International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)
ISSN: 2305-0543 (Online), ISSN: 2411-6173 (Print)
Vol. 12 (43), Jan. 2022, pp. 5101-5105

Breast Cancer Detection using Modified Particle Swarm Optimization

Zeinab Hassani
Department of Computer Science
Kosar university of Bojnord
Bojnord, Iran
Hassani@kub.ac.ir
Tel Number: +98-058-32262478

Vahid Hajihashemi
Faculdade de Engenharia, Universidade do Porto,
Porto, Portugal
Hajihashemi.vahid@ieee.org
Tel Number: +351-961-700648

***Abstract*— Early detection of cancer is the greatest approach for curing cancer and increasing the survival rate. people suffer from cancer in the world. Now, advanced Artificial intelligence algorithms help to treat and detection of cancer. This paper is applied a Modified Particle Swarm Optimization using Support vector machine for diagnosed breast cancer on Wisconsin Breast Cancer UCI Dataset. moreover, the proposed model can identify effective features of breast cancer. The results demonstrate that feature selection can improve accuracy of classifiers and the proposed model supplies the superior diagnosis of breast cancer to the existing methods.**

Keywords—Breast cancer; Supoort vector machine; Modified Particle Swarm Optimization; classification.

I. INTRODUCTION (*HEADING 1*)

Breast cancer is cancer that starts in the cellular units whose function is to secrete milk, the ducto-lobular units of the breast, mainly in women. 8 out of 10 breast cancers occur after the age of 50 [1]. Early diagnosis of breast cancer can be a determining point between life and death. This disease incidence and mortality rates vary by race and age, however, it is highly curable when it is diagnosed early and before it metastasizes. Modern approaches to the diagnosis of breast cancer use supervised learning (SL) to detect tumors with high accuracy [2].

Treatment of breast cancer is performed manually but it will take a longer period of time. Moreover, traditional diagnosis of breast cancer is lead to human errors. Therefore, inteliigent system for breast cancer detection is vital in the medical. several algorithm of machine learning can be preform for breast cancer detection such as neural network, K-nearest neighber, Svm, decision tree[3,4,5].

Furthermore, Irrelevancy and redundant in the features affect classification algorithm and lead to decrese precision of diagnois approaches that feature selection can afford the problem and has important role in classification. Feature selection is one of the pre-processing techniques in machine learning and appleyd in the fields of statistics, pattern recognition and medical domain[5]. thus, feature selection and accuracy improvement is significant challenging. Feature selection tend to reconize the features most relevant for classification and can be broadly classified as either subset selection methods or ranking methods. The former type recognizes a subset of the original set of features which are considered to be the most important for classification. Ranking methods sort the features according to their usefulness in the classification task [6].

The particle swarming optimization method, created by James Kennedy and Russell Eberhart, is currently among the meta-heuristics of optimization algorithms based on patterns of nature (such as the representation of the movement of each individual within a band of birds or a school of fish) most popular in this area, and emerges as being the most promising algorithm for solving several optimization problems, both in the area of science and engineering. Since its creation, many variants have been developed to solve practical problems related to optimization [7].

In this study, feature selection method is presented based PSO algorithm and Supoort vector machine (SVM) classification. There are variants PSO that are presented to improve performance. Modified PSO is considered to diagnoise breast cancer and to select best feature of breast cancer which could lead to a decrement in breast cancer mortality rates

The remainder of the paper is organized as follows: section 1 includes algorithms. Section 3 describes the methodology which is used for the proposed model. section 4 discusses the results and Section 5 concludes the proposed model.

II. MATERIALS AND METHODS

A. Particle Swarm Optimization(PSO)

Actually, Particle swarm optimization(PSO) is a branch of artificial intelligence also classified by some authors as a branch of evolutionary computing (EC). The particle swarm method was proposed by Kennedy and Eberhart in 1995 [8]

The PSO is inspired by the behaviors of different species to meet their needs in the search space ("search space"). However, some meta-heuristics like PSO do not guarantee that an ideal solution will be found. In general terms, the algorithm is guided by personal experience (Pbest), general experience (Gbest) and the current particle movement to decide the next positions in the research space. The motion of each particle effects its position particle is influenced by its position in the best known location, but it is also guided towards the best known positions. This is to be expected when the aim is to move the swarm towards the best solution [7]

Article History:
Received Date: Sep. 21, 2021, Accepted Date: Dec. 19, 2021, Available Online: Jan. 01, 2022

This is how a group of particles are created randomly at the beginning of the work and try to find the optimal solution by updating the generations. In each step, each particle is updated using the top two values. The first is the best situation the particle has ever been able to reach. This position is known and maintained, which is also called the best amount of nostalgia for that particle, which we display with pbest. The best values used by the algorithm is the best position ever obtained by the particle population, which we call gbest (collective intelligence). Equations 1 and 2 can be used to find the best values based on the velocity and location of each particle.

$$v(t+1) = v(t) + c1 * rand(t) * (pbest(t) - position(t)) + c2 * rand(t) * (gbest(t) - position(t)) \quad (1)$$

$$position(t+1) = position(t) + v(t+1) \quad (2)$$

The right-hand side of Equation 1 consists of three parts, the first of which is the current velocity of the particle, the second and third parts are responsible for the rate of change of the particle speed and its direction towards the best personal experience and the best group experience (collective intelligence). If we do not consider the first part in this equation, The particle velocity is then determined only by the current position and the best particle experience and the best aggregate experience, and the effect of the current velocity and inertia is practically eliminated. In this way, the best particle in the group stays in place and the others move towards that particle.

In fact, the mass motion of particles without the first part of Equation 1 would be a process in which the search space gradually shrinks and a local search is formed around the best particle. In contrast, if we consider only the first part of Equation 1, the particles go their normal way to reach the wall of the range and do a kind of global search. Parameters c1 and c2 (the value is about 2) determine the importance and weight of collective intelligence [7,8]

B. *Modified Particle Swarm Optimization*

Up to now, a great variuose of PSO have been introduced to increse the performance of original PSO algorithm.The main improvement strategy of PSO can be classified as Tuning control parameters, Hybrid PSO, Changing the topological structure, Eliminating the velocity formula, Changing the learning Strategy [9,10]. however there are still problems of premature convergence and poor performance of PSO on complex optimization problems [10].

Modified Particle Swarm Optimization (MPSO) is presented Using Adaptive Strategy which avoid the premature convergence phenomenon of PSO and improve its performance on complex problems. In MPSO, the four major strategys are described first strategy try to balance exploration and exploitation better by enlarging or shrinking the search step by a chaos-based non-linear inertia weight is proposed. the logistic chaos is introduced into the inertia weight by Eq.(3), so a non-linear inertia weight is constructed, which is defined by Eq.(4)

$$r(t+1) = 4r(t)\big(1 - r(t)\big)r(0) = rand \quad (3)$$

Where r_0 6∈ {0, 0.25, 0.5, 0.75, 1}

$$\omega(t) = r(t) \cdot \omega_{min} + \frac{(\omega_{max} - \omega_{min}) \cdot t}{T_{max}} \quad (4)$$

Where $\omega_{max} = 0.9$, $\omega_{min} = 0.4$. r(t) is a random number generated by the logistic chaotic.

second strategy is Stochastic and mainstream learning strategies by Eqs.(5) and (6) which are designed to replace the self and global learning strategies, which can effectively enhance the diversity of population and avoid premature convergence. At each iteration, the better personal best will be considered from two mutually different personal best particles are randomly chosened from the population, then, the current personal best (Pbesti) is compared with the CPbest by their fitness value for the better solution. In this way, particles will learn from SPbest to update their velocities heir velocities

$$\mathrm{CPbest}(t) = \min\{\mathrm{Pbesta}(t), \mathrm{Pbestb}(t)\} a \neq b \quad (5)$$

$$\mathrm{SPbesti}(t) = \begin{cases} \mathrm{CPbest}(t) & \mathrm{fit}(\mathrm{CPbest}) < \mathrm{fit}(\mathrm{Pbesti}) \\ \mathrm{Pbesti}(t) & \mathrm{otherwise} \end{cases} \quad (6)$$

Thus, the velocity update equation will calculate according to Eq.(7).

$$Vi(t+1) = \omega(t)Vi(t) + r1c1 \otimes \big(SPbesti(t) - Xi(t)\big) + r2c2 \otimes \big(Mbest(t) - Xi(t)\big) \quad (7)$$

where $c1 = c2 = 2$, $r1 \sim U(0, 1)$, $r2 \sim U(0, 1)$, and $\otimes$ denotes the element-wise product of two vectors. items of the right hand of Eq.(٧) called inertial component, stochastic learning component, mainstream learning component respectively.

Adaptive Position Updating Strategy is the third strategy in which an adaptive position updating strategy is introduced to further balance the exploration and exploitation process. In our investigation, we found that $X = X + V$ is beneficial to local exploitation, $X = \omega X + (1 - \omega)V$ is conductive to the global exploration. This strategy is represented by Eqs.(8) and (9).

$$p_i = \frac{exp(fit(X_i(t)))}{exp(\frac{1}{N} P_{i=1}^{N} fit(X_i(t)))} \quad (8)$$

$$Xi(t+1) = \begin{cases} \omega(t)Xi(t) + (1-\omega(t))Vi(t+1) + Gbest(t) & Pi > rand \\ Xi(t) + Vi(t+1) & otherwise \end{cases} \quad (9)$$

where fit(·) is the fitness of the corresponding particle, N is the size of population. when the pi is larger, particle i is considered to be weaker than the average level of population, and the position updating strategy $X = \omega X + (1 - \omega)V$ is adopted to enhance its local exploitation ability

Terminal Replacement mechanism is the last strategy that Inspired by the survival of the fittest rule in nature, a terminal replacement mechanism is adopted to enhance the convergence precise of MPSO. The global worst particle and a new particle Nbest are defined by as Eq.(10),(11). then, if the new generated Nbest is better than the Gworst, then the corresponding personal best of Gworst will be replaced by Nbest, otherwise, the corresponding personal best of Gworst will remain unchanged for comparing Gworst with Nbest. which is defined as Eq.(12)

$$Gworst(t) = max\{Pbest1(t), Pbest2(t), \cdots, PbestN(t)\} \quad (10)$$

$$Nbest(t) = Gbest(t) + rand \cdot (Pbestj(t) - Pbestk(t))\; j \neq k \quad (11)$$

IEEE Office - SEET Complex, Federal University of Technology, Owerri, Nigeria, https://aeuso.org
EISSN: 2305-0543, PISSN: 2411-6173

$$Gworst(t) = \begin{cases} Nbest(t) & fit(Nbest(t)) < fit(Gworst(t)) \\ Gworst(t) & otherwise \end{cases} \quad (12)$$

MPSO algorithm is displyed in algorithm 1.

Algorithm 1: MPSO algorithm

Input: Population size:N, the maximum number of iterations T_{Max}, inertia weight ω, location boundary $[x^d_{min}, x^d_{max}]$, velocity boundary $[v^d_{min}, v^d_{max}]$ $d = 1, 2 \cdots, D$, $v_{max} = 0.5 \cdot (x_{max} - x_{min})$, $v_{min} = -v_{max}$;
Output: Optimal solution;

Initialize:

$x^d_i = x^d_{min} + rand \cdot (x^d_{max} - x^d_{min})$,

$v^d_i = v^d_{min} + rand \cdot (v^d_{max} - v^d_{min})$;

for $t = 1, 2, \cdots, T_{Max}$ (Update *Ubest*) **do**
  for $i = 1 : N$ **do**
    if *fit*(*Ubest*) < *fit*(*P best$_i$*) **then**
      $Rbest_i = Ubest$;
    end
  end
  Update location of particle by Eqs.(8) and (9). Check the boundaries ;
  % Adaptive Position Updating Strategy ;
  for $i = 1 : N$ **do**
    if $p_i > rand$ **then**
      $X_i(t + 1) = \omega(t)X_i(t) + (1 - \omega(t))V_i(t + 1) + Gbest$;
    else
      $X_i(t + 1) = X_i(t) + V_i(t + 1)$;
    end
  end
  Check the boundaries ;
  % Terminal Updating Strategy ;
  for $i = 1, 2, \cdots, N$ **do**
    Update $Gworst_i$ by Eq.(10);
    Update $Nbest_i$ by Eq.(11) and Bounds checking.;
    if *fit*(*Nbest*) < *fit*(*Gworst*) **then**
      Gworst = *Nbest* ;
      fit(*Gworst*) = *fit*(*Nbest*);
    end
  end
  for $i = 1, 2, \cdots, N$ **do**
    if $fit(X_i) < fit(Pbest_i)$ **then**
      $P\ best_i = X_i$;
      $fit(Pbest_i) = fit(X_i)$
    end
    if $fit(Pbest_i) < fit(Gbest)$ **then**
      $P\ best_i = Gbest$;
      $fit(Gbest) = fit(P\ best_i)$;
    end
  end
end

C. Support Vector Machine Svm

The Support Vector Machine is a distinguishing classifier defined by a separator. In other words, by receiving labelled training data (supervised training), the algorithm outputs an optimal separator that classifies new samples. In 2D space, this is a dividing line of a page into two parts that are placed on both sides in each class[12,13].

SVM uses a technique called kernel trick to convert your data and then, based on that conversion, finds the optimal boundary between possible outputs. In simple terms, it performs very complex conversions, then specifies how to separate your data based on the tags or outputs you define.

One of the methods that are currently widely used for the classification problem is the support vector machine (SVM) method. Perhaps the current popularity of the support vector machine method can be compared to the popularity of neural networks over the past decade. The reason for this is the ability to use this method to solve various problems, while methods such as the decision tree can not be easily used in various problems.

In the simplest case, the goal is to find the best line based on a linear separation function that can separate the two categories. $x = (x_1, \ldots, x_N)^T$, with a weight vector

$$\omega = (w_1, \ldots, w_N)^T : h(x) = \omega^T x + \omega_0 \quad (13)$$

It is then decided that x is of class 1 if $h(x) \geq 0$ and of class -1 otherwise. It is a linear classifier.
The decision frontier $h(x) = 0$ is a hyperplane, called a separator hyperplane, or separator. The goal of a supervised learning algorithm is to learn the function $h(x)$ through a training set:

$$\{(x_1, \iota_1), (x_2, \iota_2), \ldots, (x_k, \iota_k), \ldots, (x_p, \iota_p)\} \subset \square^N \times \{-1,1\} \quad (14)$$

where the l_k are the labels, p is the size of the training set, N the dimension of the input vectors. If the problem is linearly separable, one must then have:
$l_k \square(x_k) \geq 0 1 \leq k \leq p$, in other words

$$l_k(\omega^T x_k + \omega_0) \geq 0 1 \leq k \leq p.$$

III. RESEARCH MATERIALS AND METHOD

In this study, MPSO using SVM are presented for the diagnosis of breast cancer early detection Which can classify breast cancer and identify effective attributes. The classification model is proposed with high accuracy to predict breast cancer patients. In the following, important phases of the model are described in two steps.

First step is the preprocessed dataset which is normalized data in the interval [0,1]. Data normalization is the process of rescaling one and more attributes to the range of 0 to 1. This means that the largest value for each attribute is 1 and the smallest value is 0. Normalization is a good technique to use when you do not know the distribution of your data.

In the second step, After preprocess applied MPSO algorithm with SVM classification and compute the accuracy for finding the best attributes. Best attributes are presented to SVM classification using the MPSO in every iteration. then, the 10-fold cross-validation method is applied to test the effectiveness of the SVM classification.

Performance of the prediction model is evaluated y considering the actual and predicted classification. accuracy, precision, recall and F1- the measure is being calculated. From

IEEE Office - SEET Complex, Federal University of Technology, Owerri, Nigeria, https://aeuso.org
EISSN: 2305-0543, PISSN: 2411-6173

the confusion matrix information. in the 10-fold cross-validation method, mean of evaluation maeures are calculted and best result of them are selected.

A. Data

This study is used for the experiment the Wisconsin Breast Cancer Diagnosis (WBCD) dataset from UCI repository that consisted of 569 instances and 32 attributes, with the patient ID, cell nuclei features, and diagnosis that is two class attributes: malignant or benign [14].

The ID is the patient identification number, and the cell nuclei features were determined from a digital image of a fie needle aspirate (FNA) of a breast mass. These features describe 10 characteristics of each cell nucleus where consists of three features: mean, standard error, and worst. therefore, a total of 30 features of 569 patients were evaluated. Of all cases, there are 357 benign cases and 212 malignant ones. Table --- displays the description of the attributes.

TABLE I. DESCRIPTION OF WISCONSIN BREAST CANCER DIAGNOSIS (WBCD) DATASET

num	*Name*	*Description*
1	Id	Id Number
2	Diagnosis	The diagnosis of breast tissues (M = malignant, B = Benign)
3	Radius	Mean of distances from the center to points on the perimeter
4	Texture	Standard deviation of grayscale values
5	Perimeter	Mean size of the core tumor
6	Area	Mean area of the core tumor
7	Smoothness	Mean of local variation in radius lengths
8	Compactness	Mean of $perimeter^2/area - 1$
9	Concavity	Mean of severity of concave portion of the contour
10	Concave points	Mean for number of concave portions of the contour
12	Fractal	Mean for *coastline approximation* − 1

IV. RESULT AND DISCUTION

This section contains the results of proposed model that are compared to the result of other researchers. we used Wisconsin Breast Cancer Dataset. The main configurations of the computer for the algorithm test are as follows: Window 10 (64bit) and MATLAB R2017a for the software environment.

In this paper, we proposed an optimization algorithm with SVM classification for diagnosis and detection of breast cancer. Also, proposed model can identify effective features of detection of breast cancer. standardization method for preprocessing dataset is used then we have applied proposed model with initial parameter which displays in table II.

TABLE II. INITIAL PARAMETER OF PROPOSED MODEL

num	*Parameter*	*value*
1	Iteration	50,100
2	Population	15,25
3	c1,c2	2

This method is used 30 features for problem. After the implementation of this method, We have seen the proposed method achieved 98.77 accuracy. Attributes 4,5,8,9,11,12,16,18,20,21,22 are selected in 4 implements of this Model. Thus, 11 effective features are recognized by MPSO using SVM for the breast cancer diagnosis. Results are shown in Table III.

TABLE III. RESULTS OF PROPOSED MODEL

Population	*Iteration*	*Accuracy*	*precision*	*Recall*	*F_measure*
	50	98.25	99.17	98.09	98.61
	100	98.25	98.89	98.36	98.61
	50	98.24	99.17	98.10	98.61
	100	**98.77**	**99.71**	**98.37**	**99.03**

The results achieved of this model are compared with the results reported by another researcher in the existing literature. We mainly focused on the method used and the accuracy achieved by the other studies.

As can be seen in Table 4, the proposed model has achieved a better result than the other works on breast cancer. In addition to the accuracy, a better result has been obtained in the other three criteria(precision, recall and F1- the measure) which shows the superiority of our proposed model.

TABLE IV. COMPARED OF PROPOSED METHOD WITH RECENT WORK FOR BREAST CANCER

Ref.	*method*	*Result*
[11]	Fuzzy- ID3 algorithm	Accuracy=94.534
[2]	supervised learning and Semi-Supervised learning[2021]	Accuracy=98
[3]	Extreme Learning Machine [2021]	Accuracy=98.68 Precision= 73.26 Recall=91.30 F- measure= 81.29
Our study		Accuracy=98.77 Precision=99.71 Recall=98.37 F- measure=99.03

V. CONCLUSION

This paper proposed a model for diagnosis and identify effective attributes of breast cancer using MPSO with an SVM classifier. The proposed method is reliable and managed to generate high performances in the classification of breast cancer data. The performance of the proposed model was compared with some state-of-the-art technologies for breast cancer diagnosis. The results achieved on the Wisconsin Diagnostic Breast Cancer dataset illustrate that model outperforms other researches. The experimental results indicate that the accuracy achieved is 98.77, the recall is 99.71, the precision is 99.37, and the F1-score is 99.03 which The best performance results of the proposed model were obtained.

REFERENCES

[1] Chugh, G., Kumar, S., & Singh, N. “Survey on Machine Learning and Deep Learning Applications in Breast Cancer Diagnosis,” Cognitive Computation, 1-2,2021

[2] N. Al-Azzam , I. Shatnawi, “Comparing supervised and semi-supervised Machine Learning Models on Diagnosing Breast Cancer,” Annals of Medicine and Surgery, vol.62, pp.53–64, 2021

[3] V. Lahoura, H. Singh, A. Aggarwal, B. Sharma et al, “Cloud Computing-Based Framework for Breast Cancer Diagnosis Using Extreme Learning Machine,” Diagnostics, vol.11, 2021

[4] A. Ed-daoudy, K. Maalmi, “Breast cancer classifcation with reduced feature set using association rules and support vector machine,” Network Modeling Analysis in Health Informatics and Bioinformatics , vol.9(34), 2020

[5] Sh. Aalaei. H. Shahraki, AR. Rowhanimanesh, S. Eslami, “Feature selection using genetic algorithm for breast cancer diagnosis: experiment on three different datasets,” Iran J Basic Med Sci, vol.19 pp.476-482, 2016

[6] S.Sasikala, D.S.A. alias Balamuruganb and D.S. Geetha, “A Novel Feature Selection Technique for Improved Survivability Diagnosis of Breast Cancer,” Procedia Computer Science 50, pp.16 – 23, 2015.

[7] Li, W., Wang, G. G., A. H. Gandomi, “A survey of learning-based intelligent optimization algorithms,” Archives of Computational Methods in Engineering, pp. 1-19, 2021.

[8] Kennedy, J., & Eberhart, R. C. (1995). Particle swarm optimization. In Proceedings of the 1995 IEEE international conference on neural networks (Perth, Australia) (pp. 1942–1948). Piscataway, NJ: IEEE Service Center.

[9] Lin Xu, Baoye Song & Maoyong Cao, “An improved particle swarm optimization algorithm with adaptive weighted delay velocity,” Systems Science & Control Engineering, vol.9, pp.188-197, 2021

[10] Hao Liu, Xue-Wei Zhang, Liang-Ping Tu , “A Modified Particle Swarm Optimization Using Adaptive Strategy,” Expert Systems With Applications, 2020, doi: https://doi.org/10.1016/j.eswa.2020.113353

[11] Idris NF, Ismail MA. 2021. Breast cancer disease classification using fuzzy-ID3 algorithm with FUZZYDBD method: automatic fuzzy database definition. PeerJ Comput. Sci. 7:e427

[12] Amulya, K. J., Divya, S., Deepali, H. V., & Ravikumar, V.” A Survey on Diabetes Prediction Using Machine Learning”. In ICCCE 2020. Springer, Singapore, pp. 1049-1057, 2021.

[13] Z. Hassani, M. Alambardar Meybodi, and V. Hajihashemi, “Credit Risk Assessment Using Learning Algorithms for Feature Selection,” Fuzzy Information and Engineering, 2021, doi: 10.1080/16168658.2021.1925021.

[14] https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/.