 值。可以看出,

本文的算法的 PSNR 值高达 39.0 dB,图像质量要好于文献[18]和文献[24]。

表 3 不可感知性对比

Tab.3 Comparison of imperceptibility dB

测试图像	文献[18]	文献[24]	本文
头部 MRI 图像	30.24	37.1	38.9
肺部 MRI 图像	31.86	34.9	36.5
腹腔 MRI 图像	31.02	36.4	37.5
肝脏 MRI 图像	30.26	38.0	39.0

(3) 鲁棒性对比

本文用 BER 对不同算法进行鲁棒性对比,表 4—表 7 显示了实验结果。可以看出:文献[18]抗椒盐噪声和高斯噪声的鲁棒性比较差,本文算法要优于所对比的方法,尤其在抗 JPEG 压缩和 JPEG2000 压缩时,肝脏图像的错误率可以低至 0.7%和 0.6%。

表 4 鲁棒性对比-椒盐噪声(方差=0.005)

Tab.4 Robustness comparison-salt and pepper noise (variance=0.005) %

测试图像	文献[18]	文献[24]	本文
头部 MRI 图像	14.5	8.9	5.6
肺部 MRI 图像	10.8	9.6	8.5
腹腔 MRI 图像	10.0	4.5	2.7
肝脏 MRI 图像	8.5	5.1	3.3

表 5 鲁棒性对比-高斯噪声(方差=0.01)

Tab.5 Robustness comparison-Gaussian noise (variance=0.01) %

测试图像	文献[18]	文献[24]	本文
头部 MRI 图像	20.8	4.3	3.9
肺部 MRI 图像	18.2	6.9	7.2
腹腔 MRI 图像	18.6	2.5	1.2
肝脏 MRI 图像	17.5	3.6	2.2

表 6 鲁棒性对比-JPEG 压缩(质量因子=25)

Tab.6 Robustness comparison-JPEG compression (quality factor=25) %

测试图像	文献[18]	文献[24]	本文
头部 MRI 图像	11.7	6.9	2.6
肺部 MRI 图像	10.8	13.9	6.7
腹腔 MRI 图像	9.7	5.8	0.9
肝脏 MRI 图像	9.7	6.4	0.7

表7 鲁棒性对比-JPEG2000 压缩

Tab.7 Robustness comparison-JPEG2000 compression %

测试图像	文献[18]	文献[24]	本文
头部 MRI 图像	12.7	5.6	2.6
肺部 MRI 图像	12.8	10.7	7.2
腹腔 MRI 图像	10.8	3.1	0.9
肝脏 MRI 图像	10.4	2.4	0.6

(4) 容量对比

表8对比了不同算法的容量,算法中设置块大小为 8×8 。本文算法的容量略高于文献[24],明显高于文献[18]。因为本文算法选取了2个小波子带嵌入水印信息,而文献[18]只使用了1个小波子带,所以本文算法容量近似文献[18]的2倍。

表8 容量对比

Tab.8 Capacity comparison bit

测试图像	文献[18]	文献[24]	本文
头部 MRI 图像	252	463	495
肺部 MRI 图像	190	278	289
腹腔 MRI 图像	191	342	364
肝脏 MRI 图像	256	471	498

4 结语

针对当前医学图像可逆水印算法抗攻击能力不足问题,本文提出一种基于深度残差网络的医学图像鲁棒可逆水印算法。首先,根据像素调整策略预处理医学图像,避免像素溢出;其次,利用深度残差模型自适应确定嵌入强度;再次,采用基于直方图平移和聚类算法嵌入与提取水印信息。实验结果表明:本文算法可以无损恢复医学图像,保证了医学图像的完整性;并且对于常见的信号攻击具有较好的稳健性,提高了可逆水印的鲁棒性;同时嵌入水印后的医学图像 PSNR 值均达到 36 dB 以上,较好地均衡了水印的不可见性和鲁棒性;此外,与现有算法相比,在嵌入容量上具有一定优势。本文算法在无损恢复医学图像的基础上,还可以抵抗常见的信号攻击,实现了对医学图像的版权保护。下一步的研究工作考虑将深度学习应用于整个医学图像版权保护算法中,在保证医学无损恢复的前提下,使得该算法对于攻击具有更好的普适性。

参考文献:

- [1] SINGH S, RATHORE V S, SINGH R. Hybrid NSCT domain multiple watermarking for medical images[J]. Multimedia Tools and Applications, 2016, 76(3):3557-3575.
- [2] LEI B, TAN E L, CHEN S, et al. Reversible watermarking scheme for medical image based on differential evolution[J]. Expert Systems with Applications, 2014, 41(7):3178-3188.
- [3] 郑洪英,任雯,程惠惠.基于位平面压缩的密文医学图像可逆信息隐藏算法[J].计算机应用,2016,36(11):3088-3092.
- [4] DENG X H, CHEN Z G, LIANG D Q, et al. Region-based lossless data hiding with high capacity for medical images[J]. Journal on Communications, 2015, 36(1):189-198.
- [5] 李智,陈怡,王丽会,等.基于实质区域的医学图像双水印算法研究[J].贵州大学学报(自然科学版),2018,35(5):55-62,73.
- [6] NAIK K, TRIVEDY S, PAL A K. An IWT based blind and robust image watermarking scheme using secret keymatrix[J]. Multimedia Tools and Applications, 2018, 77(11):13721-13752.
- [7] 熊祥光.空域强鲁棒零水印方案[J].自动化学报,2018,44(1):160-175.
- [8] 项世军,杨乐.基于同态加密系统的图像鲁棒可逆水印算法[J].软件学报,2018,29(4):957-972.
- [9] 李智,陈淑琴,程欣宇,等.ASIFT 与 SVD 相结合的 Contourlet 域的抗几何攻击视频双水印算法[J].贵州大学学报(自然科学版),2019,36(4):59-67.
- [10] VLEESCHOUWER C D, DELAIGLE J F, MACQ B M. Circular interpretation of bijective transformations in lossless watermarking for media asset management[J]. IEEE Transactions on Multimedia, 2003, 5(1):97-105.
- [11] NI Z, SHI Y Q, ANSARI N, et al. Robust lossless image data hiding[C]//2004 IEEE International Conference on Multimedia and Expo (ICME). Taipei: IEEE,2004: 2199-2202.
- [12] NI Z, SHI Y Q, ANSARI N, et al. Robust lossless image data hiding designed for semi-fragile image authentication[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(4):497-509.
- [13] ZOU D, SHI Y Q, NI Z, et al. A semi-fragile lossless digital watermarking scheme based on integer wavelet transform[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(10):1294-1300.
- [14] 毕胜,梁德群.基于人类视觉特性的纹理分割方法[J].计算机应用,2006(5):1015-1017.
- [15] AN L, GAO X, LI X, et al. Robust reversible watermarking via clustering and enhanced pixel-wisemasking[J]. IEEE Transactions on Image Processing, 2012, 21(8):3598-3611.
- [16] ZHU J, KAPLAN R, JOHNSON J, et al. Hidden: hiding data with deep networks[C]//Proceedings of the European Conference on Computer Vision (ECCV). Germany: Springer, 2018: 657-672.
- [17] MUN S M, NAM S H, JANG H, et al. Finding robust domain

- from attacks: a learning framework for blindwatermarking [J]. *Neurocomputing*, 2019, 337: 191-202.
- [18] AN L, GAO X, YUAN Y, et al. Robust lossless data hiding using clustering and statistical quantity histogram [J]. *Neurocomputing*, 2012, 77(1): 1-11.
- [19] 李同强, 周天弋, 吴斌. 基于改进遗传算法的加权模糊 C-均值聚类算法 [J]. *计算机应用*, 2009, 29(S2): 260-262.
- [20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C] // *Advances in neural information processing systems*. USA: ACM, 2012: 1097-1105.
- [21] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C] // *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. USA: IEEE Computer Society, 2015: 1-9.
- [22] HE K M, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. USA: IEEE Computer Society, 2016: 770-778.
- [23] 马永杰, 云文霞. 遗传算法研究进展 [J]. *计算机应用研究*, 2012, 29(4): 1201-1206, 1210.
- [24] KANDI H, MISHRA D, GORTHI S R K S. Exploring the learning capabilities of convolutional neural networks for robust image watermarking [J]. *Computers & Security*, 2017, 65: 247-268.

(责任编辑:周晓南)

Robust Reversible Watermarking Algorithm for Medical Images Based on Deep Residual Network

LI Zhi^{1,2*}, ZHOU Xuyang^{1,2}, YIN Xinwang^{1,2}, ZHANG Li^{1,2}

(1. Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province, Guiyang 550025, China;

2. College of Computer Science and Technology, Guizhou University, Guiyang 550025, China)

Abstract: As an important basis for doctors' diagnosis, the copyright protection of medical images has always been the focus of research. In the copyright protection of medical images, it is necessary to ensure the lossless recovery after image modification. However, most of the current reversible watermarking algorithms for medical images do not consider robustness. Therefore, this paper proposes a robust reversible medical image watermarking algorithm based on Deep Residual Network, ResNet. First, the algorithm uses depth residual model to extract depth feature information of medical images, and adaptively determines the optimal embedding strength, and balances the invisibility and robustness of watermarks. In addition, the genetic fuzzy C-means clustering algorithm is used to dynamically divide the watermark region and extract the watermark information according to the clustering results, so as to effectively overcome the influence of signal attack on the watermarked images and further improve the robustness of the algorithm. The experimental results show that the PSNR value of medical images with watermarks can reach above 36 dB after embedding watermarks, with enough image quality. After extracting the watermark, the algorithm can recover the medical image completely without loss. Under the attack of Gaussian noise, salt-and-pepper noise, JPEG compression and other common signals, the algorithm can still extract the watermark information accurately, showing strong robustness.

Key words: medical image; copyright protection; deep residual network; clustering algorithm; robust reversible watermark