基于人工蜂群算法的两阶段图像隐写分析算法

穆晓芳[1] 邓红霞[2] 李晓宾[3] 赵 鹏[4]

(太原师范学院计算机系 太原 030619)[1] (太原理工大学信息与计算机学院 太原 030024)[2]

(北京航空航天大学计算机学院 北京 100191)[3] (中国社会科学院 北京 100732)[4]

摘 要 为了提高图像隐写分析的检测准确率,提出了一种基于人工蜂群算法的两阶段图像隐写分析算法。第一阶段,设计了基于模糊理论的隐写模式检测算法,检测部分已知隐写算法的隐写内容;第二阶段,基于人工蜂群算法分析了含密图像的区域与密度双重特征,通过双重特征的分析检测未知隐写算法的嵌入内容。基于公开隐写图像数据集的实验结果表明,所提的两阶段隐写分析算法可获得较高的检测率,同时具有理想的计算效率。

关键词 人工蜂群算法,图像隐写分析,模糊理论,邻接像素,多特征分析

中图法分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.06.026

Two-phase Image Steganalysis Algorithm Based on Artificial Bee Colony Algorithm

MU Xiao-fang[1] DENG Hong-xia[2] LI Xiao-bin[3] ZHAO Peng[4]

(Department of Computer Science, Taiyuan Normal University, Taiyuan 030619, China)[1]

(College of Information and Computer, Taiyuan University of Technology, Taiyuan 030024, China)[2]

(School of Computer Science and Engineering, Beihang University, Beijing 100191, China)[3]

(Chinese Academy of Social Sciences, Beijing 100732, China)[4]

Abstract In order to improve the detection accuracy of the image steganalysis, this paper proposed a two-phase image steganalysis algorithm based on Artificial Bee Colony. In the first phase, steganography pattern detection algorithm based on fuzzy theory is designed to discover steganography content of some known steganography algorithms. In the second phase, dual features of regions and density of stego images are analyzed based on Artificial Bee Colony algorithm, and the embedded content of unknown steganography algorithms is analyzed by dual features. Experimental results on the public steganography images show that the proposed algorithm performs high detection accuracy, and it has desirable computational efficiency.

Keywords Artificial Bee Colony algorithm, Image steganalysis, Fuzzy theory, Adjacent pixels, Multi-feature analysis

1 引言

隐写分析是针对隐写算法的逆向分析技术,其目标是根据载体图像的统计特性判断其中是否存在隐蔽信息,进而估计嵌入的秘密信息量与隐写工具,并破坏或截获隐蔽信息[1-2]。目前,隐写分析算法主要分为专用隐写分析算法与通用隐写分析算法,专用隐写分析算法[3-4]对特定的隐写方法具有较高的准确性,但对其他方法无效;通用隐写分析算法[5-6]则可以同时检测不同的已知隐写算法,并且对未知隐写算法也具有一定效果。

通用隐写分析算法由于具有广泛的适用性,目前已成为隐写分析领域的研究重点。当前的通用隐写分析算法主要有3种类型:基于图像质量度量标准的隐写分析[7]、基于统计矩的隐写分析[8]以及基于相邻像素相关性的隐写分析[4]。文献[9]构建了一种能抵抗基于运动矢量的时空相关性隐写分析的视频隐算法,该算法将秘密信息嵌入到视频压缩过程中的熵编码之前的运动矢量残差中,能较好地保持运动矢量残差在隐写前后的直方图特征,具有较好的视觉不可见性。文献[10]提出了结合旋转森林变换与多分类器集成的隐写分析算法,相比于传统集成分类器和集成极限学习机分类器,该算法分别降低了3.2%与1.1%的误检率,能够有效提升集成分类器的检测精度。文献[11]提出了一种基于空间域富模型的彩色图像隐写分析方案,该方案对非自适应LSB匹配隐写算法与WOW隐写算法的检测性能较好。此类隐写分析算法对

到稿日期:2018-11-20 返修日期:2019-01-17 本文受国家自然科学基金项目(F020308),山西省重点研发计划项目(201803D31055),山西省自然科学基金项目(201801D121135)资助。

穆晓芳(1974-),女,硕士,副教授,硕士生导师,CCF会员,主要研究领域为分布式计算、图像处理、数据分析,E-mail:mu_xiao_fang@163.com(通信作者);邓红霞(1976-),女,博士,副教授,硕士生导师,CCF会员,主要研究领域为智能信息处理、脑认知研究、图像处理;李晓宾(1991-),男,博士生,主要研究领域为深度压缩、嵌入式智能硬件和视频图像处理;赵 鹏(1973-),男,博士生,教授,硕士生导师,CCF会员,主要研究领域为软件工程、数据分析、算法研究。

低嵌入率的检测性能较差[9-11]，此外，基于分类器的隐写分析方法[10-11]需选取有限个参与训练的隐写算法所生成的含密图像作为样本，这些样本无法涵盖所有含密图像的统计特性，因此分类器仅能检测来源于已训练隐写算法的含密图像。

为了提高通用隐写分析的准确率，设计一种两阶段的隐写分析算法。第一阶段隐写模式检测阶段，设计了封装型演化模式检测算法，对已知的隐写模式进行检查与分析；第二阶段为基于人工蜂群算法（Artificial Bee Colony algorithm，ABC）提取两种特征，通过融合两种特征来分析邻接像素之间的关系，并检测含密图像的含密区域。基于 BOSS benchmark 数据集的实验结果表明，所提算法具有较高的隐写分析准确率，优于 DRL 和 BS_RSVC 等优秀的隐写分析算法。

2　隐写模式检测

图 1 给出了本文隐写模式检测算法的流程框图。假设已知隐写方案的数量为 $S_i(i=1,\cdots,I)$，嵌入的隐秘数据量为 $L_j(j=1,\cdots,J)$，隐写分析方法的数量为 $A_k(k=1,\cdots,K)$，隐写方法的模式数量为 P_{ijk}，此外，需要为无密图像建立（$i\times J\times K$）个模型 M_{ijk}。

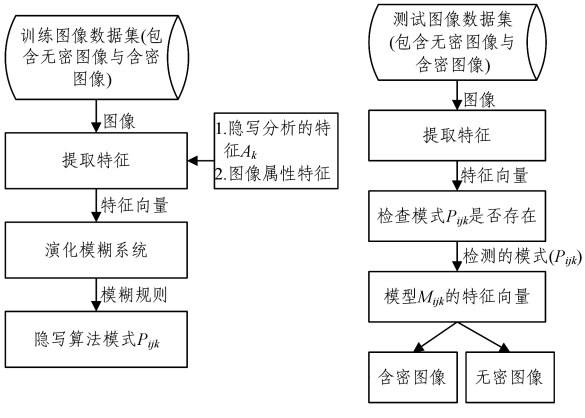


图 1　本文隐写模式检测算法的流程框图

Fig. 1　Flow diagram of proposed steganographic detection algorithm

2.1　生成特征向量

本文采用文献[12]和文献[13]中两个常用的隐写算法的特征，表 1 所列为两种隐写算法统计的所有特征。

表 1　两种隐写算法统计的特征结果

Table 1　Total features of two types of steganographic algorithms

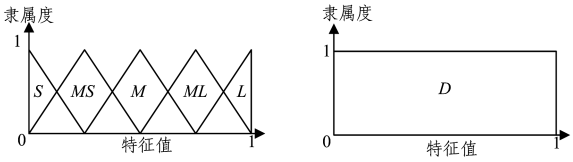
隐写算法	特征数量	特征类型
文献[12]	11	全局直方图
	66	5 AC 直方图
	99	11 个双直方图
	1	差异
	2	块效应
	25	共生矩阵
	81	Markov 特征
文献[13]	39	空域与离散小波变换直方图
	39	预测误差的直方图
	39	JPEG 格式的直方图
	78	JPEG 的水平二维直方图
	78	JPEG 的垂直二维直方图
	78	JPEG 的正交二维直方图
	39	JPEG 格式预测误差的直方图

2.2　生成模糊规则

将每个模糊规则编码为一条字串，从图像的特征向量生成模糊规则，采用以下简称表示 6 个编码：忽略（D）、小（S）、较小（MS）、中等（M）、较大（ML）、大（L）。

规则 R_j：IF（x_1 IS A_{j1} && … && x_n IS A_{jn}）THEN 图像为无密图像，且 $CF=CF_j$。其中，R_j 为第 j 个模糊规则的序号，$x_1,\cdots,x_n$ 为从图像中提取的特征，$A_{j1},\cdots,A_{jm}$ 则为区间[0，1]的值，代表{{S，MS，M，ML，L}，D}，如图 2 所示。如果一个特征的值为 D，则隶属函数 D 的值为 1，CF_j 是模糊规则的置信度因子，每个模糊规则具有一个置信度因子，表示该规则的置信度。

图 2 中每个置信度的隶属函数均匀地分区，将每个特征域指定为对称的三角模糊集。因为 n 维特征向量的可能模糊规则的总数量为 6^n，所以如果 n 值较大，将产生大量的模糊规则，本方法对于小规模模糊规则的性能较优。



(a) S，MS，M，ML，L 隶属函数　　(b) D 隶属函数

图 2　隶属函数

Fig. 2　Membership functions

通过以下 3 个步骤计算模糊规则的置信度。

步骤 1　计算模糊规则 R_j 每个训练样本 $x_p=(x_{p1},x_{p2},\cdots,x_{pn})$ 的兼容性：

$$\mu_j(x_p)=\mu_{j1}(x_{p1})\times\cdots\times\mu_{jn}(x_{pn}),P=1,2,\cdots,M \tag{1}$$

其中，$\mu_{ji}(x_{pi})$ 是第 p 个样本、第 i 个特征的隶属函数；M 表示样本总数量。

步骤 2　对于无密图像与含密图像，分别根据规则 R_j 计算训练样本的兼容度之和：

$$\beta_{\text{clean}}(R_j)=\frac{\sum\limits_{x_p\in\text{Clean}}\mu_j(x_p)}{N_{\text{clean}}} \tag{2}$$

$$\beta_{\text{stego}}(R_j)=\frac{\sum\limits_{x_p\in\text{Stego}}\mu_j(x_p)}{N_{\text{Stego}}} \tag{3}$$

其中，$\beta_{\text{clean}}(R_j)$ 与 $\beta_{\text{stego}}(R_j)$ 分别是无密图像与含密图像的训练样本兼容度之和。

步骤 3　无密图像的置信度 CF_j 的计算式为：

$$CF_j=\frac{(\beta_{\text{Clean}}(R_j)-\beta_{\text{Stego}}(R_j))}{(\beta_{\text{Clean}}(R_j)+\beta_{\text{Stego}}(R_j))} \tag{4}$$

2.3　演化模糊算法

本文所提演化模糊算法每轮迭代中优化一条模糊规则、迭代地学习全部模糊规则。首先，所有训练样本的权重相同，通过图像的特征向量初始化每个模糊规则，每轮迭代中将适应度最高的模糊规则作为该迭代的输出结果；然后，在学习过程中降低正确学习的训练样本权重，由此实现演化过程。

本学习系统的演化过程采用的适应度函数为：

$$f_p=\frac{\sum\limits_{x^k\in\text{Clean}}\omega^k\mu_{R_i}(x^k)}{\sum\limits_{x^k\in\text{Clean}}\omega^k} \tag{5}$$

$$f_N=\frac{\sum_{x^k\in \text{Stego}}\omega^k\mu_{R_i}(x^k)}{\sum_{x^k\in \text{Stego}}\omega^k} \tag{6}$$

$$fitness(R_j)=\omega_P f_P-\omega_N f_N \tag{7}$$

其中，f_P 为规则 R_i 所覆盖训练样本的阳性率；f_N 为规则 R_i 所覆盖训练样本的阴性率；w_k 为权重参数，表示样本 x_k 在训练数据库中的频率；w_p 与 w_N 分别是规则阳、阴性的权重。

本文迭代演化方法的步骤如下。

步骤 1(初始化) 基于训练样本的权重生成一个模糊规则的初始化种群。

步骤 2(遗传操作) 通过选择、交叉与变异 3 个遗传操作生成新的模糊规则。

步骤 3(替换操作) 使用新生成的规则替换当前种群的部分成员。

步骤 4(内部循环结束条件) 如果满足内部结束条件，则结束步骤 2－步骤 3，否则返回步骤 2。

步骤 5(外部循环结束条件) 如果满足外部结束条件，则结束步骤 1－步骤 6，否则转至步骤 6。

步骤 6(权重调节) 降低新获得模糊规则的样本的权重，返回步骤 1。

算法每轮迭代的输出是一条模糊规则。

2.4 隐写模式决策

对于一个给定的规则库 S，为了决定特征向量 $x_p=(x_{p1},x_{p2},\cdots,x_{pn})$ 的图像是否含有隐秘信息，使用下式计算参数 τ^j_{stego}：

$$\tau_{\text{Stego}}=\sum_{R_j\in S}\mu_j(x_{ps})CF_j \tag{8}$$

采用式(9)计算图像 I 的隐写模式(DP)：

$$DP(I)=\arg\max(\{\tau^n_{\text{Stego}},\tau_{\text{Clean}}\}),n=1,\cdots,N \tag{9}$$

3 基于 ABC 的隐写分析算法

正常图像的相邻像素间具有一定的关联性，隐写算法则改变了相邻像素间的关联性，本文尝试降低所提取的特征之间的关联性，从而提高隐写分析的准确率。本文提出一个采用 ABC、基于区域的图像隐写分析算法(RAB)，图 3 和图 4 是本算法两个阶段的结构，分别包含了 7 个训练阶段与 6 个测试阶段。

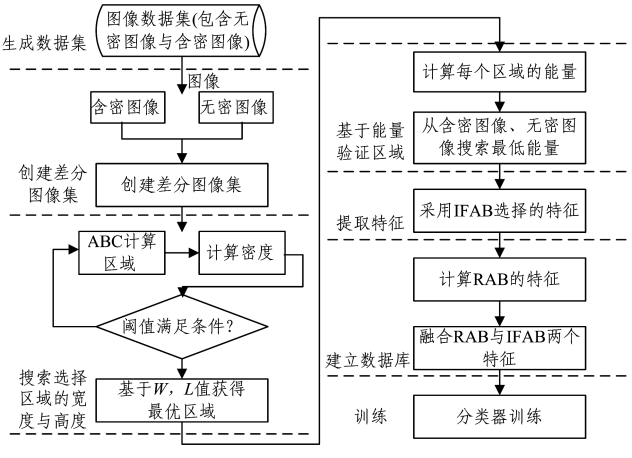


图 3 基于 ABC 的隐写分析算法的训练阶段

Fig. 3 Train phase of ABC based steganalysis algorithm

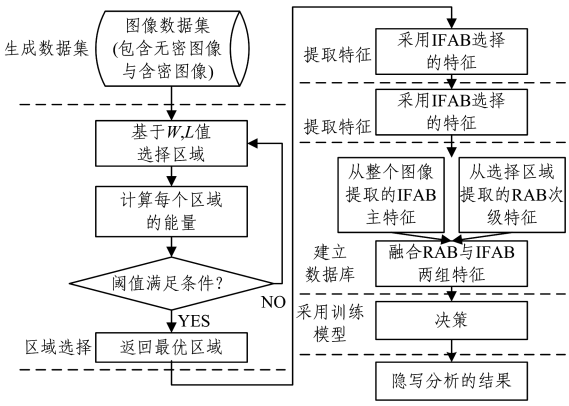


图 4 基于 ABC 的隐写分析算法的测试阶段

Fig. 4 Test phase of ABC based steganalysis algorithm

本算法的目标是检测满足条件的最优区域。首先，寻找图像的采样区域；然后，评估区域的密度和能量，以选择最优的样本(子图像)，通过 IFAB[14] 从原集合中选择最优的样本，从整个图像中提取相同的特征，融合提取的两个特征集合生成一个新的训练数据集；最终，通过学习模型从覆盖图像重构含密图像，此外，通过 RAB 方法提取判别性的特征。

IFAB 采用 ABC 算法进行特征选择，RAB 根据密度评估生成的子图像来进行区域选择。ABC 算法中包含 3 组蜂群：雇佣蜂、观察蜂与侦查蜂。设置雇佣蜂占种群的一半，剩下的蜂群为观察蜂，每个雇佣蜂必须对应一个实物源，即雇佣蜂的数量等价于食物源的数量，放弃食物源的雇佣蜂变为侦查蜂。将图像表面密度作为区域密度，计算方法为：

$$D=\frac{m}{A} \tag{10}$$

其中，D 表示平均区域密度，m 表示子图像的总质量，A 表示子图像的总区域大小。

3.1 训练阶段

3.1.1 生成图像数据集

从图像库收集、生成可靠的覆盖图像与含密图像：采用一个隐写算法对图像进行处理，以获得完整的含密图像与覆盖图像数据库，采用 HUGO 隐写算法[9] 在覆盖图像中嵌入隐秘消息，嵌入率为 40%(比特/像素)。在训练程序和测试程序中，该阶段的处理方式相同。

3.1.2 建立差分图像

向图像嵌入一个消息，会导致图像中相邻像素的关联性发生变化，该阶段尝试识别子图像的长度 L 与宽度 W，通过嵌入程序改变其中的像素点。为了识别嵌入的点，使用式(2)所示的函数获得覆盖图像与含密图像，式(2)有两个输入参数：含密图像与覆盖图像。其计算结果为一个图像：嵌入点为白色点，其余点为黑色点。因为测试阶段并不知道该图像是覆盖图像还是含密图像，所以测试过程会跳过该阶段。

$$EP=\sum_{i=1}^{M}\sum_{j=1}^{N}|S(i,j)-c(i,j)| \tag{11}$$

其中，EP 表示嵌入点，M 与 N 分别表示子图像的长度与宽度，S 与 C 分别表示含密图像与覆盖图像。

3.1.3 基于 ABC 搜索最优区域的长度与宽度

1)子图像选择算法

如图 5 所示，子图像的选择方法包含 4 个部分：生成子图

像，评价函数，ABC 的结束条件，验证子图像。通过 ABC 评估子图像选择程序的总过程，子图像选择程序的关键目标是搜索图像的大嵌入空间，训练阶段基于每个差分子图像搜索长度 L 与宽带 W。

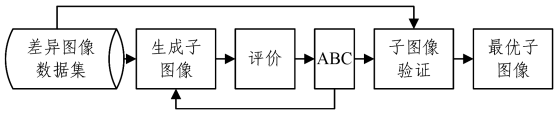


图 5　子图像选择程序的总体流程

Fig. 5　Overall flow of sub-image selection process

2）基于 ABC 的区域选择算法

ABC 算法在基于 ABC 的区域选择方法中对区域选择程序进行优化，算法 1 为区域选择程序的伪代码。伪代码中 M 与 N 分别表示图像的高度与宽度，P_i 表示选择一个解的概率，X_i 表示分配到雇佣蜂的区域密度，X_j 表示选择的观察蜂的区域密度。采用 ABC 寻找最优区域，采用评价函数来评估观察蜂创建的每个区域。

算法 1　基于密度的区域选择算法

```
1. iterate=1;
2. 初始化 ABC 参数;
3. FOREACH 像素 IN 图像 // 像素总量为 M×N
4. 基于密度评估各像素的适应度;
5. WHILE (iterate++ < 最大迭代次数) {
       /* 雇佣蜂 */
6. 为每个雇佣蜂分配一个生成的子图像;
7. 生成新的子图像 Vi;
8. 将生成的子图像传递至基于密度的函数,使用密度评估当前的适应
   度,计算当前解的概率 Pi;
       /* 观察蜂 */
9. 基于概率 pi 选择一个新的食物源;
10. 根据 Xi 与 Xj 计算 Vi;
11. 使用贪婪算法选择 vi 与 xi;
12. 生成一个食物源;
13. 根据密度评估新解的适应度;
14. IF (新解适应度 > 当前解适应度)
15.     替换为新解;
        /* 侦查蜂 */
16. IF (迭代解的适应度未能提高)
17.     放弃当前的食物源,搜索新的食物源;
18.     保存最优的子图像;
19. }
20. 采用上述的蜂群搜索程序生成预定数量的最优子图像;
```

1）种群初始化。本文基于密度的方案搜索像素的搜索空间。随机生成初始化种群，ABC 的参数包括食物源数量、种群大小、下界、上界、约束与最大迭代次数。雇佣蜂与观察蜂的种群大小分别等于覆盖图像与含密图像的维度。

2）计算与评估像素和食物源。该步骤的目标是寻找有价值的子图像（区域）。本算法是一种基于封装的方法，因此评估像素和食物源极为重要，有助于获得最优的子图像。将像素的密度和食物源作为适应度函数的输入参数，评估子图像的质量。

一个特征向量有 4 个维度：子图像的起始点、结束点位置，图 6 所示是食物源的结构。根据图 6，I 与 J 表示子图像的开始点，I' 与 J' 表示子图像的结束点。

I	J	I'	J'
图像长度-像素值	图像宽度-像素值	图像长度	图像宽度

图 6　食物源的结构

Fig. 6　Structure of food source

将各雇佣蜂分配至子图像，并使用式(3)评估每个特征的适应度，然后采用基于密度的函数评价各个候选解。

3）雇佣蜂。人工蜂群算法中每个雇佣蜂生成当前位置的食物源，获得食物源适应度的准确率，食物源中每个像素是候选解。雇佣蜂负责开采食物源，并根据开采的食物源质量与观察蜂分享信息。使用式(12)评估解的适应度，其中 D 表示子图像的密度。

$$fitness(R)=\begin{cases}\dfrac{1}{1+D}, & D<0\\ 1+|D|, & D>0\end{cases} \tag{12}$$

4）观察蜂。观察蜂选择适应度最高的食物源，根据每个食物源的概率(式(13))选择一个食物源，然后仅更新当前的合适解。

$$P(R)=\frac{fitness(R)}{\sum_{s=1}^{m} fitness(s)} \tag{13}$$

其中，m 表示图像的高度，s 表示每个解的适应度，R 表示观察蜂所选择的解的适应度。

观察蜂基于子图像的密度分析解，雇佣蜂指向观察蜂选择的区域。分配合适解 V_i 的方法为：

$$j=rand[1,N],V_i^*=f_i+\theta(f_i-f_j) \tag{14}$$

其中，$i=\{0,1,2,\cdots,N\}$，$j=\{0,1,2,\cdots,N\}$，N 表示像素的上界，f_i 表示雇佣蜂的子图像密度，f_j 表示观察蜂的子图像密度，θ 表示 $[-1,1]$区间的一个随机实数。根据式(5)，假设 V_i 是观察蜂选择的一个新节点密度，若 $V_i<f_i$，则可得 f_j 的密度高于 V_i，此时应当增强食物源的总密度。

每当为一个子图像分配雇佣蜂，观察蜂则生成一个子图像，当探索多个可行的子图像之后，蜂蜜的累积量达到一个临界点，从而选择一个更优的子图像。如果未分配雇佣蜂，则将雇佣蜂变为侦查蜂，然后将侦查蜂分配到式(6)的一个像素空间中。如果候选解的适应度高于当前解，则改变最近解：

$$X_i=X_{\min}+\theta'(X_{\max}-X_{\min}) \tag{15}$$

其中，$X_{\max}$ 和 $X_{\min}$ 分别代表种群数量的上界和下界，θ' 是[0，1]区间的一个随机实数。结束条件为迭代次数达到预设的最大迭代次数或者满足预定的密度增量值。

5）侦查蜂。搜索像素空间中的侦查蜂随机地生成一个食物源。若经过预定的迭代次数后最近食物源的适应度值并未增加，则放弃该食物源，然后，侦查蜂在所有的像素维度中随机生成一个食物源(式(6))，由此防止选择次优解。

6）结束条件。迭代次数达到预定的最大值或者满足结束条件。

7）基于能量的验证。基于上文获得的 L 与 W，根据覆盖图像与含密图像的能量值验证子图像的优劣，从而保证图像

始终具有最优的能量，最优能量是含密图像、覆盖图像之间的能量最小值。图 7 所示是覆盖图像、含密图像与差分图像的实例。

(a)覆盖图像 (b)HUGO 隐写算法获得的含密图像(负载为 0.4) (c)差分图像(矩形框内为所选子图像)

图 7 图像实例

Fig. 7 Image instances

8)使用选择的特征集。本文 IFAB 算法选择的特征可以提高隐写预测的精度。

9)IFAB 的总体结构。图 8 所示是 IFAB 的总体结构，包含 3 个重要的步骤：特征提取器、ABC 特征选择、分类器(支持向量机)。

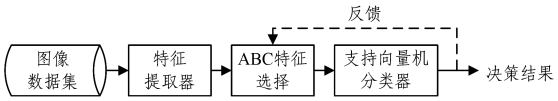


图 8 IFAB 的总体结构

Fig. 8 Overall structure of IFAB

10)基于提取的特征建立数据库。最终将 IFAB 与 RAB 的特征融合成一个数据集。该数据集中特征的维度不同，DRL 特征选择算法中数据集的特征数量为 160。

最终，在各属性的结尾增加一个类属性(名字)以指明每个实例的类型。

3.1.4 学习模型

在训练图像中创建数据集，设置一个分类器学习数据集的模型。

3.2 测试阶段

在测试阶段，评估各覆盖图像以识别出含密图像。

1)生成数据集。采用 HUGO 隐写算法在图像中嵌入消息，嵌入比例为 40%(比特/像素)，该阶段与训练程序相同。

2)区域选择。根据能量值(式(16))、学习的子图像宽度 W 和长度 L 决定最优区域。从每个输入图像中选择 $W\times L$ 的子图像，然后计算每个子图像的能量值，若一个子图像的能量达到能量阈值，则选择该子图像为最优的子图像；否则，选择另一个子图像直至达到阈值。

$$E=\sum_{r=0}^{L-1}P(r)^2,P(r)=\frac{n(r)}{L\times W} \tag{16}$$

其中，$W\times L$ 表示子图像的像素总量，$n(r)$是幅度 r 的像素总量，$L-1$ 是子图像的最大强度。

计算各子图像的密度与能量值之后，总结子图像的含密图像与覆盖图像之间的关系。训练阶段测试了点的密度，可计算出对应子图像的能量，然后选择一个能量与密度值均较高的子图像，该过程的能量值优先级高于密度值。因为测试阶段无法获得密度值，所以根据能量值决定子图像的优劣，图 9 所示是测试图像和选择的子图像的实例。

(a)测试图像 (b)基于能量值选择的子图像

图 9 基于能量选择子图像的实例

Fig. 9 Instance of sub-image selection based on energy

3)特征提取器。利用 IFAB 减少提取的特征量，从每个选择的子图像中尝试提取相同的特征，然后从整个图像中提取相同的特征，该测试阶段与训练程序相同。

4)创建数据集。融合两种特征，建立最终的数据库，该阶段的测试程序与训练程序相同。

5)使用训练模型。使用训练阶段学习的模型预测测试数据集，检查正确预测的含密图像的准确率。使用训练阶段的学习模型，从测试数据集中预测图像实例的类标签。

6)结果。将所有的测试图像标记为含密图像或覆盖图像，然后统计其中的检测率与误检率。

4 仿真实验与结果分析

采用公开的隐写系统(BOSS 系统 V1.01)获得隐写图像，该系统的嵌入率为 0.4(比特/像素)，BOSS 数据库共包含 10000 幅覆盖图像与 10000 幅隐写图像，随机、均匀地从 20000幅图像中选择 900 幅图像作为本文仿真实验的 benchmark 数据集。

将本算法与 DRL(深度残差模型)[15]、TIFS[16]、BS_RSVC[17-18]3 种隐写分析算法进行比较，综合评价本算法的性能。DRL 是空域性能最好的检测方法之一，该方法首先计算空域相邻像素在 8 个方向的像素之差，然后建立相应的马尔可夫链模型，最后根据转移概率矩阵构建一个 686 维的特征向量。CC-PEV 与 BS_RSVC 则是近两年性能较好的两个隐写分析算法。

4.1 参数设置

因为小于 160×160 像素的子图像过小，难以提取其特征，所以将 160 设为子图像的最小尺寸。

表 2 所列为与 IFAB 和 RAB 相关的参数设置。

表 2 本文 ABC 的参数设置

Table 2 Parameters of proposed ABC algorithm

参数	*RAB*	*IFAB*
种群大小 P	2×图像大小	2×特征数量
食物源	$P/2$	$P/2$
像素值 PV	160	*
特征维度 D	4	80
下界	1	1
上界	$N=548-PV$	N
运行次数	20	20
限制	100	100

4.2 性能指标

采用 3 个性能指标评估 4 个隐写分析算法的性能，3 个性能指标分别为：MAE(Mean Absolute Error)，RMSE(Root Mean Squared Error)，RME(Relative Absolute Error)。

MAE 的计算方法为：

$$MAE=\frac{1}{n}\sum_{i=1}^{n}|f_i-y_i|=\frac{1}{n}\sum_{i=1}^{n}|e_i| \tag{17}$$

MAE 表示绝对误差 $e_i=f_i-y_i$ 的均值，其中 f_i 表示预

测值，y_i 表示真实值。

RMSE 的计算方法为：

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}e_i^2} \tag{18}$$

RMSE 测量了预测值与真实值之间的差异。

RME 计算 N 次实验的绝对误差之和，再除以实际值与均值差值的和。

4.3　实验结果与分析

每种隐写分析算法对 BOSS benchmark 数据集独立地进行 20 次实验，取平均值作为最终的实验结果。表 3 所列是 4 种隐写分析算法对 BOSS benchmark 数据集的平均实验结果。总体而言，DRL 算法的性能优于 TIFS 算法，由此可见，DRL 算法依然是空间域性能最优的算法之一。在 4 种隐写分析算法中，本算法的检测性能最优，略高于 BS_RSVC 算法，主要原因在于本算法通过两阶段的隐写分析实现了严格的隐写检测，在第二阶段的隐写分析中，本文融合了图像的区域密度特征与图像的能量特征，综合地分析了图像的隐密信息，具有较好的检测精度。

表 3　4 种隐写分析算法对 BOSS benchmark 数据集的平均实验结果

Table 3　Average experimental results of four steganalysis algorithms applied to BOSS benchmark dataset

性能指标	DRL	TIFS	BS_RSVC	本文算法
MAE	45.11	52.78	38.5	32.23
RMSE	67.16	72.65	61.2	56.37
RAE	90.24	105.55	83.4	65.89

计算效率是隐写分析算法另一个重要的性能指标，表 4 所列是 4 种隐写分析算法的平均计算时间。由于 TIFS 算法需要处理图像集所有的特征，因此其计算效率最低，BS_RSVC 算法则是一种基于奇异值曲线分析的隐写分析算法，该算法无须分析所有的特征，计算效率最高。本文算法则通过人工蜂群算法对特征进行了优化，不仅提高了隐写分析的准确率，而且减少了所分析的像素的数量，提高了计算效率，本文算法的计算时间与 BS_RSVC 算法接近，但本文算法在检测率上的性能优于 BS_RSVC 算法。

表 4　4 种隐写分析算法的计算时间

Table 4　Computational times of four steganalysis algorithms

（单位：s）

隐写分析算法	DRL	TIFS	BS_RSVC	本文算法
评价计算时间	0.66	0.92	0.36	0.39

结束语　文章设计了一种两阶段的隐写分析算法。第一阶段为隐写模式检测阶段，该阶段通过模糊理论检测隐写算法的模式；第二阶段为基于 ABC 的隐写分析阶段，该阶段采用 ABC 对图像的密度与能量特征进行了优化，检测出了图像的含密区域。训练阶段测试了点的密度，可计算出对应子图像的能量，然后选择一个能量与密度值均较高的子图像，该过程的能量值优先级高于密度值。因为在测试阶段无法获得密度值，所以根据能量值判断子图像的优劣。

未来将考虑引入图像其他时空域与稀疏表示的特征，建立基于人工蜂群算法的模型，进一步提高图像隐写分析的准确率。

参 考 文 献

[1]　SI Y F, WEI L X, ZHANG Y N, et al. Revised Steganography Scheme Based on SI-UNIWARD[J]. Computer Science, 2016, 43(5): 108-112.

[2]　SUN X, ZHANG W M, YU N H, et al. Steganography based on parameters' disturbance of spatial image transform[J]. Journal on Communications, 2017, 38(10): 166-174.

[3]　ZHANG Y W, ZHANG W M, YU N H. Specific Testing Sample Steganalysis[J]. Journal of Software, 2018, 29(4): 987-1001.

[4]　ZHANG Y, LIU F, YANG C, et al. Steganalysis of content-adaptive JPEG steganography based on Gauss partial derivative filter bank[J]. Journal of Electronic Imaging, 2017, 26(1): 013011.

[5]　JIAN Y, NI J, YANG Y. Deep Learning Hierarchical Representations for Image Steganalysis[J]. IEEE Transactions on Information Forensics & Security, 2017, 12(11): 2545-2557.

[6]　ZENG J, TAN S, LI B, et al. Large-Scale JPEG Image Steganalysis Using Hybrid Deep-Learning Framework[J]. IEEE Transactions on Information Forensics & Security, 2018, 13(5): 1200-1214.

[7]　KARAMPIDIS K, KAVALLIERATOU E, PAPADOURAKIS G. A review of image steganalysis techniques for digital forensics[J]. Journal of Information Security & Applications, 2018, 40(4): 217-235.

[8]　WANG Y J, NIU K, YANG X Y. Information hiding scheme based on generative adversarial network[J]. Journal of Computer Applications, 2018, 38(10): 2923-2928.

[9]　DUAN R, CHEN D. Video steganography algorithm uses motion vector difference as carrier[J]. Journal of Image and Graphics, 2018, 23(2): 163-173.

[10]　CAO Z, ZHANG M Q, SUN W J, et al. Novel Steganalysis Algorithm Combine Rotating Forest Transformation with Multiple Classifi-ers Ensemble[J]. Journal of Chinese Computer Systems, 2017, 38(10): 2297-2302.

[11]　HAO Z, TAO Z, CHEN H. Revisiting weighted Stego-image Steganalysis for PVD steganography[J]. Multimedia Tools & Applications, 2018, 3(2): 1-19.

[12]　SONG X, LIU F, LUO X, et al. Steganalysis of perturbed quantization steganography based on the enhanced histogram features[J]. Multimedia Tools and Applications, 2015, 74(24): 11045-11071.

[13]　SURYAWANSHI G R, MALI S N. Universal steganalysis using IQM and multiclass discriminator for digital images[C]// International Conference on Signal Processing. 2017.

[14]　WU S, ZHONG S, LIU Y. Deep residual learning for image steganalysis[J]. Multimedia Tools & Applications, 2017, 77(9): 1-17.

[15]　HAO Z, PING X J, MANKUN X U, et al. Steganalysis by subtractive pixel adjacency matrix and dimensionality reduction[J]. Science China Information Sciences, 2014, 57(4): 1-7.

[16]　BOROUMAND M, FRIDRICH J. Applications of Explicit Non-Linear Feature Maps in Steganalysis[J]. IEEE Transactions on Information Forensics & Security, 2018, 13(4): 823-833.

[17]　NOURI R, MANSOURI A. Blind image steganalysis based on reciprocal singular value curve[C]// Iranian Conference on Machine Vision and Image Processing. IEEE, 2015: 124-127.

[18]　CHANG K K, HUO J Y, MEI K. A Gbest-Guided Aritificial Bee Colony Algorithm with Hunting Factor[J]. Journal of Chongqing University of Technology (Natural Science), 2017 (6): 160-165, 187. (in Chinese)
常扣扣，火久元，梅凯. 一种带搜索因子的全局最优人工蜂群算法[J]. 重庆理工大学学报(自然科学版)，2017(6)：160-165，187.