IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 23, 2020, accepted July 7, 2020, date of publication July 10, 2020, date of current version July 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3008473

AI-Enabled Diagnosis of Spontaneous Rupture of Ovarian Endometriomas: A PSO Enhanced Random Forest Approach

MINGYAN ZHOU[1,*], FENG LIN[2,*], QIAN HU[1], ZHENZHOU TANG[1], (Senior Member, IEEE), AND CHU JIN[3]

[1]College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou 325035, China
[2]Department of Gynecology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou 325015, China
[3]Renji College, Wenzhou Medical University, Wenzhou 325015, China

Corresponding authors: Zhenzhou Tang (mr.tangzz@gmail.com) and Chu Jin (634137089@qq.com)

*M. Zhou and F. Lin contributed equally to this work.

This work was supported in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LZ20F010008 and Grant LQ18H280007, in part by the Natural Science Foundation of China under Grant 81803781, and in part by the Fundamental Scientific Research Project of Wenzhou City under Grant G20180008 and Grant G20190021

ABSTRACT Spontaneous rupture of ovarian endometriomas (OEs) may cause serious injury to patients. However, in traditional clinical diagnosis, it is vulnerable to be ignored or misdiagnosed for the symptoms of the acute abdomen caused by it are quite similar to those of some common gynecological emergencies, which leads to serious complications to patients. In view of this, this study investigates the AI-enabled, early and accurate diagnosis of spontaneous rupture of OEs. Although artificial intelligence (AI) has been proved to be a powerful tool to help to make accurate clinical diagnosis, however, as far as we know, there is so far no report on AI-enabled diagnosis of spontaneous rupture of OEs yet. Specifically, this study proposes a particle swarm optimization (PSO) enhanced random forest (RF) classification model, called PSO-RF, to make diagnosis, where RF is used to rank feature importance and make diagnosis considering that an OE is ruptured or not is a typical 0-1 classification problem, and PSO is leveraged to fine-tune the essential parameters of RF. The performance of the proposed PSO-RF model is evaluated with practical data collected from a local hospital and fully benchmarked by comparing with eight other machine learning models whose key parameters are sufficiently optimized as well by grid search or PSO, for the sake of fairness. The experiment results show that the proposed PSO-RF model outperforms all the other models, with the accuracy of 97.47%, the area under the ROC curve (AUC) of 0.996, the sensitivity of 94.12% and the specificity of 98.39%. It can be concluded that the PSO-RF model is a highly effective AI-enabled tool for preoperatively diagnosing spontaneous rupture of OEs.

INDEX TERMS Ovarian endometrioma, spontaneous rupture, random forest, particle swarm optimization, machine learning.

I. INTRODUCTION

Ovarian endometrioma is a kind of endometriosis, patients of which have ectopic endometrial tissues in ovary [1] and suffer from various clinical symptoms including pelvic pain, irregular menstrual bleeding, dysmenorrheal, dyspareunia and infertility [2]. It is one of the most common diseases in reproductive women aging from 25 to 45 years old, with a high incidence of 10%-15% [3]. Although ovarian endometrioma is a common benign disease, the size of the cyst typically increases gradually, and with a small probability, spontaneous rupture of cyst may occur during or after menstruation [4], [5]. If the rupture can not be diagnosed and treated correctly in time, the severe cases may cause massive hemorrhage in the abdominal cavity, leading to hemorrhagic shock [6]. However, the symptoms of the acute abdomen caused by spontaneous rupture of ovarian endometrioma are similar to those of ectopic pregnancy, rupture and hemorrhage of corpus luteum, torsion of ovarian cyst pedicle and appendicitis [7], [8]. Consequently, it is vulnerable to be ignored or misdiagnosed clinically which causes serious complications to patients. Early and accurate diagnosis the spontaneous rupture of ovarian endometrioma could significantly help to avoid serious injury.

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

Considering the high misdiagnosis of spontaneous rupture of ovarian endometrimas, researches on it has been widely carried out. Some of them have tried to find suitable biomarkers to assist diagnosis of rupture of ovarian endometrimas preoperatively and have reported that spontaneous rupture of ovarian endometrimas are typically accompanied by high levels of serum CA125 and CA19-9 [1], [7], [9]–[12]. In particular, Xinyu Dai *et al.* have evaluated the clinical significance of serum CA125 and CA19-9 in the patients with spontaneous rupture of ovarian endometromas by statistical methods [2]. The results have shown that the levels of CA125 and CA19-9 in patients with spontaneous rupture of ovarian endometromas increased significantly. Currently, it is believed that the combined biomarkers of CA125 and CA19-9 is helpful for the early diagnosis of spontaneous rupture of ovarian endometriomas.

In the other hand, artificial intelligence (AI), especially machine learning (ML), has been widely used in medical diagnosis [13]. The innovative AI-enabled methods are important assistances in precision medicine, which may help to make accurate clinical diagnosis at low cost and early. It has been frequently reported recently that the innovative AI-enabled approaches are important tools in precision medicine, which may help to make accurate clinical diagnosis at low cost and early.

By leveraging various machine learning methods, Kawakami *et al.* have developed an ovarian cancer-specific predictive framework for clinical stage, histotype, residual tumor burden, and prognosis based on multiple biomarkers [14]. The results have revealed that AI-enabled models can provide accurate diagnoses and prognostic predictions for patients with epithelial ovarian cancer prior to initial intervention. In [15], an optimized random forest (RF) classifier has been used to detect nodules inside the lungs. In [16], Y. Shi et. al have proposed a hybrid QPSO-RF (Quantum particle swarm optimization - RF) model to predict the disease progression of Idiopathic Pulmonary Fibrosis using Computed Tomography (CT). Specifically, QPSO algorithm has been used to search the optimal feature subsets and RF has been used for classification. In [17], four machine learning algorithms, namely support vector machine (SVM), partial least squares discriminant analysis (PLS-DA), RF, and logistic regression (LR), have been applied to develop the model for identifying metabolomic biomarkers in cervicovaginal fluid for EC detection. Pergialiotis *et al.* have investigated the diagnostic accuracy of three different machine learning algorithms, i.e. LR, artificial neural networks and classification and regression trees (CARTs) for the prediction of endometrial cancer in postmenopausal women with vaginal bleeding or endometrial thickness $\geq$ 5 mm, as determined by ultrasound examination [18]. E. Pashaei *et al.* have proposed a binary version of black hole algorithm for solving feature selection problem in biological data [19]. A novel one-dimensional deep densely connected neural network (DDNN) has been constructed in [20] to detect atrial fibrillation in 12-lead electrocardiogram waveforms. The proposed model has been proved to be sufficiently effective to be applied to clinical diagnosis of atrial fibrillation. In the study of Obrzut *et al.*, the usefulness of six AI-enabled methods have been evaluated for 5-year overall survival prediction in patients with cervical cancer treated by radical hysterectomy [21]. It has demonstrated that the best model, namely the probabilistic neural network model, achieves a high prediction accuracy of 0.892 and sensitivity of 0.975.

Machine learning models have also been used in the diagnosis of endometrial tumor related diseases. In [22], decision trees have been used to analyze the effectiveness of treatment of patients with recurrent pelvic cyst who underwent surgical intervention. The study of Günakan has constructed a naive Bayes model to make predictions of lymph node involvement in endometrial cancer [23]. In [24], automated image analysis and RF models have been used to classify normal, premalignant, and malignant endometrial tissues. However, as far as we known, no study on ML-enabled diagnostics of spontaneous rupture of ovarian endometriomas has been reported so far.

Inspired by the above observations, this paper proposes a machine learning model to make accurate diagnosis of spontaneous rupture of ovarian endometriomas preoperatively. Specifically, we first model the problem of cyst rupture as a 0-1 classification problem. Practical physiological data involving plenty of features are collected from women with ovarian endometriomas who have treated in the First Affiliated Hospital of Wenzhou Medical University between 2006 and 2017. Considering the effectiveness of random forest algorithm [25] in feature selection and classification and therefore it has been widely used to make intelligent medical diagnoses [15]–[17], [26], [27], we design a random forest model to rank the importance of different features and further solve the 0-1 classification problem. In order to improve the performance of random forest model, we further leverage particle swarm optimization (PSO) algorithm [28] to optimize the parameters of the random forest model. To the best of our knowledge, this is the first time to adopt the machine learning technique in the diagnosis of spontaneous rupture of ovarian endometriomas.

The proposed PSO enhanced RF model, PSO-RF for short, is comprehensively benchmarked by comparing with other fine-tuned machine learning models, including the gridsearch optimized random forest model (GO-RF), the naive Bayesian classification model [29] (NBC), the gridsearch optimized k-nearest neighbor [30] model (GO-KNN), the PSO enhanced KNN (PSO-KNN), the gridsearch optimized lightGBM [31] model (GO-lightGBM), the PSO enhanced lightGBM model (PSO-lightGBM), the gridsearch optimized logical regression [32] model (GO-LR), and the PSO enhanced logical regression model (PSO-LR). The results show that the PSO-RF model outperforms all the other models, with the accuracy of 95.57%, the area under the ROC curve (AUC) of 0.996, the sensitivity of 94.12% and the specificity of 98.39%.

TABLE 1. Baseline characteristics of endometriosis patients.

Features	Overall		Training set		Test set		P value
	Ruptured 53 patients	Unruptured 140 patients	Ruptured 32 patients	Unruptured 82 patients	Ruptured 21 patients	Unruptured 58 patients	
Age	32.09±7.11	33.32±6.19	32.22±7.1	32.98±5.9	31.9±7.29	33.81±6.6	0.57
FIGO Stage							0.84
I	6	12	4	6	3	9	
II	6	24	2	16	2	8	
III	12	54	5	34	3	23	
IV	29	50	21	26	9	22	
Maximum cyst diameter	8.14±3.29	6.82±2.32	7.9±2.89	6.74±2.4	8.53±3.9	6.94±2.24	0.51
<4 cm	48	93	30	55	18	38	
≥4 cm	1	2	0	1	1	1	
CA19-9	318.21±450.13	42.56±77.33	77.25±237.72	137.39±341.38	270.21±283.39	58.81±112.15	0.89
CA-125	450.38±549.83	64.53±48.48	77.84±57.6	182.73±352.38	567.35±599.87	60.59±45.65	0.4
Leukocyte	9.22±4.56	6.26±1.72	7.04±2.97	7.03±3.34	8.89±3.46	6.22±1.7	0.71
Alpha-fetoprotein	1.77±0.81	1.08±0.67	1.33±0.85	1.27±0.71	1.83±0.96	1.07±0.68	0.93
Carcinoembryonic antigen	1.34±0.92	2.44±1	2.38±1.12	2.05±1.15	1.32±0.81	2.41±0.93	0.95
Cyst location							0.61
Left side	8	25	3	18	5	7	
Right side	20	34	14	18	6	16	
Both sides	21	36	13	20	8	16	

The rest of this paper is organized as follows. The methodology of the proposed PSO-RF is given in detail in Section II. Section III presents the experimental verification and analysis. Finally, Section IV concludes the paper and proposes the future work.

II. METHODS

A. DATA ACQUISITION AND PREPROCESSING

Physiological data were collected by complete blood counts, laparoscopic surgery and laparotomy from premenopausal female patients with ovarian endometriomas. The dataset includes 193 records where 53 patients have been diagnosed with spontaneous ruptured ovarian endometriomas and 140 patients with unruptured ovarian endometriomas. Each sample in the dataset includes the age, the leukocyte count and the physiological features obtained by laparoscopic surgery or laparotomy, such as the carcinoembryonic antigen (CEA) level, the CA125 level, the alpha fetoprotein (AFP) level, the CA19-9 level, the position of the ovarian endometrioma, the endometrioma stage[1] and the size of the ovarian endometrioma, as shown in Table 1. All the data have been normalized by $x^* = \frac{x-\min}{\max-\min}$.

The dataset has been divided into the training set and the test set by repeated random sampling until the P values with respect to all the features between the two sets are greater than 0.2. This results in a division of 114 patients in the training set (32 patients with ruptured ovarian endometriomas and 82 patients with unruptured ovarian endometriomas in the training set) and 79 patients in the test set (21 patients with ruptured ovarian endometriomas and 58 patients with unruptured ovarian endometriomas), as listed in Table 1.

[1]These patients were classified into four stages according to the revised classification made by the American Society for Reproductive Medicine (ASRM).

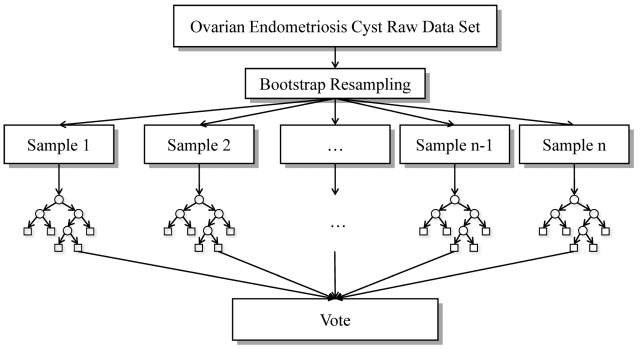


FIGURE 1. Structure of a RF model.

B. RANDOM FOREST MODEL

Random forest (RF) is an ensemble learning model for classification and regression. RF runs efficiently on large data bases, seldom overfits, and has the ability of giving estimates of which features are important in the classification. In particular, RF is able to achieve a good accuracy rate for the classification of discrete high-dimensional datasets. With this in mind, in this work, RF is leveraged to diagnose the spontaneous rupture of ovarian endometriomas.

The RF model consists of multiple classification and regression trees (CARTs) [33]. Each CART is a binary decision tree that is grown by recursively partitioning the data in a parent node into child nodes. The structure of the RF model is shown in Fig. 1.

1) SAMPLING AND SPLITTING

The overall samples in the training set are further divided into two parts. Specifically, 60% of the samples are used to train the CARTs, the set of which is denoted as $\mathbb{D} = \{\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{|\mathbb{D}|}\}$ where $\mathbf{s}_i, i = 1, 2, \ldots, |\mathbb{D}|$ is the ith

multi-dimensional sample in $\mathbb{D}$. The rest are the pretest samples used to measure the classification accuracy rate of this CARTs, the set of which is denoted as $\hat{\mathbb{D}} = \{\hat{\mathbf{s}}_1, \hat{\mathbf{s}}_2, \ldots, \hat{\mathbf{s}}_{|\hat{\mathbb{D}}|}\}$ where $\hat{\mathbf{s}}_i, i = 1, 2, \ldots, |\hat{\mathbb{D}}|$ is the ith multi-dimensional sample in $\hat{\mathbb{D}}$.

Each CART is grown with a bootstrap sample set which is taken by randomly sampling with replacement from $\mathbb{D}$. Concretely, one sample is randomly drawn from $\mathbb{D}$ as a training sample at a time, and then, this sample is put back. This process of sampling with replacement is repeated m times so that to generate a new sample set with the same size as $\mathbb{D}$. The new created sample set is then used as the root, e.g. the training set of a CART, denoted as $\mathbb{R}_i, i = 1, \ldots, N_T$, where N_T is the number of CATRs in RF. Apparently, in this way, some samples may appear multiple times in the new training set, while others, namely about 36.8%[2] of the samples in the original dataset, may never appear. Therefore, the samples that do not appear in the new dataset are taken as the test set.

The training set of each CART will be split as a binary tree. Let N_F be the number of features of each sample. At each node of the decision tree, randomly select $N_f, N_f \ll N_F$ features for splitting. In the other words, at each node, the splitting feature is selected among a random subset of the original set of features. This randomness further enhances the generalization of RF model.

The criteria of splitting in this work is the Gini index of impurity. Formally, denote the sample set at node t as $\mathbb{S}_t = \{\mathbf{s}_{t,1}, \ldots, \mathbf{s}_{t,|\mathbb{S}_t|}\}$ where $\mathbf{s}_{t,i} = \{f_{t,i,1}, \ldots, f_{t,i,N_F}\}$ is the ith sample in node t and $f_{t,i,j}$ is the value of the jth feature in $\mathbf{s}_{t,i}$. The class assignment of $\mathbf{s}_{t,i}$ is denoted as $C_{t,i} \in \{-1, 1\}$ where -1 indicates an unruptured ovarian endometrioma and 1 indicates a ruptured ovarian endometrioma. The Gini index at node t is given by

$$\mathbb{G}_t = \sum_{k=-1,1} p_{k|t}(1 - p_{k|t}), \tag{1}$$

where $p_{k|t}$ is the probability of class k estimated from the samples in node t, e.g.

$$p_{k|t} = \frac{1}{|\mathbb{S}_t|} \sum_{\mathbf{s}_{t,i} \in \mathbb{S}_t} C_{t,i} = k. \tag{2}$$

At each step in the algorithm, the node t is split at the feature j and value $f_{t,j}$ into a pair of child nodes, denoted as $t_L = \{x \in t | x_j \leq f_{t,j}\}$ and $t_R = \{x \in t | x_j > f_{t,j}\}$, respectively, where j and $f_{t,j}$ are determined by

$$\arg\max_{j, f_{t,j}} \quad \Delta\mathbb{G}_t|_{j,f_{t,j}}, \tag{3}$$

where

$$\Delta\mathbb{G}_t|_{j,f_{t,j}} = \left(\mathbb{G}_t - p_{L|t}\mathbb{G}_{L|t} - p_{R|t}\mathbb{G}_{R|t}\right)\Big|_{j,f_{t,j}}, \tag{4}$$

[2]Assuming there are m samples in the original data set, the probability of each sample being selected is $1/m$, and the probability of a sample not being selected in all of the m sampling is $(1 - 1/m)^m$. Moreover, $\lim_{m\to\infty}(1 - 1/m)^m = e^{-1} \approx 0.368$

$p_{L|t}$ and $p_{R|t}$ are the probabilities of assigning a sample to the left child and right child, respectively.

The process of splitting repeats until the terminal condition is satisfied, that is, 1) the depth of the CART reaches the predefined maximum value $d_{\max}$, or 2) $|\mathbb{S}_t|$ is less than the predefined minimum value $N_{s,\min}$. In this way, a CART is formed. The set of the CARTs in RF is denoted as $\mathbb{T} = \{T_1, \ldots, T_{N_T}\}$.

2) VOTING

Finally, all the CARTs involved in a RF cast votes for their inputs, after which RF aggregates their results and determines the output by majority voting of the CARTs. Taking in consideration the different classification abilities of CARTs in RF model, in this work, the voting weight of each CART is set separately according to its classification accuracy. In essence, after the training, the weight of each CART is estimated by the pretest samples in $\hat{\mathbb{D}}$ as follows:

$$w_{T_i} = \frac{N_i^{\text{correct}}}{\left|\hat{\mathbb{D}}\right|}, i = 1, 2, \ldots, N_T, \tag{5}$$

where N_i^{correct} is the number of samples classified correctly by T_i.

The voting strategy is to summarize the classification results of all the CARTs, and then take the weighted mode of the classifications as the final classification results as follows:

$$C_{\mathbf{s}_i} = \arg\max_{k=-1,1} \sum_{T_i \in \mathbb{T}} w_{T_i} \cdot \left(C_{\mathbf{s}_i} \,|\, _{T_i} = k\right). \tag{6}$$

where $C_{\mathbf{s}_i}|_{T_i}$ is the classification result of feature $\mathbf{s}_i$ made by CART T_i.

In summary, the procedure of random forest model is shown in Algorithm 1.

3) FEATURE IMPORTANCE

One of the most attractive advantages of RF is that it gives estimates of what features are important in the classification. There are many biomarkers and factors associated with the rupture of ovarian endometriomas among which some may have very weak correlation with it. So, it will be of great significance to rank the importance of all the features to eliminate irrelevant or weak correlated features, so as to reduce the complexity of RF model and improve the accuracy of the model.

RF uses Gini index to evaluate the importance of a feature. Specifically, the importance of feature j to node t of T_i, denoted as $V_{j,t,i}^{im}$, is evaluated as follows.

$$V_{j,t,i}^{im} = \max_{f_{t,j}} \quad \Delta\mathbb{G}_{j,t,i}|_{f_{t,j}}. \tag{7}$$

Let $\mathbb{M}_{j,i}$ be the set of nodes where feature j is included in T_i. The importance of feature j to T_i can be calculated as

$$V_{j,i}^{im} = \sum_{t \in \mathbb{M}_{j,i}} V_{j,t,i}^{im}. \tag{8}$$

Further, the importance of feature j to RF model is

$$V_j^{im} = \sum_{T_i \in \mathbb{T}} V_{j,i}^{im}. \tag{9}$$

Algorithm 1 Procedure of Random Forest Model

Initialize $N_T, N_F, N_f, d_{\max}, N_{s,\min}$
Preprocess the data set and obtain $\mathbb{D}$ and $\hat{\mathbb{D}}$
for i in $[1, N_T]$ **do**
  for j in $[1, |\mathbb{D}|]$ **do**
    Randomly draw $\mathbf{s}_j$ from $\mathbb{D}$ with replacement.
    Insert $\mathbf{s}_j$ to $\mathbb{R}_i$
  end for
  $t \leftarrow \mathbb{R}_i$
  SPLIT($\mathbb{S}_t, N_f$)
end for
Aggregates the results by Eq. (6).

function split($\mathbb{S}_t, N_f$)
  if Termination condition is satisfied **then return**
  else
    Randomly select N_f out of N_F features at t.
    Determine the feature among N_f features and the value for splitting by Eq. (3) and Eq. (4).
    Split the tagged node into t_L and t_R.
    SPLIT($\mathbb{S}_{t_L}, N_f$)
    SPLIT($\mathbb{S}_{t_R}, N_f$)
  end if
end function

Finally, V_j^{im} is normalized as follows:

$$\tilde{V}_j^{im} = \frac{V_j^{im}}{\sum_{f_i \in \mathbb{F}} V_i^{im}}, \tag{10}$$

where $\mathbb{F}$ is the set of features.

C. PARAMETER OPTIMIZATION BASED ON PSO

Generally, in RF, some parameters such as the maximum depth of a CART ($d_{\max}$), the minimum number of samples in a node ($N_{s,\min}$) and the number of CARTs (N_T), etc., need to be tune for achieving the optimal performance of RF. A feasible solution is to fine-tune these parameters by an appropriate optimization algorithm, such as various heuristic algorithms.

In this study, we leverage particle swarm optimization [34], which is a metaheuristic algorithm simulating the behavior of birds catching food, to tune the essential parameters of the RF model to further improve the accuracy. The essence of PSO is to use the current position, global extremum and individual extremum information to guide the next iteration position of particles, which enables it to approach the optimal solution with a fast convergence speed, and hence effectively optimize the parameters of the model [16], [35]. Another significant advantage of PSO is that it can adjust the maximum step size at each iteration, making it possible to find an approximate optimal solution in a wide range of possible parameters [36]. In addition, PSO has some attractive features, such as easy to implement, less parameters, computationally cheap and so on.

PSO searches the optimal solution through agents, or called particles. A large number of particles form a swarm. A particle i is characterized by its location vector L_i and velocity vector v_i which are updated iteratively by

$$L_i(t+1) = L_i(t) + v_i(t+1), \tag{11}$$

and

$$v_i(t+1) = w_i v_i(t) + c_1 r_1 \left(L_{i,best} - L_i(t)\right) + c_2 r_2 \left(L_{gbest} - L_i(t)\right), \tag{12}$$

where w_i is the selected weighting factor for particle i, $L_{i,best}$ is the location at which particle i previously had the best fitness measure, L_{gbest} is the global optimal location of the whole swarm, c_1 and c_2 are the cognitive acceleration constant and the social acceleration constant, respectively, generally with the value of 2, and r_1 and r_2 are two random parameters within the range [0, 1]. The iteration terminates if 1) reach the maximum number of iterations, or 2) convergence is reached based on the fitness measure.

D. PSO-RF

In summary, the procedure of the proposed PSO-RF is as follows:

- S1 Determine the parameters of PSO, including the maximum number of iterations, the maximum speed of particles and the search space, i.e., the feasible region of solutions.
- S2 Randomly generate a swarm where each particle is a three dimensional vector $\{d_{\max}, N_{s,\min}, N_T\}$. The size of the swarm is denoted as N_{PSO}.
- S3 Start PSO iteration procedure to obtain the optimal $d_{\max}$, $N_{s,\min}$ and N_T.
  - S3.1 For each particle in the swarm, perform the RF algorithm defined in Algorithm 1 to evaluate the fitness, i.e., the diagnostic accuracy.
  - S3.2 For each particle $i, i \in \{1, 2, \ldots, N_{\text{PSO}}\}$ in the swarm, compare the particle's current fitness with the fitness of $L_{i,best}$. If the current location fits better, then replace $L_{i,best}$ with the current location.
  - S3.3 Update L_{gbest} if the new global optimal location of the current swarm fits better.
  - S3.4 For each particle $i, i \in \{1, 2, \ldots, N_{\text{PSO}}\}$ in the swarm, update v_i and L_i by Eq.(12) and Eq.(11).
  - S3.5 Repeat S3.1 - S3.4 until convergence is reached.
- S4 Perform the RF algorithm defined in Algorithm 1 with the optimal $d_{\max}$, $N_{s,\min}$ and N_T obtained in Step S3 to make diagnosis.

The Pseudo code and the flow diagram of the PSO-RF procedure are shown in Algorithm 2 and Fig. 2, respectively.

III. RESULTS

All the involved algorithms and models have been implemented using Python 3.8. The computational analyses have been conducted on a Server running Windows Server 2016 Standard 64 bits operating system with Intel (R) Xeon (R) CPU E5-2620 v4, 32GB of RAM and Nvidia GeForce GTX 1080 Ti graphics card. Unless otherwise stated,

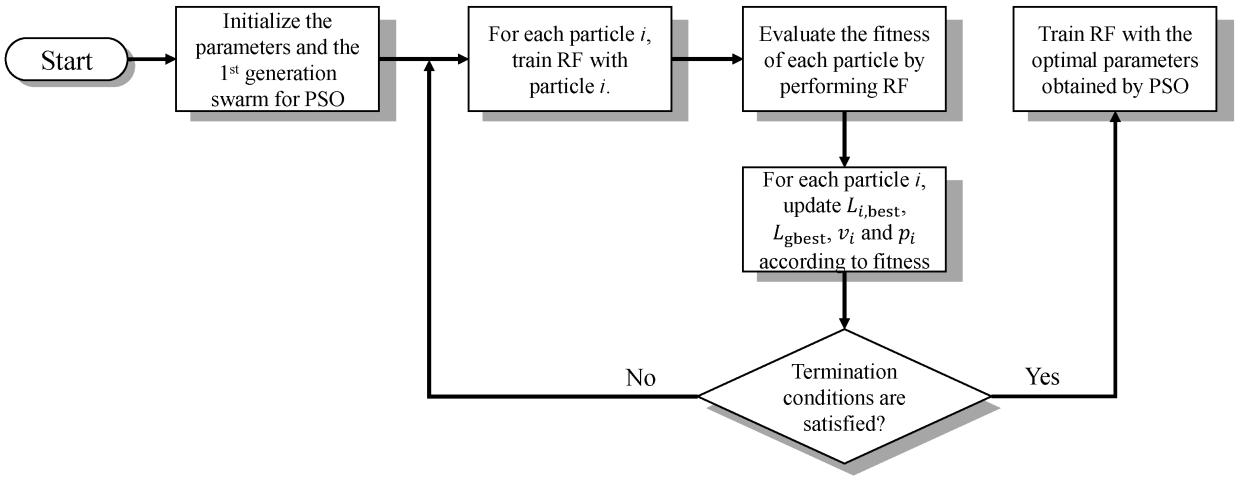


FIGURE 2. PSO-RF flowchart.

Algorithm 2 Procedure of PSO-RF

1: Initialize the parameters for PSO.
2: **for each** particle $i, i \in \{1, 2, \ldots, N_{\text{PSO}}\}$ **do**
3: Randomly initiate v_i and L_i.
4: **end for**
5: **repeat**
6: **for each** particle $i, i \in \{1, 2, \ldots, N_{\text{PSO}}\}$ **do**
7: Perform Algorithm 1 to evaluate the fitness.
8: **if** $fitness(particle\ i) > fitness(L_{i,best})$ **then**
9: $L_{i,best} \leftarrow L_i$
10: **end if**
11: **if** $fitness(particle\ i) > fitness(L_{gbest})$ **then**
12: $L_{gbest} \leftarrow L_i$
13: **end if**
14: **end for**
15: **until** Termination conditions are satisfied.
16: Perform Algorithm 1 with L_{gbest}.

the parameter settings of all the involved algorithms and models are listed in Table 2.

A. PERFORMANCE INDICATORS

1) ACCURACY

Diagnosis of spontaneous rupture of ovarian endometrioma is a typical binary classification problem, in which the diagnosis results are labeled either as positive (P), indicating a ruptured ovarian endometrioma, or negative (N), indicating an unruptured ovarian endometrioma. There are four possible combinations of the diagnosis results and the actuality, namely TP, FP, TN and FN which are explained as follows:

- TP: both the diagnosis result and the actuality are P.
- FP: the diagnosis result is p and the actuality is N.
- TN: both the diagnosis result and the actuality are N.
- FN: the diagnosis result is n and the actuality is P.

TABLE 2. Parameters and settings of algorithms.

Algorithms	Parameters	Settings†
PSO	Maximum velocity	0.5
	Size of swarm	20
	Number of variables	3
	Number of iterations	500
	Social learning factors	2
	Individual learning factors	2
RF	Number of features for splitting	sqrt
	Max depth of a CART*	2
	Min number of samples in a node*	6
	Number of CARTs*	94
KNN	Number of neighbors*	7
NBC	N/A.	
LR	Regularization coefficient*	0.4062
	Penalty	l1
LightGBM	Metric	MSE
	Objective	Regression
	Boosting type	GBDT
	Learning rate*	0.1
	Frequent of bagging	10
	Number of iterations	200
	Number of leaves*	9
	Early stopping rounds	10
	Samples of sampling rate*	0.9
	Feature selection of proportion*	0.8

† The settings for classifiers are obtained by grid search.
* These parameters are the optimization objects of PSO.

The classification accuracy rate, or called the diagnostic accuracy rate, of the proposed PSO-RF model is evaluated by

$$\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{13}$$

2) SENSITIVITY & SPECIFICITY

Besides accuracy rate, sensitivity and specificity are also used to measure the effectiveness of PSO-RF model. Sensitivity, also known as true-positive rate (TPR), is used to measure the ability of correctly identifying the cases with ruptured ovarian endometriomas, whereas specificity, also known as true-negative rate (TNR), is used to measure the ability of the identifying those without the disease. The higher the sensitivity or the specificity is, the more effective the model is. Sensitivity (TPR) and specificity (TNR) can be calculated by

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} = \frac{\mathrm{TP}}{\mathrm{P}}, \tag{14}$$

and

$$\mathrm{TNR} = \frac{\mathrm{TN}}{\mathrm{FP} + \mathrm{TN}} = \frac{\mathrm{TN}}{\mathrm{N}}, \tag{15}$$

respectively.

3) ROC AND AUC

Moreover, the receiver operating characteristic curve (ROC) is used to illustrate the diagnostic ability of the PSO-RF. ROC is a comprehensive indicator that is capable of evaluating TPR and false-positive rate (FPR) at various thresholds. It is a two-dimensional graph in which TPR is plotted on the Y axis and FPR is plotted on the X axis. FPR can be calculated by

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}} = \frac{\mathrm{FP}}{\mathrm{N}} = 1 - \mathrm{TNR}. \tag{16}$$

The AUC (area under the curve) is a concomitant of ROC obtained by calculating the area under the ROC. AUC tells how much the model is capable of distinguishing between ruptured and unruptured ovarian endometriomas. Higher the AUC, better the model is at predicting positives as positives and negatives as negatives.

B. RESULTS

1) FEATURE SELECTION

The ranking of the feature importance for diagnosing the rupture of ovarian endometriomas derived by PSO-RF is shown in Fig. 3. The P values and the Pearson correlation coefficients (PCC) for the features related to the rupture of OE are listed in Table 3.

It can be observed the ranking obtained by PSO-RF is on the whole consistent with those obtained by statistical analyses. The level of CA19-9 and CA125 are the top two important biomarkers, with the normalized importance of 0.316 and 0.273, respectively, the P values of 4.08×10^{-11} and 2.47×10^{-14}, respectively, and the PPC values of 0.4522 and 0.5128, respectively. This conclusion is consistent with the clinical experience. Serum CA19-9 and CA125, tumor associated antigens, have been used clinically to identify the recurrence and severity of endometriosis. The endometrioma cyst fluid was suspected to be rich of biomarkers such as CA19-9 and CA125. However, in normal conditions, the large CA19-9 and CA125 glycoprotein molecules were preventing by the thick wall of endometrioma cyst from entering the peripheral circulation. The spontaneous rupture of the OE may lead the biomarkers into the blood circulation, resulting the high elevation of the CA19-9 and CA125 in the serum [11]. The level of CEA, with the normalized importance of 0.231, P value of 7.35×10^{-11}, and PCC value of -0.4468, is the third important biomarker. The leukocytes count and AFP level are also significant biomarkers, however, are less important than CA19-9, CA125 and CEA.

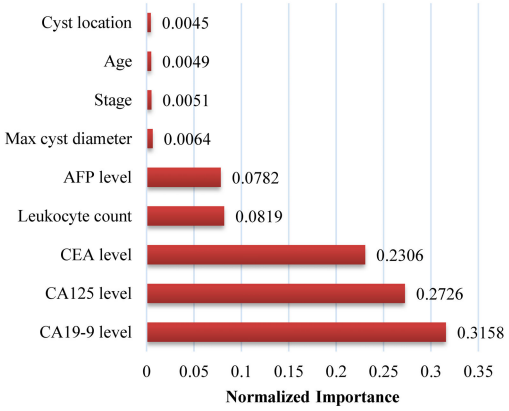
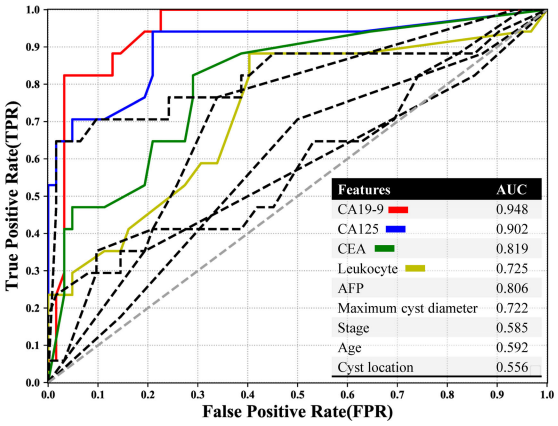


FIGURE 3. Ranking of the importance of features of ovarian endometriomas.

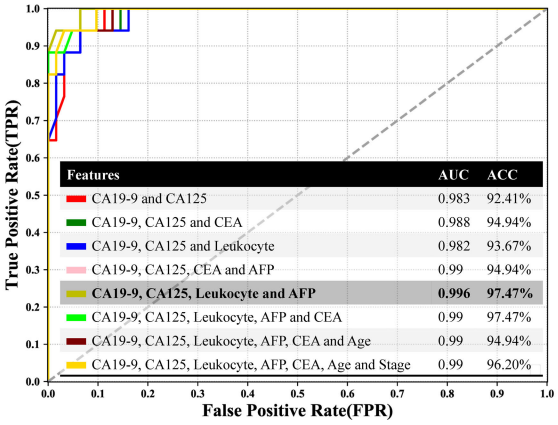
TABLE 3. Association of ruptured/unruptured OEs with clinicopathologic features.

Features	P value	PCC
Age	0.2398	-0.0850
Stage	0.2169	0.0893
Location	0.1250	0.1108
Maximum diameter	1.74×10^{-03}	0.2240
Leukocyte	2.07×10^{-08}	0.3900
AFP Level	1.23×10^{-08}	0.3957
CEA Level	7.35×10^{-11}	-0.4468
CA19-9	4.08×10^{-11}	0.4522
CA125	2.47×10^{-14}	0.5128

In order to investigate the optimal combinations of features to be input into the PSO-RF model, the effectiveness of any single feature and various feature combinations on diagnosis of ruptured ovarian endometriomas is evaluated. Specifically, Fig. 4 plots the ROC curves and the corresponding AUC values of the classification results derived from PSO-RF using a single feature. Not surprisingly, CA19-9, CA125 and CEA output the most effective results, with the values of AUC of 0.948, 0.902 and 0.819, respectively. In addition, although the importance of the Leukocyte level is a bit higher than that of the AFP level, the AUC value of AFP level (0.806) is much greater than that of Leukocyte level (0.725). The ROC Curves, AUC values and accuracy rates of classifications using various features combinations derived from PSO-RF are depicted in Fig. 4(b). It can be seen that the best performance is achieved in the case of inputting

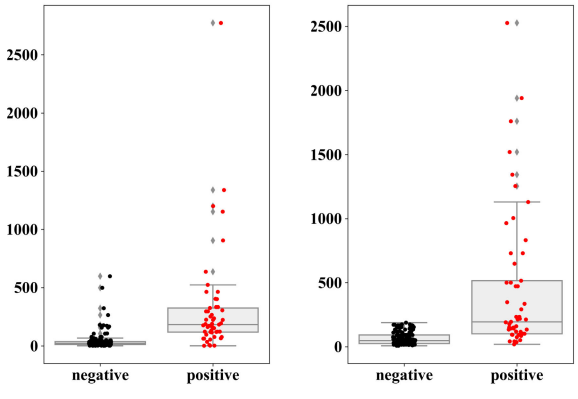


(a)

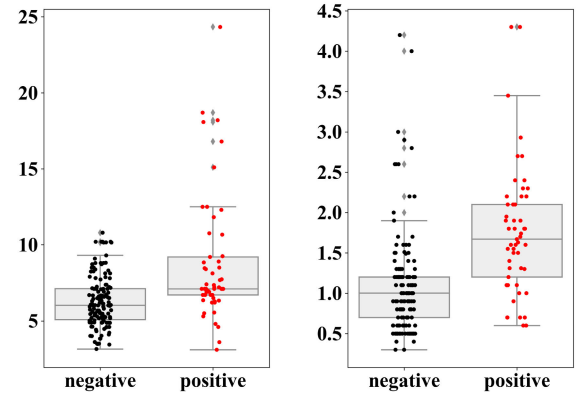


(b)

FIGURE 4. (a) ROCs Curves and values of AUC of all the features derived from PSO-RF for segregating ruptured ovarian endometriomas from ovarian endometriomas, (b) ROCs Curves, values of AUC and accuracy of various features combinations derived from PSO-RF for segregating ruptured ovarian endometriomas from ovarian endometriomas.



(a) CA19-9 (b) CA125



(c) Leukocyte (d) AFP

FIGURE 5. Box and jitter plots representing distribution of the selected biomarkers between the ruptured and unruptured ovarian endometriomas.

the following four features: CA19-9 level, CA125 level, leukocyte count and AFP level, with the AUC of 0.996 and the accuracy of 97.47%. The feature combination of CA19-9 level, CA125 level, leukocyte count, AFP level and CEA level also results in the same high accuracy of 97.47%, however, the AUC value is a bit lower, with the value of 0.99.

The box and jitter plots representing the distributions of the selected four features for PSO-RF-based classification of ruptured and unruptured OEs are shown in Fig. 5. It can be observed that the distributions are significantly different between the ruptured and unruptured OEs. The levels of CA19-9, CA125, leukocyte count and AFP of the patients with ruptured OEs are all significantly higher than those of patients with unruptured OEs.

2) COMPARISONS WITH OTHER ALGORITHMS

The proposed PSO-RF model is comprehensively benchmarked by comparing with eight fine-tuned machine learning models, including four gridsearch optimized (GO) models, namely the GO-RF, GO-lightGBM, GO-LR and GO-KNN, three PSO enhanced models, namely PSO-lightGBM, PSO-LR and PSO-KNN, and the NBC model. The parameters obtained by grid search are shown in Table 2. NBC model has no parameter tuning for there is no parameter to be fine-tuned in it.

Table 4 summarizes the performance of all the nine models, including the classification accuracy rates, the sensitivity and the specificity. Besides the performance results obtained from the dataset fragmentation shown in Table 1, Table. 4 also lists the accuracy rates obtained by hold-out cross-validations (40% holdout samples, 100 times) and 10-fold cross-validations, respectively, to evaluate the performance more comprehensively. The distributions of the cross-validation results are presented as box plots, as shown in Fig. 6. For each test in the 40% holdout cross-validations, the dataset is divided into a training set and a test set by repeated random sampling until the P values with respect to all the four features between the two sets are greater than 0.2 to ensure that there is no significant correlation between the samples of the two sets. As to the 10-fold cross-validations, the P values are not guaranteed to be greater than 0.2.

TABLE 4. Performance comparisons among PSO-RF and benchmark models.

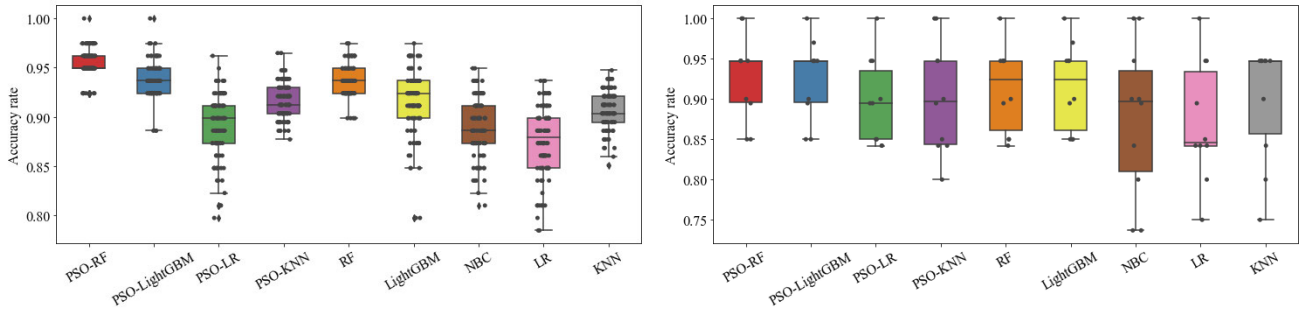
Algorithms	PSO enhanced					Gridsearch optimized				
	Accuracy Holdout†	Accuracy 10-fold‡	Accuracy*	Sensitivity TPR*	Specificity 1-FPR*	Accuracy Holdout†	Accuracy 10-fold‡	Accuracy*	Sensitivity TPR*	Specificity 1-FPR*
RF	**95.47%**	**92.84%**	**97.47%**	94.12%	98.39%	93.58%	91.26%	96.2%	94.12%	98.39%
LightGBM	93.68%	92.54%	96.2%	100%	91.94%	91.56%	91.57%	93.67%	100%	88.71%
NBC*						88.91%	87.58%	93.67%	88.24%	98.39%
LR	89.10%	89.68%	96.2%	88.24%	93.55%	87.58%	87.16%	93.67%	88.24%	83.87%
KNN	91.62%	90.24%	93.7%	94.12%	90.32%	90.61%	89.76%	92.41%	94.12%	93.55%

* The results are obtained from the dataset segmentation shown in Table 1.

† The accuracy rates are the average results of holdout cross-validations (40% holdout samples, 100 times).

‡ The accuracy rates are the average results of 10-fold cross-validations.

* The gridsearch does not applied to the NBC.



(a) Holdout validations (40% holdout samples, 100 times)

(b) 10-fold validations

FIGURE 6. Box plots of cross-validations.

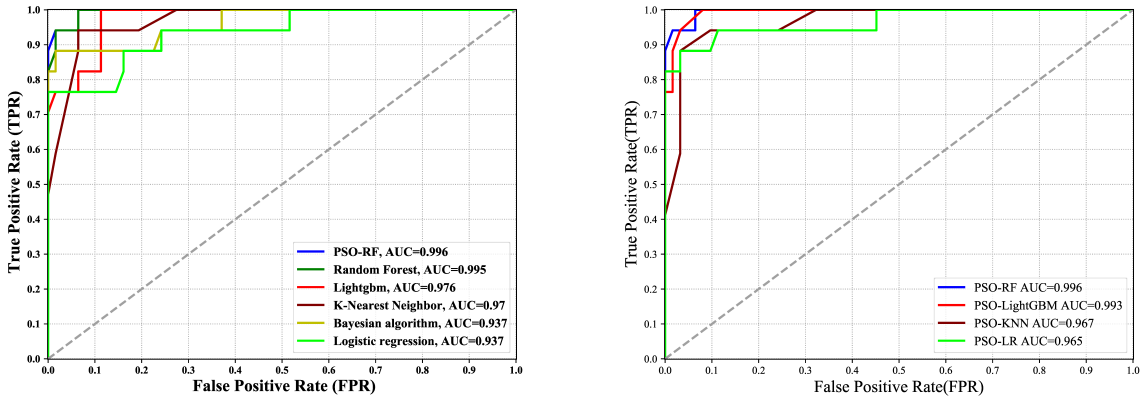
It can be observed that the performances of PSO-RF in both cross-validations are the highest. Specifically, for 40% holdout cross-validations, PSO-RF achieves the highest average accuracy rate of 95.47%, with the minimum standard deviation (S.D.) of 1.36%. The red box of PSO-RF in Fig. 6 is the highest one, and also the narrowest one, with an interquartile range (IQR) of 1.27% and a min-max range of 7.59%. PSO-RF also performs the best in the 10-fold cross-validations, with the highest average accuracy rate of 92.84%. The GO-RF model and the PSO-lightGBM model also achieve high average accuracy rates, i.e., 93.58% and 93.68% in 40% holdout cross-validations, and 91.26% and 92.54% in 10-fold cross-validations, respectively.

Table 4 also shows that PSO does improve the performance of these models. It slightly improves the accuracy rates of the RF model, the lightGBM model, the LR model and the KNN model by 1.89%, 2.13%, 1.52% and 1.01%, respectively, in the 40% holdout cross-validations, and 1.58%, 0.97%, 2.53% and 0.47%, respectively, in the 10-fold cross-validations. The reason why the performance improvements are limited is that the parameters of the original models have also been optimized by grid search. However, the time complexity of PSO is $O(N_{\text{iter}}N_{\text{PSO}}\log N_{\text{PSO}})$ where N_{iter} is the number of iterations and N_{PSO} is the size of the swarm. Whereas that of the gridsearch approach is $O\left(\prod_i n_i\right)$ where $n_i, i = 1, 2, \ldots, j$ is number of potential values of feature i and j is the number of features to be fine-tuned, which is much higher than that of PSO.

The superiority of the PSO-RF model is also presented in Fig. 7 where the ROCs of these models are plotted. It can be observed intuitively that the curve of PSO-RF is the closest one to the upper left corner (AUC = 0.996), implying that it outperforms all the others for its highest the overall accuracy. The ROCs of the PSO-lightGBM model and the GO-RF model are also quite closer to the upper left corner, with slightly lower AUC values, i.e., 0.996 for the PSO-lightGBM model and 0.995 for GO-RF model, respectively.

The confusion matrices of the PSO-RF model and the benchmark models are shown in Fig. 8. PSO-RF performs pretty well. There is only one false positive case and one false negative case, respectively. By contrast, one false negative and two false positives are made by the GO-RF model and the PSO-lightGBM model. The PSO-LR model, the GO-lightGBM model and the NBC model recognize all the patients with unruptured ovarian endometriomas, however, make three, four and five false positives, respectively. In addition, the PSO-KNN model makes two false positives and two false negatives, the GO-LR makes one false negative and four false positives, whereas the GO-KNN model makes four false negatives and two false positives, all of which are inferior to the proposed PSO-RF model.

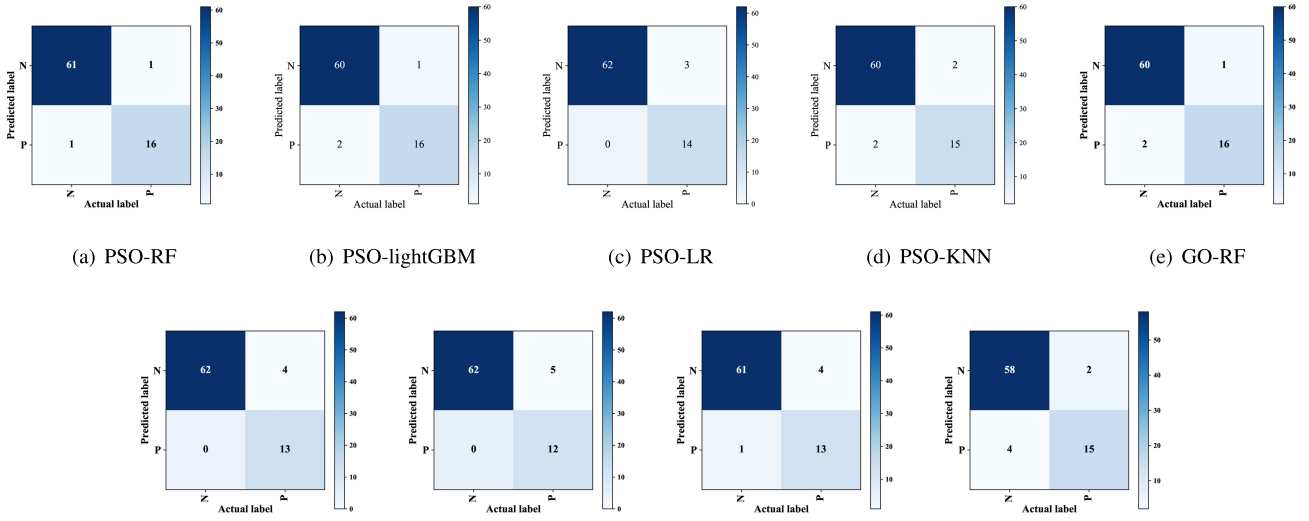
In addition to the classification accuracy, the efficiency of these classifiers is also evaluated. For PSO enhanced models, the convergence performances are evaluated since they train the models iteratively to approach the sub-optimal solutions. Fig. 9 shows the comparisons of the convergence



(a) PSO-RF versus benchmark models without PSO optimization

(b) PSO-RF versus PSO enhanced benchmark models

FIGURE 7. ROCs of PSO-RF and benchmark models.



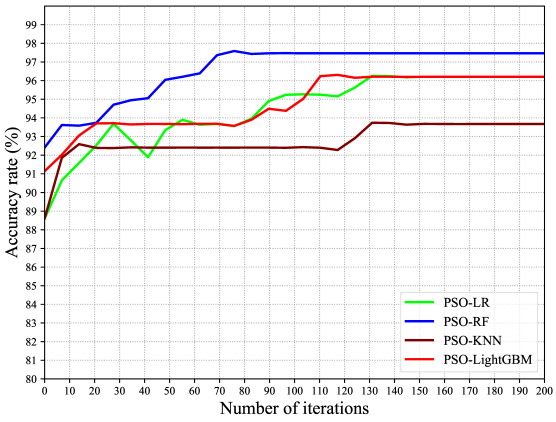
(a) PSO-RF (b) PSO-lightGBM (c) PSO-LR (d) PSO-KNN (e) GO-RF

(f) NBC (g) GO-lightGBM (h) GO-LR (i) GO-KNN

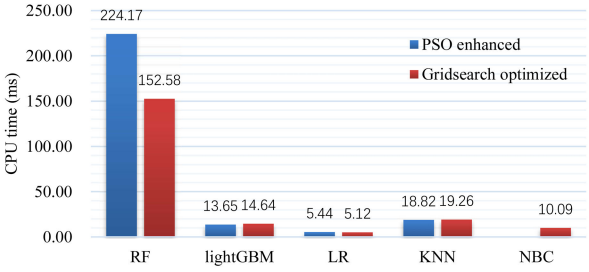
FIGURE 8. Confusion matrices of PSO-RF and benchmark models.

performance among PSO enhanced models. The fitness of all the algorithm is the accuracy rate. It can be observed that the accuracy rate of the PSO-RF model gradually rises in the beginning (within 80 iterations). Then it converges to the stable value of 97.47%. The other three PSO enhanced models, namely the PSO-LR model, the PSO-KNN model and the PSO-lightGBM model converge after the 120th iteration, moreover, with lower accuracy rates.

Fig. 9 only presents the convergence performance along with iterations. However, the time complexities per iteration of difference models vary greatly. In view of this, next, we evaluate the per iteration time complexities of these models. Considering that RF, lightGBM, KNN, NBC and LR are not iterative algorithms, the CPU time required to train a model and then to make a classification are measured instead of the convergence rate. The results averaged from 100 tests are shown in Fig. 9(b), where it can be observed that the CPU time needed by RF is the longest. The reason is that the time complexity of RF is $O(N_T d_{\max} \cdot mn)$, that is, it is determined not only by the number of samples (m) and the number of features (n), but also by the number of CARTs (N_T) and the depth of a CART ($d_{\max}$). As a contrast, the time complexity of LR is only $O(mn)$; that of NBC is $O(cmn)$ where c is the number of classes; that of KNN is $O(kmn)$ where k is the number of neighbors; and the time complexity of lightGBM is co-determined by the procedures of gradient-based one-side sampling and exclusive feature bundling, with the time complexities of $O(lmn)$ and $O(lmn_b)$, respectively, where l is the number of leaves and $n_b \gg n$ is the number of bundles. Specifically, in this work, $N_T = 94$, $d_{\max} = 2$, $c = 2$, $k = 7$ and $l = 9$, so that $N_T \cdot d_{\max} \gg k, l, 2, 1$, results in a much longer CPU time of RF than those of other models. Nevertheless, about 0.2 second CPU time training and classification can fully satisfy the requirements of the classification system.



(a) Convergence comparison among PSO-RF and benchmark models.



(b) CPU times of various models.

FIGURE 9. **Evaluations on convergence rates and CPU times.**

IV. CONCLUSIONS

Spontaneous rupture of OE is vulnerable to be ignored or misdiagnosed clinically, which may probably lead to serious complications to patients. In this work, we proposed a PSO enhanced RF model, namely PSO-RF, to assist in the preoperative diagnosis of spontaneous rupture of ovarian endometriomas. As far as we know, it is the first work on ML-enabled diagnostics of spontaneous rupture of OEs. The data used in this work were collected by complete blood counts, laparoscopic surgery and laparotomy from premenopausal female patients with ovarian endometriomas who have treated in the First Affiliated Hospital of Wenzhou Medical University. We first leveraged the RF model to rank the importance of the features and select the proper feature combination. Then, we employed PSO to fine-tune the essential parameters of the RF model to further improve the diagnosis accuracy.

The PSO-RF model was comprehensively benchmarked by comparing with other machine learning models, including three PSO enhanced models and five gridsearch-optimized models. The results showed that the PSO-RF model outperforms all the other models in accuracy, with the accuracy of 97.47%, the AUC value of 0.996, the sensitivity of 94.12% and the specificity of 98.39%, due to the excellent classification ability of RF and parameter tuning based on PSO. Although the PSO-RF model costs the highest time complexity, the overall complexity is totally acceptable practically. Our work concludes that the PSO-RF model is a highly effective AI-enabled tool for diagnose spontaneous rupture of ovarian endometriomas. It is fully qualified to assist surgeons in diagnosing whether a patient's ovarian endometrioma is ruptured preoperatively.

V. ACKNOWLEDGMENT

(Mingyan Zhou and Feng Lin contributed equally to this work.)

REFERENCES

[1] A. Simonelli, R. Guadagni, P. De Franciscis, N. Colacurci, M. Pieri, P. Basilicata, P. Pedata, M. Lamberti, N. Sannolo, and N. Miraglia, "Environmental and occupational exposure to bisphenol a and endometriosis: Urinary and peritoneal fluid concentration levels," *Int. Arch. Occupational Environ. Health*, vol. 90, no. 1, pp. 49–61, Jan. 2017.

[2] X. Dai, C. Jin, Y. Hu, Q. Zhang, X. Yan, F. Zhu, and F. Lin, "High CA-125 and CA19-9 levels in spontaneous ruptured ovarian endometriomas," *Clinica Chim. Acta*, vol. 450, pp. 362–365, Oct. 2015.

[3] T. A. Gelbaya and L. G. Nardo, "Evidence-based management of endometrioma," *Reproductive Biomed. Online*, vol. 23, no. 1, pp. 15–24, Jul. 2011.

[4] J. H. Pratt and W. R. Shamblin, "Spontaneous rupture of endometrial cysts of the ovary presenting as an acute abdominal emergency," *Amer. J. Obstetrics Gynecol.*, vol. 108, no. 1, pp. 56–62, Sep. 1970.

[5] K. Tanaka, Y. Kobayashi, K. Dozono, H. Shibuya, Y. Nishigaya, M. Momomura, H. Matsumoto, and M. Iwashita, "Elevation of plasma D-dimer levels associated with rupture of ovarian endometriotic cysts," *Taiwanese J. Obstetrics Gynecol.*, vol. 54, no. 3, pp. 294–296, Jun. 2015.

[6] M. Králíčková, A. S. Laganà, F. Ghezzi, and V. Vetvicka, "Endometriosis and risk of ovarian cancer: What do we know?" *Arch. Gynecol. Obstetrics*, vol. 301, pp. 1–10, Nov. 2019.

[7] V. F. D. Amaral, R. A. Ferriani, M. F. S. D. Sá, A. A. Nogueira, J. C. R. E. Silva, A. C. J. D. S. R. E. Silva, and M. D. D. Moura, "Positive correlation between serum and peritoneal fluid CA-125 levels in women with pelvic endometriosis," *Sao Paulo Med. J.*, vol. 124, no. 4, pp. 223–227, 2006.

[8] F. Zullo, E. Spagnolo, G. Saccone, M. Acunzo, S. Xodo, M. Ceccaroni, and V. Berghella, "Endometriosis and obstetrics complications: A systematic review and meta-analysis," *Fertility Sterility*, vol. 108, no. 4, pp. 667–672, Oct. 2017.

[9] D. B. Manders, S. C. Purinton, J. S. Lea, D. L. Richardson, D. S. Miller, and S. M. Kehoe, "Predictive value of carcinoembryonic antigen and cancer antigen 125 in identifying mucinous ovarian tumors," *Gynecol. Oncol.*, vol. 123, no. 2, pp. 445–446, Nov. 2011.

[10] T. Ren, S. Wang, J. Sun, J.-M. Qu, Y. Xiang, K. Shen, and J. H. Lang, "Endometriosis is the independent prognostic factor for survival in Chinese patients with epithelial ovarian carcinoma," *J. Ovarian Res.*, vol. 10, no. 1, p. 67, Dec. 2017.

[11] B.-J. Park, T.-E. Kim, and Y.-W. Kim, "Massive peritoneal fluid and markedly elevated serum CA125 and CA19-9 levels associated with an ovarian endometrioma," *J. Obstetrics Gynaecol. Res.*, vol. 35, no. 5, pp. 935–939, Oct. 2009.

[12] P.-H. Wang, Y.-T. Huang, K.-K. Ng, H.-H. Chou, H.-Y. Lu, S.-H. Ng, and G. Lin, "Detecting recurrent ovarian cancer: Revisit the values of whole-body CT and serum CA 125 levels," *Acta Radiol.*, vol. 60, no. 10, pp. 1360–1366, Oct. 2019.

[13] K. R. Foster, R. Koprowski, and J. D. Skufca, "Machine learning, medical diagnosis, and biomedical engineering research-commentary," *Biomed. Eng. OnLine*, vol. 13, no. 1, p. 94, 2014.

[14] E. Kawakami, J. Tabata, N. Yanaihara, T. Ishikawa, K. Koseki, Y. Iida, M. Saito, H. Komazaki, J. S. Shapiro, C. Goto, Y. Akiyama, R. Saito, M. Saito, H. Takano, K. Yamada, and A. Okamoto, "Application of artificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian cancer based on blood biomarkers," *Clin. Cancer Res.*, vol. 25, no. 10, pp. 3006–3015, May 2019.

[15] M. P. Paing, K. Hamamoto, S. Tungjitkusolmun, S. Visitsattapongse, and C. Pintavirooj, "Automatic detection of pulmonary nodules using three-dimensional chain coding and optimized random forest," *Appl. Sci.*, vol. 10, no. 7, p. 2346, Mar. 2020.

[16] Y. Shi, W. K. Wong, J. G. Goldin, M. S. Brown, and G. H. J. Kim, "Prediction of progression in idiopathic pulmonary fibrosis using CT scans at baseline: A quantum particle swarm optimization–random forest approach," *Artif. Intell. Med.*, vol. 100, Sep. 2019, Art. no. 101709.
[17] S.-C. Cheng, K. Chen, C.-Y. Chiu, K.-Y. Lu, H.-Y. Lu, M.-H. Chiang, C.-K. Tsai, C.-J. Lo, M.-L. Cheng, T.-C. Chang, and G. Lin, "Metabolomic biomarkers in cervicovaginal fluid for detecting endometrial cancer through nuclear magnetic resonance spectroscopy," *Metabolomics*, vol. 15, no. 11, p. 146, Nov. 2019.
[18] V. Pergialiotis, A. Pouliakis, C. Parthenis, V. Damaskou, C. Chrelias, N. Papantoniou, and I. Panayiotides, "The utility of artificial neural networks and classification and regression trees for the prediction of endometrial cancer in postmenopausal women," *Public Health*, vol. 164, pp. 1–6, Nov. 2018.
[19] E. Pashaei and N. Aydin, "Binary black hole algorithm for feature selection and classification on biological data," *Appl. Soft Comput.*, vol. 56, pp. 94–106, Jul. 2017.
[20] W. Cai, Y. Chen, J. Guo, B. Han, Y. Shi, L. Ji, J. Wang, G. Zhang, and J. Luo, "Accurate detection of atrial fibrillation from 12-lead ECG using deep neural network," *Comput. Biol. Med.*, vol. 116, Jan. 2020, Art. no. 103378.
[21] B. Obrzut, M. Kusy, A. Semczuk, M. Obrzut, and J. Kluska, "Prediction of 5–year overall survival in cervical cancer patients treated with radical hysterectomy using computational intelligence methods," *BMC Cancer*, vol. 17, no. 1, p. 840, 2017.
[22] Y.-F. Wang, M.-Y. Chang, R.-D. Chiang, L.-J. Hwang, C.-M. Lee, and Y.-H. Wang, "Mining medical data: A case study of endometriosis," *J. Med. Syst.*, vol. 37, no. 2, p. 9899, Apr. 2013.
[23] E. Günakan, S. Atan, A. N. Haberal, İ. A. Küçükyildiz, E. Gökçe, and A. Ayhan, "A novel prediction method for lymph node involvement in endometrial cancer: Machine learning," *Int. J. Gynecol. Cancer*, vol. 29, no. 2, pp. 320–324, Feb. 2019.
[24] M. J. Downing, D. J. Papke, S. Tyekucheva, and G. L. Mutter, "A new classification of benign, premalignant, and malignant endometrial tissues using machine learning applied to 1413 candidate variables," *Int. J. Gynecol. Pathol.*, vol. 39, no. 4, pp. 333–343, 2020.
[25] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[26] A. Subasi, E. Alickovic, and J. Kevric, "Diagnosis of chronic kidney disease by using random forest," in *Proc. CMBEBIH*, Singapore, 2017, pp. 589–594.
[27] E. Alickovic and A. Subasi, "Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier," *J. Med. Syst.*, vol. 40, no. 4, p. 108, Apr. 2016.
[28] S. M. Raafat, A. M. Hasan, T. W. A. Khairi, and K. G. Ali, "Optimized performance of consensus algorithm in multi agent system using PSO," *Al-Nahrain J. Eng. Sci.*, vol. 21, no. 2, pp. 292–299, Apr. 2018.
[29] M. A. P. Andrade and M. A. M. Ferreira, "Bayesian networks use in simple maternity problems," *Appl. Math. Sci.*, vol. 8, pp. 6963–6967, 2014.
[30] P. Sala, A. Colatutto, D. Fabbro, L. Mariuzzi, S. Marzinotto, B. Toffoletto, A. R. Perosa, and G. Damante, "Immunoglobulin k light chain deficiency: A rare, but probably underestimated, humoral immune defect," *Eur. J. Med. Genet.*, vol. 59, no. 4, pp. 219–222, Apr. 2016.
[31] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3146–3154.
[32] S. Dreiseitl and L. Ohno-Machado, "Logistic regression and artificial neural network classification models: A methodology review," *J. Biomed. Informat.*, vol. 35, nos. 5–6, pp. 352–359, Oct. 2002.
[33] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, *Classification Regression Trees*. Boca Raton, FL, USA: CRC Press, 1984.
[34] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. Neural Netw.*, vol. 4, Nov./Dec. 1995, pp. 1942–1948.
[35] Y. Zhang, S. Yu, R. Xie, J. Li, A. Leier, T. T. Marquez-Lago, T. Akutsu, A. I. Smith, Z. Ge, J. Wang, T. Lithgow, and J. Song, "PeNGaRoo, a combined gradient boosting and ensemble learning framework for predicting non-classical secreted proteins," *Bioinformatics*, vol. 36, no. 3, pp. 704–712, Aug. 2019.
[36] B. Xue, M. Zhang, W. N. Browne, and X. Yao, "A survey on evolutionary computation approaches to feature selection," *IEEE Trans. Evol. Comput.*, vol. 20, no. 4, pp. 606–626, Aug. 2016.

MINGYAN ZHOU is currently pursuing the bachelor's degree with the College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou, China. She is currently working on machine learning applications in intelligent medicine.

FENG LIN received the B.S. degree in obstetrics and gynecology and the M.S. degree in gynecology from Wenzhou Medical University, Zhejiang, China, in 2003 and 2011, respectively. She is currently an Associated Physician and a Professor with the First Affiliated Hospital of Wenzhou Medical University. She mainly specializes in malignant tumors of gynecology and women infertility, and is good at surgical techniques in laparoscopy and hysteroscopy.

QIAN HU received the B.S. degree in communication engineering from the Nanjing Institute of Posts and Telecommunications, Nanjing, China, in 2001, and the M.S. degree in circuits and systems from Zhejiang University, Hangzhou, China, in 2005. She is currently an Associated Professor with the College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou, China. Her research interests are in the areas of machine learning, intelligent systems, and wireless networks.

ZHENZHOU TANG (Senior Member, IEEE) received the B.S. degree in electronic engineering and the M.S. degree in communications and information system from Zhejiang University, in 2001 and 2004, respectively, and the Ph.D. degree in communications and information system from the Dalian University of Technology, in 2015. In March 2004, he joined Wenzhou University. He is currently a Professor with the College of Computer Science and Artificial Intelligence, Wenzhou University, and the Director of the Intelligent Network Innovation Research Center, Wenzhou University. From February 2015 to August 2015, he was a Visiting Scholar with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA, USA. From November 2015 to June 2018, he was a Postdoctoral Research Fellow with Zhejiang University. Since September 2018, he has been a joint Ph.D. Supervisor with Chonnam National University, South Korea. He is currently serving as an Associate Editor of IEEE Access. His current research interests include machine learning, intelligent systems, and wireless networks.

CHU JIN received the M.Sc. degree in information security from the Royal Holloway and Bedford New College, University of London, U.K., in 2002. He is currently a Lecturer with the Renji College, Wenzhou Medical University, Wenzhou, China. His research interests are in the areas of computer security and network security.