THEORETICAL ADVANCES

Hybrid ABC and black hole algorithm with genetic operators optimized SVM ensemble based diagnosis of breast cancer

Indu Singh[1] · K. G. Srinivasa[2] · Mridul Maurya[1] · Aditya Aggarwal[3] · Himanshu Sheokand[1] · Harsh Gunwant[4] · Mohit Dhalwal[1]

Abstract

Forever and a day, breast cancer has caused significant negative impacts on the quality of lives of number of women, more often than not turning into a fatal disease. The growth in the number of such cases has constantly been a major concern for the community as well as medical experts. To prevent irreversible damages caused by the disease, early identification of breast cancer is essential. Various researches and techniques have been devised in the past as an attempt to achieve this task with appreciable accuracy. As an advancement to these pre-existing algorithms and methods, we have devised a model by exploiting the techniques of nature-inspired metaheuristics in order to efficiently detect breast cancer at an early stage while maintaining acceptable levels of accuracy. In this paper, we propose a hybrid model, namely "hybrid artificial bee colony and black hole with genetic operators (GBHABC)", for the early detection of breast cancer. In the proposed model, we employed a support vector machine (SVM) ensemble technique, optimized using the proposed GBHABC model. This model combines the techniques of two major algorithms, namely artificial bee colony (ABC) and black hole (BH), guided through crossover and mutation genetic operators. Datasets from the well-known UCI breast cancer repository have been used to train the models and evaluate test result. For a fair and accurate evaluation of the model, a number of metrics have been examined including accuracy, sensitivity, specificity, F1-score and precision. An impeccable accuracy of 99.42% was obtained on the UCI dataset, clearly outperforming any literature in the same field.

Keywords Breast cancer diagnosis · Genetic operators · Artificial bee colony · Black hole algorithm · Ensemble learning

1 Introduction

Cancer is the uncontrolled regeneration and reproduction of untypical and unwanted cells in different parts of a body. Cancer really is a comprehensive term for a group of diseases caused when untypical or unusual cells grow up in different parts of the body, thus also known as neoplasm (literally meaning, a new growth). There are hundreds of strains of cancer diseases such as colon cancer, oral cancer, skin cancer, prostate cancer, breast cancer and lung cancer, depending upon the site of reproduction of the cells. Among all these types of cancers, breast cancer

✉ Indu Singh
indusingh@dtu.ac.in

K. G. Srinivasa
srinivasa@iiitnr.edu.in

Mridul Maurya
mridulmaurya99@gmail.com

Aditya Aggarwal
aditya99.iitd@gmail.com

Himanshu Sheokand
sheokandhimanshu@gmail.com

Harsh Gunwant
hgunwant@gmail.com

Mohit Dhalwal
mohitdhalwal12@gmail.com

[1] Department of Computer Science and Engineering, Delhi Technological University, Delhi, India

[2] Department of Data Science and Artificial Intelligence, International Institute of Information Technology, Naya Raipur, Chhattisgarh, India

[3] Department of Mechanical Engineering, Indian Institute Of Technology, Delhi, India

[4] Department of Computer Science and Engineering, Maharaja Agrasen Institute of Technology, GGSIPU, Delhi, India

(B-Can) is a sort of cancer with contrasting appearances among women. It is the most common type of cancer that affects women throughout the world [3]. A breast lump is the most classic presenting manifestation among women with B-Can and has a relatively outrageous predictive value for it to be malignant [27].

Based on histological and molecular evidences, B-Can is tabulated into three crucial sub-types based on the presence or absence of molecular indicators for progesterone (PR+) or oestrogen receptors (ER+) and human epidermal growth factor 2 (ERBB/HER2+): hormone receptor-positive (ERBB2-), ERBB2+ and triple-negative B-Can (TNBC) (lacking all 3 standard molecular markers). The treatment and the duration of it vary from the type of B-Can [50]. There are numerous ways of diagnosis and treatment of breast cancer. Some of the methods for treatment are radiotherapy, chemotherapy, immunotherapy (using immune cells already present in breast tissue) [17], combining both radiotherapy and immunotherapy [38], breast conservation therapy (BCT) and mastectomy which offer analogous survival and can be regarded as equivalent therapy in early-stage B-Can. The last two therapies, BCT and mastectomy, are the most conflicting ones for women to decide [18]. Digital mammography (DM) is the most recommended method for diagnosing breast tumours [33]. The diagnosis of breast cancer using IOT-enabled classifier systems is also available in the literature [31]. Following these methods and earlier diagnosis, there is an improvement in B-Can treatment leading to an increase in survival rate. However, such therapies and diagnosis may have some side effects, majorly in pregnant women who may have the chance of miscarriage and terminations [16]. Due to these reasons, patients have started refusing to undergo the treatment which increases the risk of mortality to 2.42 times higher according to the study [15].

Typically being the most commonly occurring cancer type among women and second-most periodic cancer overall, B-Can accounted for as much as 19.3 million new cases and 10 million cancer deaths around the world in the year 2020 (estimate by Cancer Journal) [46]. Though there are some studies that focus on male breast cancer, it still exists to be rare and understudied when compared to female breast cancer, which shows that the 5-year survival rate for male counterparts was lower than that for female patients. In the subcontinent of India itself, it accounted for almost 14 of total cancer cases among Indian women. Figures show that breast cancer has now outgrown even the most frequent forms of cancers including cervical cancer. Being on the rise in both rural and urban parts of the country, every four minutes an Indian woman is diagnosed with this disease in India with post-survival rates reported to be as low as 80 among US women and an even worse 60 among Indian women. B-Can is the second leading cause of cancer-associated mortality in the USA and has been reckoned that 1

out of 8 women will experience breast cancer at some point of their life [15].

It is well known to many that as the stages of this disease grow, the chances of survival of the patient diminish. With more than half of Indian women diagnosed with the disease suffering from stage three–four of breast cancer, the chances of survival seem very far-fetched.

It is seen that a lot of factors affect the chances of developing breast cancer in women. Correlation of different factors includes body mass index (BMI) [8], pregnancy obesity [25], alcohol consumption patterns and intensities [49], active–passive smoking [35], age groups [30] among others helping in more accurately evaluating the risks of developing breast cancer in later ages and chances of recovery. As close as 50% of the total cases fall in the age group of 25-50 years. Making the situation worse, more than 70% of the advanced-stage cases have low survival rates and high mortality rates. Various studies have been done in the past discussing the incidence and mortalities of female breast cancer, particularly for the Asia-Pacific regions.

Due to the most recent COVID-19 pandemic, the patients of B-Can are more susceptible to be infected by the virus, leading to more deaths. Also, due to lockdown and shortage of medical supplies, their treatments and therapies are given a pause, increasing the chances of augmentation of raging cancer cells. The aftereffects of this pandemic included high rates of anxiety, insomnia, distress and depression, especially in breast cancer patients [22]. To beat these effects as well as the disease itself, an active and healthy lifestyle is believed to be very helpful. Physical exercise has been shown to lower the risk of breast cancer, reduce the incidence of recurrence and improvise the survival rate of breast cancer patients. It can also aid with mental wellness and the immune system [10]. The low survival rates of breast cancer patients in the world are due to the failure in the detection of the disease in the early phase. As time passes and the stages of the disease progress, the chances of survival become less. Apart from spreading awareness, the only possible ways to change this situation are to develop and deploy some state-of-the-art accurate and efficient algorithms that can detect and raise alarms when the disease is at its early phase itself. Breast cancer can be cured and the survival rates plausible if detection is done in time.

1.1 Past approaches for detection in brief

Various works and researches have been done in the past for the early phase detection of cancer by the tools and algorithms of nature-inspired metaheuristics.

Kiyan et al. [26] proposed an algorithm which used radial basis function (RGF), general regression neural network (GRNN) and probabilistic neural network (PNN) to classify Wisconsin Breast Cancer Dataset (WBCD) and compare

the result with multilayer perceptron (MLP). Yeh et al. [55] devised a data mining method using the concept of discrete PSO. This method created a novel PSO that has a feasible system structure, and every particle is coded in positive integer numbers. Statistical analysis is used to eliminate feature variables that are insignificant.

Nazarian et al. [37] introduced a model for the classification of binary issues through utilizing information mining procedures, based on artificial bee colony algorithms.

Krawczyk et al. [28] presented a work in progress on a computer-aided breast cancer determination. They utilized a multi-objective memetic algorithm to choose the pool of one-class predictors that show simultaneously high diversity and consistency. Dheeba et al. [12] have advocated a PSO optimized WNN model. The model attempted to increase the classification accuracy of WNN by tuning the initial network parameter using PSO. By using PSOWNN, they evolved connected WNN and optimized it by using the best network architecture which was created by optimizing the momentum factor, learning rate and number of neurons in the hidden layer.

Bhardwaj et al. [4] came up with a new algorithm, genetically optimized neural network (GONN), for problem-solving. They modified the neural network to improve its architecture (structure and weight) of planning. They used the GONN algorithm to classify breast cancer tissue as malignant or benign. To demonstrate their results, they took a WBCD database.

Wang et al. [52] have proposed the CMWOAFS-SVM model which is an improved WOA based on the chaos and control engineering theories which are lodged to optimize two critical parameters for SVM with the RBF kernel and simultaneously select features in practical classification tasks. Similarly systems have also been developed for heart disease prediction and Brainstroke Classification using machine learning [21, 41].

1.2 Brief about our model

Machine learning models integrating metaheuristic optimization algorithms have been employed for various applications including detection of breast cancer. However, such models can often face issues like less generalization and overfitting and need better parameter optimization. Though metaheuristic algorithms have been effective in optimizing SVM ensembles and proposed better ability to find global optima for SVM, synchronization of multiple metaheuristic algorithms needs more exploration for optimizing the combination strategies in SVM ensembles. This should be implemented by considering ensemble diversity and performance improvement. The SVM can be made suitable for specific characteristics of the breast cancer dataset, by optimizing the parameters using the ABC algorithm for better model

performance. In the model, we have employed the artificial bee colony that is based on the communication and cognitive skills of honeybees [24]. The issue with ABC is that it has an inherent problem of stagnation. This suggests that in the first phase, the location gradually changes, and so, the exploration phase moves at a slow pace. We have implemented an improved black hole [11] algorithm as a means to increase convergence of the model and to expand the exploration phase. The classic black hole algorithm possesses difficulty in maintaining a balance between the exploitation and exploration phases; thus, it was modified to tackle the very same limitation. It has a tendency to get stuck in the local optima. The improved black hole algorithm did not face any such limitation, since the stars orbiting around the focal point were at random angles and move to a better position, and therefore, exploration was done efficiently. This algorithm was incorporated into the first phase of the artificial bee colony's exploration, thus making it more efficient, hence establishing a balance between the exploration and exploitation phases. Once a star enters the event horizon, it is destroyed. Traditional black hole algorithms create points randomly in the search space which can possibly make the search space uneven and very different from the initial distribution. To rectify this drawback, we applied mutation and crossover genetic operators. The crossover operation is one of the well-known evolutionary algorithms, such as differential evolution (DE) [45], that can significantly increase an algorithm's effectiveness. Over the binary coded GA, we applied a two-point crossover. The extreme points were rectified using a two-point crossover. In order to arrive at a wider range of options, we employed the crossover operator. After crossover, we also employed swap mutation which is also one of the most prominent genetic operators. This makes sure that no two points, which are generated in different iterations, are exact replicas of each other. This model is then used to optimize a majority voting ensemble technique which classifies the data points as per the task. All these improvements over respective base algorithms work in harmony with each other and work towards combating their limitations. The proposed model can improve the overall performance of SVM ensemble and increase model accuracy. Better parameter optimization and reduction in errors in the proposed breast detection model can produce excellent results.

The contributions of research work are summarized as follows:

1. ABC algorithm is synchronized with improved black hole algorithm to make it more efficient and overcome their respective drawbacks.
2. Crossover and mutation genetic operators are used to generate new stars instead of randomized generation to counter the limitation of black hole algorithm, thereby

making a diversified population and a better newly generated star.

3. A novel GBHABC model is proposed to optimize the weights of the ensemble model for the best possible solution.

4. Majority voting ensemble is used to set priorities for different kernels available in the model.

2 Related work

In the past, a diverse range of datasets were utilized for the early diagnosis of various diseases such as diabetes, breast cancer, hepatitis and others. Many studies used the Wisconsin Prognostic Breast Cancer Dataset (WBCD) to detect breast cancer. This dataset is still used to create and validate various algorithms in the field of breast cancer detection today. In an earlier study, the use of data mining techniques in conjunction with nature-inspired algorithms was seen to solve diagnosis and detection challenges. Some of the major state-of-the-art techniques are illustrated in Table 1.

Yeh et al. [55] proposed a data mining technique using the discrete PSO on WBCD dataset. The proposed method created a novel PSO with a feasible system structure, with every particle being coded in positive integer numbers. Statistical analysis was used to eliminate insignificant feature variables, and after that, only six feature variables were remaining. The proposed hybrid data mining method reduced the computational complexity and sped up the whole data mining process. The proposed method showed satisfactory results on a range of performance metrics including sensitivity, accuracy and specificity.

Chen et.al. [6] showed the superiority of a hybrid SVM-based model rough set-support vector machine (RS-SVM) over other relevant state-of-the-art algorithms for tackling the problem of breast cancer diagnosis. In the proposed RS-SVM model, the RS feature reduction algorithm was employed for feature selection to eradicate the redundant features of the dataset and add to the diagnostic accuracy of the model by SVM. RS feature reduction algorithm was employed on the whole dataset for optimal subset generations, and only a small fraction was retained for the complete study. The combination subset of features used by the model for categorizing the breast tumours was pinned down to be the most promising feature of the reduction algorithm. Moreover, experiments done on altered slices of the WBCD showed that the RS-SVM model performed satisfactorily in distinguishing between the two classes of breast tumour. The performance of this model greatly depended on the optimal parameters of the SVM. Thus, the need for new algorithms to accurately calculate the optimal parameters arose. Moreover, a single dataset like WBCD was used for evaluating the model, thus undermining the accuracy and real-life application of this model.

Chen et al. [7] proposed the PSO-SVM Classifier for breast cancer diagnosis on the WBCD dataset. They incorporated feature selection and parameter optimization in the suggested technique. The proposed approach is divided into two parts. The first stage involves using the PSO algorithm to firmly conduct SVM parameter optimization with feature selection. The second stage involves conducting a classification job with the help of a selected feature and an ideal value using the tenfold approach and the SVM model. In the proposed model, the particle swarm framework simultaneously solved the issue of feature selection and model selection in SVM. For the designing purposes of the objective function of particle swarm optimization a weighted function was adopted in which simultaneously accounted for a number of support vectors, selected features and average accuracy rates of SVM, but the model was not feasible in practical diagnosis as it costed a lot of CPU time for PSO-based system training.

Nazarian et.al [37] put forward a novel classification method by using artificial bee colony algorithm for classification. This algorithm was another straightforward, new, simple and proficient algorithm dealing with real-world scenarios for optimization. In this paper, analysis of this ailment is conceived through utilizing data related to B-Can samples. The aftereffects of this technique were contrasted with different strategies for breast cancer determination.

Krawczyk et al. [28] proposed a CAD-based breast cancer determination, tested on genuine and proven medical datasets of 675 pictures, gathered from 75 patients out of which 25 were amiable, 25 were dangerous and 25 were fibroadenoma. In the study, each patient was addressed by 9 pictures chosen by a pathologist self-assertively. They proposed a productive medical choice support structure that permitted recognizing amiable, dangerous and fibroadenoma cases from the dataset. The nuclei recognition methodology depends on the firefly calculation algorithms. The methodology produced nuclei markers that were utilized in the marker-controlled watershed division. In the final step, an image recognition step was finished by a novel classifier. The model was built utilizing a one-class classifier trouped for a multi-class decomposition, permitting to capture proficiently the unique attributes of inspected classes and henceforth to precisely classifying them. The outcomes introduced in this paper showed that a modernized medical determination system dependent on the proposed strategy would be successful and could give precise diagnostic data.

Sheikhpour et. al. [44] proposed a PSO-KDE-based model: hybridizing the PSO and classifier based on non-parametric kernel density estimation for breast cancer determination on the WBCD. In this model, they used PSO to simultaneously find a subset of features that could minimize a number of features as well as the classification error in the

Table 1 Related work comparison

Authors	Year	Dataset	Proposed algorithm	Benefits	Limitation
Chang et al. [55]	2009	WBCD	A new hybrid approach of discrete PSO and statistical method for mining breast cancer pattern	Reduced computational complexity	Only patterns and correlations may be generated from the data, and their validity cannot be validated. Inaccurate or skewed data gathering might result in erroneous statistical analysis
Hui-Ling Chen et al. [6]	2011	WBCD	Feature selection using rough set-based SVM classifier for breast cancer diagnosis	A subset of features was the most promising feature of the reduction algorithm	The results and accuracy of this model greatly depend on the parameter tunings of the SVM
Yang et al. [7]	2012	WBCD	SVM parameter optimization with feature selection using PSO and classification using SVM with the help of tenfold technique	Particle swarm framework simultaneously solved the enigma of feature selection and model selection in SVM	For practical diagnosis, problem costs a lot of CPU time for PSO-based system
Nazarian et al. [37]	2013	WBCD	Classification of breast cancer using the ABC algorithm	This strategy is easier than different strategies	The success percentage in this proposed method is comparatively low
Krawczyk & Filipczuk [28]	2014	675 pictures were gathered from 75 patients of the Regional Hospital in Zielona Góra, Poland	Analysis of cytological image using hybrid one-class classification decomposition with firefly nuclei detection	Utilizing one-class classifier troupes for a multi-class decomposition permits us to seize proficiently the unique attributes of inspected classes and in result to precisely separate between them	Use of small dataset could lead to overfitting. Fewer data points can improve training set accuracy, but reduce test set accuracy as the model generalizes. An overfit statistical model fits its training data exactly
Dheeba et al. [12]	2014	Real clinical database of 216 mammograms collected from mammogram screening centres	PSO optimized WNN model for determination of breast cancer on mammograms	Improvement in the classification accuracy of computer-aided analysis of the digital mammogram	Accuracy of the model is 93.671%. It cannot be practically used for the diagnosis because of the low accuracy
Sheikhpour et al. [44]	2016	WBCD, WDBC	PSO for bandwidth determination and feature selection of KDE-based classifiers for diagnosis of breast cancer	The proposed approach KDE-PSO has a better mean performance than the mean performance of GA-KDE model in breast cancer detection	The PSO used for minimizing the classification error and number of features is not multi-objective
Shahbeig et al. [43]	2017	Breast cancer microarray dataset	Selection of the associated genes with breast cancer using fuzzy multi-objective hybrid TLBO-PSO approach	Using this model, they were able to enhance the sensitivity to 90.55%, specificity to 93.33% and accuracy to 91.88% by choosing 195 genes	This method can only be used to do prognosis at very early stages only
Wang & Chen [52]	2020	Wisconsin Breast Cancer Dataset (WBCD), Diabetes Data Set, Dermatology Data Set	Medical diagnosis using chaotic multi-swarm whale optimizer boosted SVM	Tackle feature selection and parameter optimization simultaneously	It has a complex search process because it is based on multi-swarm which has stratified mechanisms

Table 1 (continued)

Authors	Year	Dataset	Proposed algorithm	Benefits	Limitation
Thawkar et al. [48]	2021	Database of 651 mammograms obtained from Digital Database for Screening Mammography (DDSM)	Hybrid feature selection method based on butterfly optimization algorithm and ant lion optimizer	Achieves high level of accuracy while using fewest possible features	It failed to show better accuracy on WBC dataset and result based on ANFIS sometimes provide inaccurate prediction as they converge slowly
Afolayan et al. [2]	2022	Wisconsin breast Cancer dataset (WBCD)	PSO optimized using decision tree machine learning technique for breast cancer diagnosis	Efficient and beneficial for decision making	Accuracy of the model is 92.26%, and hence, it cannot be practically used for the diagnosis because of the low accuracy

classifier based on kernel density estimation. This model contained feature selection using PSO and classification with the selected feature subset and optimal kernel bandwidth determination. KDE-PSO model had better average performances than the average performance of the GA-KDE model in breast cancer diagnosis. Though results could show the supremacy of this model over other proven methods, the PSO used for minimizing the classification error and number of features was not for multi-objective function. On the contrary, some researchers have also proposed a TLBO-PSO [43]-based model in their study having a multi-objective optimization with two primary objectives: the first one for attaining the highest classification accuracy and the second to select the least number of informative genes. They used the breast cancer microarray dataset for implementation of their study. Using the proposed TLBO-PSO model, they were able to enhance various performance metrics in the study including sensitivity, accuracy and specificity.

Another study proposed by Wang et al. [53], based on CNN- and CAD-based methods, was proposed in a study that used a dataset of 219 patients with 614 ABUS—automated breast ultrasound—volumes that included 745 cancer regions, and 144 healthy women with a total of 900 volumes, without abnormal findings. In this study, a 3D CNN architecture was proposed for CAD cancer diagnosis in ABUS volumes. The proposed system used the densely deep supervision (DDS) mechanism to learn more discriminative features for malignant representations of cancer cells and to enhance the sensitivity of detection, and utilize the recommended threshold map with a unaccustomed threshold loss to streamline the cancer probability map for the false-positive reduction while maintaining a high record of sensitivity metric. Although the proposed detection system presented astonishing results, it still posed a few limitations. In the model, cystic lesions and fatty lumps would be misclassified as cancerous lesions and small cancerous lesions

may be neglected by considering them a common ultrasound shadow.

Wang et al. [52] proposed the CMWOAFS-SVM model which was an enhanced WOA model based on the control and chaos engineering theories, proposed to optimize two principal parameters for SVM with the RBF kernel and simultaneously select features in practical classification tasks. They tested the prototype on WBCD, diabetes and dermatology datasets. The proposed model consisted of two parts: the first part included feature selection and parameter optimization of SVM that were simultaneously conducted by an improvised WOA algorithm. In the latter part, to evaluate the suitability of SVM for the classification task using a tenfold CV analysis, a subset of the optimal parameter pair and feature was obtained from the initial part. The proposed model could prove supremacy over other models by better performances and feature subset sizes, and the ability to tackle feature selection and parameter optimization simultaneously. Along with these astonishing results, it is not obscured that CMWOA failed to have any advantage in terms of computational cost over the original WOA. Moreover, it also has a complex scavenging process being based on multi-swarm having stratified mechanisms.

Some authors relied more on other datasets than WBCD, seen in studies like one done by Dheeba et. al. [12]. They proposed a PSO optimized WNN model on the real-world clinical database of 216 mammograms collected from mammogram screening centres. In this model, they have attempted to enhance the classification accuracy of WNN by tuning the initial network parameter using PSO. By using PSOWNN, they evolved connected WNN and optimized it by using the best network architecture which was created by optimizing the momentum factor, learning rate and the number of neurons in the hidden layers. The proposed PSOWNN classifier showed improvement in the classification accuracy of CAD analysis of the digital mammogram for diagnosis of breast cancer.

3 Methods and data

This section is divided into four primary subsections: a discussion of the artificial bee colony algorithm, a description of the black hole algorithm, a discussion of genetic operators and a discussion of the WBCD dataset.

3.1 ABC algorithm

Algorithm 1 Pseudo Code for ABC Algorithm

1: **Initialization:**
2:     Population of stars with a random location in the search space using the formula:

$$x_{ij}; \text{ where,} (i = 1, 2, \ldots, SN) \text{ and } (j = 1, 2, \ldots, D)$$

3:     Evaluate the fitness function fitiof all solutions in the population.
4:     Set trial = 0
5: **The Employed Bee Phase**
6: **for** $i = 1 \ TO \ SN \ STEP \ 1$ **do**
7:     Update a new candidate solution v_i

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj})$$

8:     Apply a greedy selection process between v_i and x_i to select the better one
9:     **if** solution x_i does not improves **then**
10:         $trial_i = trial_i + 1$
11:     **else**
12:         else $trial_i = 0$
13:     **end if**
14: **end for**
15: **Calculate the probability** p_i

$$p_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n}$$

16: **The onlooker bee phase**
17: **for** $(i = 1 \ TO \ SN \ STEP \ 1)$ **do**
18:     **if** $rand(0, 1) \leq p_i$ **then**
19:         Update a new candidate solution v_i for the onlookers bees
20:         Evaluate the fitness value of the candidate solution v_i
21:         Apply a greedy selection process between v_i and x_i to select the better one
22:         **if** the solution does not improve **then**
23:             $trial_i = trial_i + 1$
24:         **else**
25:             $trial_i = 0$
26:         **end if**
27:     **end if**
28: **end for**
29: **The scouts bee phase**
30: **if** max $(trial_i) >$ limit **then**
31:     Replace x_i with a new randomly produced candidate solution by using:

$$x_i^j = x_{min}^j + rand(0, 1) \times (x_{max}^j - x_{min}^j)$$

32: **end if**
33: **if** stop criterion is satisfied **then**
34:     Stop and Output the best solution achieved
35: **else**
36:     return to Step 16
37: **end if**

One of many nature-inspired algorithms, "artificial bee colony (ABC)" [24] proposed by Karaboga, is inspired by honeybees' foraging behaviour. The algorithm proposes to see any possible solution to the problem as a food source for the bees, whereas the amount of nectar in the food source corresponds to the solution's quality/fitness obtained. The artificial bee colony consists of three groups, namely employed bees, scouts and onlooker bees. The employed bees exploit local information to extract the locally best possible solution. Once all employed bees are done with the search process, onlooker bees evaluate and choose a food source based on the information supplied by the employed bees. This process is then used until a global best possible solution is found to the optimization problem. The proposed equation, to produce a candidate food (feasible possible solution) source from the information of an old one present in memory, is:

$$v_{ij} = x_{ij} + \Phi_{ij}(x_{ij} - x_{kj}), \tag{1}$$

where $k \in \{1, 2, ..., SN\}$ and $j \in \{1, 2, ..., D\}$ are the randomly chosen indices. Each solution denoted by $x_i (i = 1, 2, \ldots, SN)$ is a D-dimensional vector, where D is the number of optimization parameters. Φ_{ij} is a randomly chosen value lying between $[-1, 1]$, controlling the production of neighbouring food sources around x_{ij} and representing a comparison among two food positions as seen by a bee. A food source is chosen by an artificial onlooker bee based on the probability value attached with that food source, p_i given by Eq. (2).

$$p_i = \text{fit}_i / \sum_{n=1}^{SN} \text{fit}_n, \tag{2}$$

where fit$_i$ is the fitness value of the solution i and SN is the number of food sources which is equal to the number of employed bees.

In this algorithm, scout bees replace an old food source, of which the nectar is rejected by the bees, with a new food source. This process is simulated by the algorithm by generating a random position in dimension space and producing a position at random and replacing that with the rejected food source. In the ABC algorithm, after a prefixed number of cycles, if the quality of food source cannot be improved any further, then that food source is abandoned by the bees. This prefixed number of cycles, called as limit for abandonment, is a very important and sometimes a deciding control parameter in determining the efficiency and accuracy of the algorithm. Taking the abandoned source as x_i, while $j \in \{1, 2, \ldots, D\}$, the scout bees discover a new food source to be replaced with x_i. This operation can be defined by using the following equation:

$$x_i^j = x_{min}^j + \text{rand}(0, 1)\left(x_{max}^j - x_{min}^j\right), \tag{3}$$

The ABC algorithm is tried to be represented with a pseudo-code shown in Algorithm 1.

3.2 Black hole algorithm

Algorithm 2 Pseudo Code Black Hole Algorithm

1: **Initialization:** Population of stars with a random location in the search space using the formula:

$$P(t) = \{x_1^t, x_2^t, x_3^t, \ldots, x_n^t\}$$

2: These randomly generated stars are placed in the search space of function.

3: **for** $i = 1$ to the number of stars **do**

4:     **for** each ith star objective function is evaluated according to the following formula **do**

$$f_i = \sum_{i=1}^{pop\_size} eval(p(t)$$

$$f_{BH} = \sum_{i=1}^{pop\_size} eval(p(t)$$

5:     **end for**

6: **end for**

7: Selecting the best star that has the best fitness value as the black hole

8: **while** max iteration or convergence criteria is not met **do**

9:     **for** $i = 1$ to numbers of stars **do**

$$X_i(t) = X_i(t) + \text{rand} \times (X_{BH} - X_i(t))$$

10:         **if** $f_i > f_{BH}$ **then**

11:             $\text{Swap}(X_i), X_{BH})$

12:         **end if**

13:         Replace the new fitness value of the star (X_i) with the previous value

14:         Update and Calculate :

$$R_{BH} = \frac{f_{BH}}{\sum_{i=1}^{N} f_i}$$

15:         **if** $\sqrt{(X_{BH} - X_i)} < R_{BH}$ **then**

16:             replace X_i with a new star in an optional location in the search scope

17:         **end if**

18:     **end for**

19: **end while**

The black hole algorithm [29, 19, 20] is a bio-inspired population-based method having similar features with another population-based algorithm. This method was inspired by the black hole phenomenon. A black hole is essentially a region of space with so much mass concentrated inside it that if an entity approaches it, there is no way of escaping its gravitational pull. If anything falls into black hole including light, then that object is forever gone from the universe. In the black hole algorithm, we start with an initial population of candidates that are a feasible solution to optimization problems, for every iteration, the best candidate out of the population is selected, and the remaining becomes the normal star after the initialization process the normal star starts moving towards black hole, i.e. best candidate. Whenever a normal star is swallowed by the black hole then for replacing the candidates that have entered the black hole new candidates are generated in the search space.

3.2.1 The absorption rate of stars by a black hole

After the initialization, absorption of the star around the black hole is started and all the normal stars start moving to the black hole. The absorption rate of stars by a black hole is calculated as follows:

$$X_i(t) = X_i(t) + \text{rand} \times (X_{BH} - X_i(t)) \tag{4}$$

where $i=1,2,3,4\ldots n$, X_i^t and X_i^{t+1}, respectively, are the location of ith star at t and $t+1$. X_{BH} is the location of the black hole in the search space and rand is a randomly generated number between the interval of [0,1] and N is the number of normal stars.

3.2.2 Probability of crossing the event horizon during moving star

We use the likelihood of crossing the event horizon (E_H) of the black hole while moving the normal star towards the black hole to obtain the most optimal data from the problem's search space. Every star which crosses the event horizon is absorbed by the black hole that means it dies and a new star is generated in the search space. To calculate the radius of the event horizon, the following equation is used:

$$R_{\text{BH}} = \frac{f_{\text{BH}}}{\sum_{i=1}^{N} f_i} \tag{5}$$

where f_{BH} and f_i are the fitness values of black hole and ith star, respectively, is the number of candidate stars. The pseudo-code for black hole algorithm is shown below in algorithm.

3.3 Genetic algorithms

Genetic algorithms (GAs) are adaptive heuristic search algorithms that are a subset of evolutionary algorithms. They are working on a concept based on natural selection and genetics. The random search is given historical data to direct it into the region of greater performance in the solution space. They are frequently utilized to produce high-quality solutions to optimization and search challenges. GAs work on a population of simulated chromosomes. Typically, these are binary coded. Each chromosome represents a solution to a problem and has a fitness, which is a real number that indicates how good a solution it is to the specific problem. It is a population-based search method that employs the survival of the fittest idea. A genetic algorithm uses operators that instruct the algorithm to solve a particular issue. During the search procedure, the GAs employed a number of operators. Encoding scheme, crossover, mutation and selection are among these operators.

3.3.1 Encoding scheme

In many computational issues, the encoding technique, or the method used to change data into a certain form, is critical. The information provided must be transformed into a certain bit string. The encoding strategy is different depending on the issue domain. They can be used in neural networks to identify the best weights, generate programmes or expressions and so forth. Binary, octal, hexadecimal, permutation, value-based and tree encoding techniques are well known.

3.3.2 Selection techniques

In the genetic algorithm, selection plays a vital role in selecting whether or not a certain string will participate in the reproduction process. It basically chooses parents who will mate and recombine to produce offspring for the future generation. Parental selection is critical to the GA's convergence rate because good parents lead individuals to better and fitter solutions. This operator is also known as the reproduction operator. The rate of convergence of GA is determined by the selection pressure.

3.3.3 Crossover operators

Reproduction and biological crossover are analogous to the crossover operator. In crossover, more than one parent is chosen, and one or more offspring are generated using the parents' genetic material. Crossover is typically used with a high probability in a GA. Single-point, two-point, k-point, uniform, partially matched, order, precedence preserving crossover, shuffle, reduced surrogate and cycle are examples of well-known crossover operators. These crossover operators are fairly generic, and they can also be used to create a problem-specific crossover operator.

For our proposed model, we have used a two-point crossover. The concept of two-point crossover is identical to that of one-point crossover. Instead of a single cut point, two are randomly selected at the same places in two parents, and two offspring are formed from them.

Process for a two-point crossover:

(i) Choose two cut locations at random between each parent's genes.
(ii) Copy the sub-string between the two cut places from Parent 2 to Offspring 1, as indicated in Fig. 1.
(iii) The remaining values are copied from the first parent and placed exactly as they are in Offspring 1, excluding duplicate values. (Genes 1 and 8 have already been inserted in the middle of the Offspring 1 in Fig. 1. As a result, by being copied to it, they are omitted.)
(iv) Empty areas in Offspring 1 are filled with genes that did not occur in the first parent, but did appear in the second parent after the second cut point, as illustrated in Fig. 2. (As gene 4 comes after the second cut point of Parent 2, it appears first in Offspring 1, followed by gene 5.)
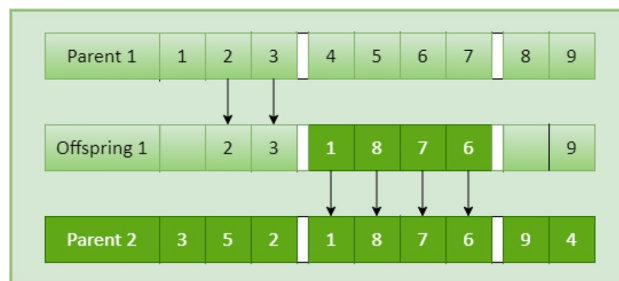

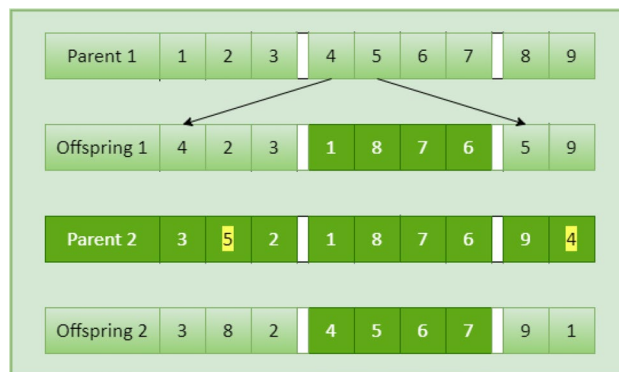**Fig. 1** Diagram representation of two-point crossover

Fig. 2 Diagram representation of two-point crossover

(v) Similarly, Offspring 2 can be formed by swapping the roles of the parents as seen in Fig. 2, which is represented as Offspring 2.

3.3.4 Mutation operators

Mutation is a genetic operator that preserves genetic variation from one population to the next. It can be characterized as a minor random change in the chromosome that results in a new solution. It is typically used with a low probability. The GA is simplified to a random search if the probability is sufficiently high. The "exploration" of the search space is related to mutation. In several circumstances, it has been discovered that mutation is required for GA convergence, whereas crossover is not. Displacement, simple inversion and scramble mutation are well-known mutation operators.

3.4 SVM

SVM is one of the most successful binary classification algorithms due to its excellent generalization capabilities. SVM was developed in the 1990 s by Vapnik et al. [9] with the goal of minimizing structural risk and making the model very resistant to overfitting. If a model classifies testing or unknown data with a lower accuracy than the training data, it is said to be overfit.

By creating a hyperplane or a group of hyperplanes, SVM divides the data into distinct classes. A hyperplane is defined in SVM as a surface that entirely divides data points into two classes so that no data point from one class falls into the other [47].

The hyperplane with the greatest distance to the nearest data points in both classes, out of an unlimited number of potential hyperplanes, is deemed the most optimum [9].

A hyperplane is the set of points **x** satisfying the equation:

$$\mathbf{w} \cdot \mathbf{x} + \mathbf{b} = 0 \tag{6}$$

where b is the bias and **w** is the weight vector which in linear case can be easily solved using Lagrangian multipliers.

SVM may efficiently classify high-dimensional data by mapping it onto a high-dimensional feature space such that the mapped data can be readily separated using lower-order polynomial functions, despite its origins in linearly separable situations. A kernel function is used to map original data into a high-dimensional feature space [5].

The following common kernels are used in SVM classifier:

1. Linear kernel:

$$f(\alpha_i, \alpha_j) = \alpha_i \cdot \alpha_j \tag{7}$$

2. Polynomial kernel:

$$f(\alpha_i, \alpha_j) = (\beta \alpha^t \alpha_j + r^2)^2 \tag{8}$$

3. RBF kernel:

$$f(\alpha_i, \alpha_j) = e^{[\gamma ||\alpha_i - \alpha_j||^2]} \tag{9}$$

4. Sigmoid kernel:

$$f(\alpha_i, \alpha_j) = \tan h(\beta \alpha_i^t \alpha_j + r) \tag{10}$$

3.5 WBCD dataset

Breast cancer is one of the major reasons for the death of women around the world. It is an absolute urgency of finding treatment solutions and some accurate diagnosing. To deal with this problem, the raw breast cancer data are taken from Wisconsin Diagnosis Breast Cancer dataset from UCI machine learning repository. Features in the dataset are computed from a digitized image of a fine needle aspirate (FNA) of a breast mass. They describe the characteristics of the cell nuclei present in the image [14].

These data are preprocessed by replacing the missing values with 0 instead of removing them as it should not introduce a noticeable impact on statistical data patterns of the dataset. The dataset is too small for our use after dividing the dataset into training and testing subsets. Another common issue with the dataset is skewness. So, in order to resolve the small dataset size issues and the data unbalance new data are generated of the same dimensions by introducing normal distributed random noise. Almost 75% of data are considered for training and 25% for testing. A piece of detailed information about the WBCD dataset is presented in Table 2.
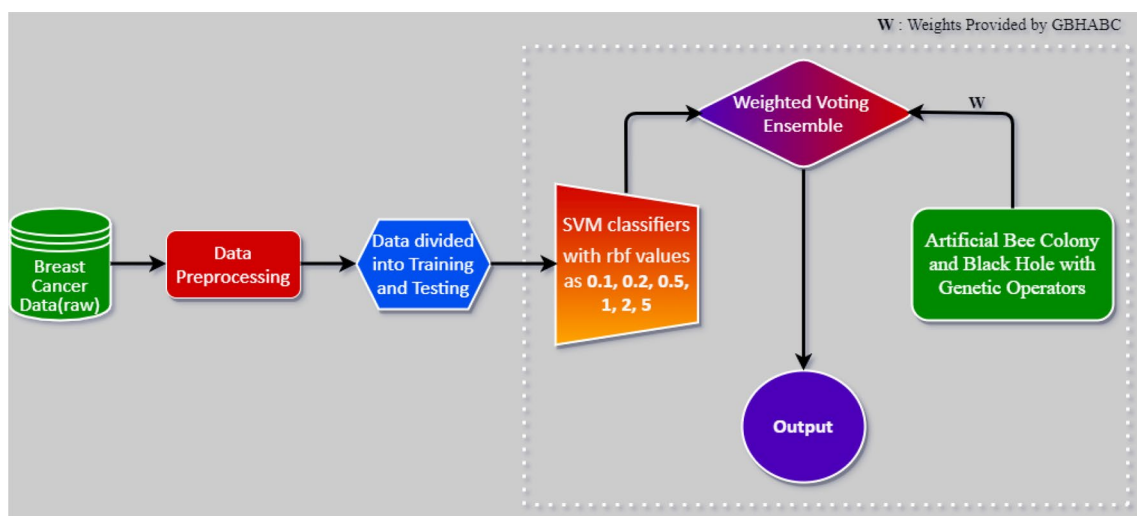
4 Proposed model

ABC faces a notable challenge in the form of stagnation, where the exploration phase experiences a gradual progression in the initial stages due to the evolving nature of the location. This indicates that the exploration process encounters difficulties in advancing at a desirable pace during

Table 2 Description of the WBCD dataset

Attribute number	Attribute description	Minimum value	Maximum value	Standard deviation	Mean value	Total count
1	Clump thickness	1	10	2.82	4.42	699
2	Uniformity of cell size	1	10	3.05	3.13	699
3	Uniformity of cell shape	1	10	2.97	3.21	699
4	Marginal adhesion	1	10	2.85	2.81	699
5	Single epithelial cell size	1	10	2.21	3.22	699
6	Bare nuclei	1	10	3.64	3.46	699
7	Bland chromatin	1	10	2.44	3.44	699
8	Normal nucleoli	1	10	3.05	2.87	699
9	Mitoses	1	10	1.71	1.59	699

the early phases of ABC. Finding an appropriate balance between the exploitation and exploration stages is one of the main difficulties addressed by the basic black hole algorithm. The algorithm frequently runs into problems because it tends to become stuck in local optima, which makes it difficult for it to adequately traverse the search space, which subsequently results in sub-optimal solutions. Another property of black hole algorithm is that when a star crosses the event horizon it gets destroyed. To keep the number of data points same as initial it generates new points randomly. Some areas of the search space may become densely inhabited as a result of the algorithm's random creation of points, while other areas may remain mostly unexplored. As a result, the algorithm can find it difficult to fully explore the whole search area, which could result in biased findings or a failure to find the best answers. The algorithm's efficiency and efficacy may be hampered by the uneven distribution produced by the random point creation, making it more difficult for it to successfully navigate the search environment. All these shortcomings led to the development of the proposed algorithm.

The architecture of the proposed GBHABC (hybrid artificial bee colony and black hole with genetic operators) model is given in Fig. 3. The pseudo-code for GBHABC is illustrated in Algorithm 3. In the first step, a support vector machine (SVM) is applied. SVM gives the best outcome when we use the RBF kernel function for classification. Different kernel functions can be specified for the decision function. Common kernels are provided, but we can also specify custom kernels. We are training six different kinds of SVM experts with varied values of RBF parameters. Values designated to the RBF variable for different SVM classifiers are [0.1, 0.2, 0.5, 1, 2, 5]. Later, in the result section, it has been shown that based on the selected RBF value, the performance measures of each SVM classifier differ. In the second step, the outcomes of the SVM expert system are ensemble using weighted majority voting. Littlestone and Warmuth [34] showed the minimization of errors made by the ensemble system, by providing weights to the majority voting procedure. Based on their classification accuracy, all the experts are provided with some weights.

Fig. 3 Architecture of proposed GBHABC

Algorithm 3 Pseudo Code For Proposed Model

1: **Initializing Parameters:** m = Maximum Number of Cycles, d = Dimension of the Problem, α = Swarm Size, l = Lower Bound, μ = Upper Bound, n, a;
2: Initialise the swarm having food sources/stars, x_i where $(i = 1, 2, \ldots, \alpha)$ by using $Eq.(1)$;
3: $t=1$;
4: **while** $t < m$ **do**
5:     Employed Bee Segment:
6:     Evaluate fit_i for each x_i
7:     Select the best star from the population as BH(black hole)
8:     **for** each x_i **do**
9:         Move the star toward the BH using $Eq.(2)$
10:         Re-evaluate fit_i
11:         **if** $(fit_i < fit_{BH})$ **then**
12:             Set x_i as BH
13:         **end if**
14:         Calculate the RBH(event horizon radius) using $Eq.(3)$
15:         **if** $(t/m * 10\%2) = 0$ **then**
16:             $e_r = 0.65$
17:         **else**
18:             $e_r = 0.35$
19:         **end if**
20:     **end for**
21:     **for** each x_i **do**
22:         **if** $(x_i - x_{BH}) < R_{BH}$ **then**
23:             Remove x_i
24:             **if** $r < e_r$ **then**
25:
26:                 r: random number in the uniform range of $[0, 1]$
27:                 Perform two-point crossover to generate new star
28:             **else**
29:                 Generate the new star randomly using $Eq.(1)$
30:                 Add the generated star to the population
31:             **end if**
32:         **end if**
33:         **if** $fit_i < fit_{BH}$ **then**
34:             Set x_i as BH
35:         **end if**
36:     **end for**
37:     Onlooker Bee Segment:
38:     Scout Bee Segment:
39:     $t = t + 1$
40: **end while**
41: return x_{BH}

We have proposed GBHABC model to generate weights for the six (SVM) experts for weighted majority voting. ABC has the limitation of being stagnant due to which the ABC algorithm gets stuck to the local optima or converges prematurely. So, black hole is introduced in order to make ABC more explorative. To create a simulation of the black hole phenomena, all candidate solutions are chosen as stars, and the solution with the highest fitness among the candidates is chosen as the black hole. The flowchart of our proposed algorithm is represented in Fig. 4.

4.1 Initialization and updation of positions

All the solutions are initialized in a random manner in the search space as shown in Eq. (11).

$$x_{ij} = x_{lb} + \text{rand}[0, 1](x_{ub} - x_{lb}) \tag{11}$$

where x_{ij} represents the ith food source in the populous in $j$thdimension. x_{ub} and x_{lb} are the upper bound and the lower bound, respectively, for x_{ij} and rand$[0, 1]$ is a uniformly scattered arbitrary number in the bounds $[0, 1]$.

After initializing the BH (black hole) and the stars, the updation of positions by the solutions (stars) is done using the direction and the distance of the BH. The position of the stars is updated by the equation expressed by Eq. (12).

$$x_i(t + 1) = x_i(t) + \text{rand}(x_{BH} - x_i(t)) \tag{12}$$

where $x_i(t + 1)$ and $x_i(t)$ represent the position of the ith solution during iteration $t + 1$ and t, respectively, x_{BH} is the position of the black hole (best solution) in the search space and rand$[0, 1]$ is a uniformly scattered arbitrary number in the bounds $[0, 1]$.

If a solution (star) achieves a better position than the black hole, their positions are swapped, and a new star is added as the BH of the search space. Furthermore, if a star advances towards the BH, it has the potential to cross the BH's event horizon (EH). Crossing the horizon signifies that the star will be absorbed in the BH, i.e. eliminated, and a new star will be generated, i.e. a new solution in the search space will be produced.

4.2 Generation of a new star

When a star is absorbed by the black hole, a new star is created using one of the following strategies:

1. A new star is created at random with a probability of 0.65 (by using uniform distribution).
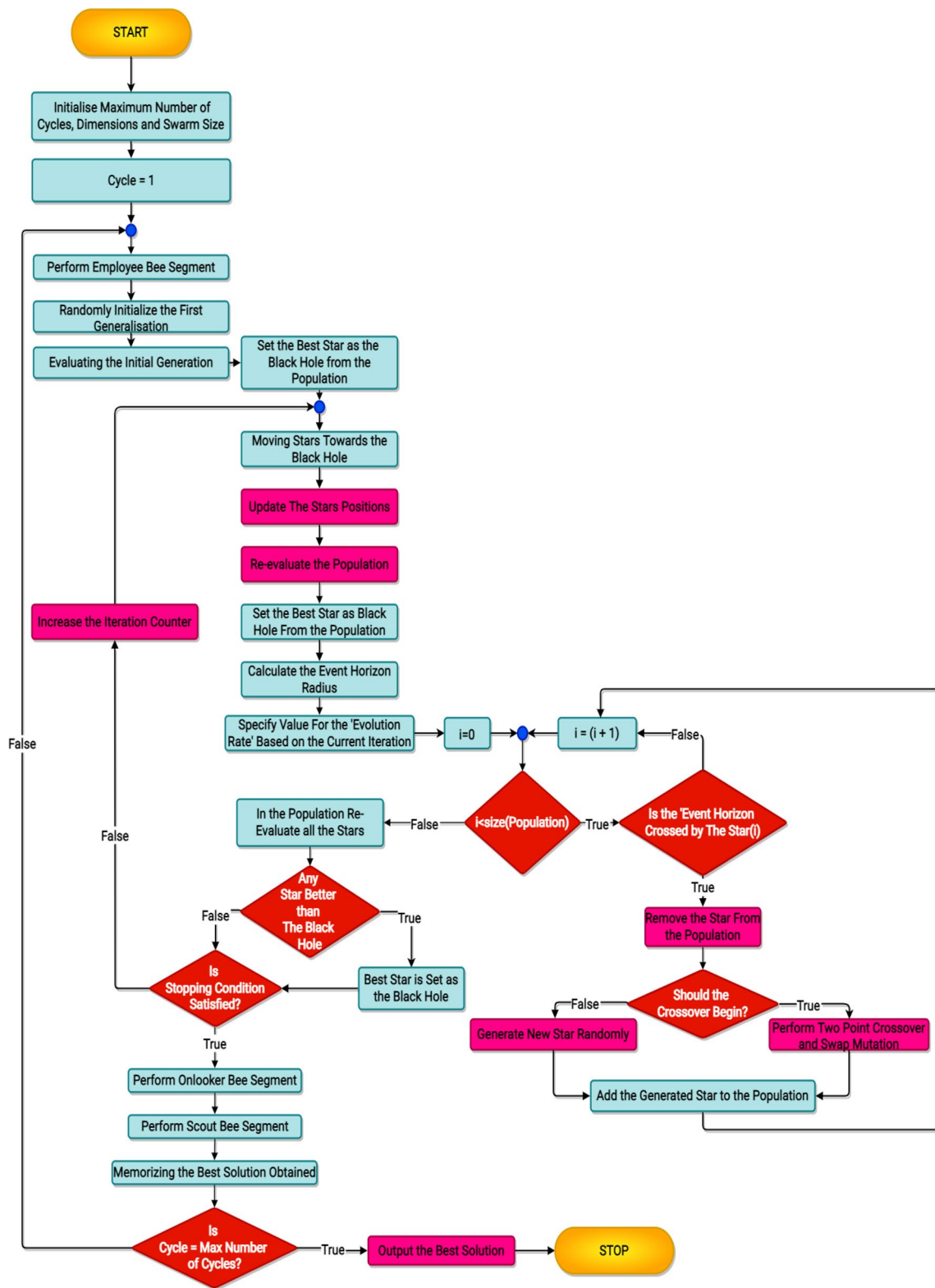
Fig. 4 Flowchart of the proposed GBHABC

2. A new star is created by selecting two stars from the population and recombining their attributes using a recombination procedure. The best of them will be added to the population later on. This approach is used with a probability of 0.35 at first.

The probabilities for regeneration [0.65, 0.35] change during the progress of the model. For instance, using the 1st strategy the model starts with a 0.65 probability of generating stars. When the ratio of the iterations to the total number of iterations reaches 0.2, BH changes the probability for generating stars using the 2nd strategy to be 0.65 and sets the probability to be 0.35 for the 1st strategy. It will alternate again when the ratio reaches 0.4. The alternation of probabilities will continue for each 20% progress of the iterations until the stop condition is satisfied. By implementing the above strategies, we can alternate the focus on exploitation and exploration during model progression and achieve the best balance between them in order to find the optimal solution.

4.3 Recombination method for generation of a new star

According to the fitness value, the current population is sorted and divided into 2 subpopulations, i.e. p_{male} and p_{female}. Randomly one star is selected from both, and a two-point crossover is applied over the selected stars. Then, among the two resulting stars created by crossover, the best one is selected. Finally, swap mutation is performed on the newly generated star to have diversity and even distribution in the search space. This star is added to search space which makes the search space better and proposed model efficient.

Then the subsequent phases of the model follow the same steps as ABC. Similar to ABC, GBHABC is also divided into three segments which are improved employed bee segments where the BH phenomenon with the genetic operators is applied, the onlooker bee segment and the scout bee segment. All other phases except the first one are kept similar as in the original ABC algorithm.

5 Experimentations and results

This section compares and contrasts the findings achieved using the suggested model and those obtained without it. First, there will be a quick review of the assessment measures. As previously stated, our research consists of two key experiments. The outcomes of traditional approaches will be given and debated in the first experiment. The outcomes

of the suggested model are then presented and discussed in the second experiment.

5.1 Evaluation metrics

One of the most crucial steps is the comparison of a model by evaluation metrics. By utilizing and studying these metrics, we can analyse the performance of the proposed model and compare it with the existing methods. For this regard, 8 metrics have been considered, which are sensitivity, precision, specificity, $F1$ score, accuracy, FPR, FNR, MCC computed using the true positive (TP), false negative (FN), false positive (FP) and true negative (TN) rates. The confusion matrix was created for different RBF values.

As mentioned above, 8 metrics have been used to compute the performance of various RBF kernels of SVM. The different values taken for RBF in SVM were 0.1,0.2,0.5,1.0,2.0 and 5.0. They were calculated using Eqs. (13)–(20):

$$\text{Specificity} = \text{TNR} = \frac{\text{TN}}{(\text{TN} + \text{FP})} \tag{13}$$

$$\text{Sensitivity} = \text{TPR} = \frac{\text{TP}}{(\text{TP} + \text{TN})} \tag{14}$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{15}$$

$$F1 = \frac{2\text{TP}}{(2\text{TP} + \text{FP} + \text{FN})} \tag{16}$$

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{17}$$

$$\text{FPR} = \frac{\text{FP}}{(\text{FP} + \text{TN})} \tag{18}$$

$$\text{FNR} = \frac{\text{FN}}{(\text{FN} + \text{TP})} \tag{19}$$

$$\text{MCC} = \frac{(\text{TP} \cdot \text{TN}) - (\text{FP} \cdot \text{FN})}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \tag{20}$$

5.2 First experiment

This experiment discusses the results obtained by using different RBF values for SVM algorithm. As previously mentioned in our first experiment, 699 records of whole WBCD dataset

Table 3 Confusion matrix for various SVM RBF values

Algorithm	TP	FP	FN	TN
SVM 0.1	61	0	5	109
SVM 0.2	61	0	8	106
SVM 0.5	61	0	15	99
SVM 1.0	61	0	24	90
SVM 2.0	61	0	30	84
SVM 5.0	61	0	52	62

Table 5 Parameter settings of the proposed GBHABC algorithm

Parameters	Settings
Maximum iterations (Max_iter)	50
Lower bound (lb)	− 1000
Upper bound (ub)	1000
Dimensions (dim)	12
SearchAgents_no	5
npop	10
A	1

were used. The confusion matrices are shown in Table 3 that shows the results for different RBF values of SVM. The table shows the confusion matrix of the testing set used for this experiment. 25 per cent of the dataset is used for this step to verify the feasibility of SVMs in the current study. It can be observed that in the best case, as shown in Table 3, 61 malignant records were detected as malignant and were classified as TP, and in the best case, 109 benign records were detected as benign and were classified as TN. Moreover, in the best case, 0 benign records were classified as malignant and 52 malignant records were classified in the benign class. The calculated $F1$ score, specificity, sensitivity, precision and accuracy values are illustrated in Table 4, for training steps.

According to Table 4, the evaluation metrics with the greatest values were produced in SVM with RBF value 0.1. It was able to obtain a 97.14 per cent accuracy rate. It can also be observed that the different RBF values tested were 0.1, 0.2, 0.5, 1.0, 2.0 and 5.0. There are two important factors in RBF SVM that determine the final results: gamma(γ) and C. The C parameter strikes a compromise between the proper classification of training examples and the maximization of the decision function's margin. If the decision function is better at reliably categorizing all training points, a narrower margin will be acceptable for higher values of C. A lower C encourages a larger margin and, as a result, a simpler decision function; however, this comes at the cost of training accuracy. This means that the C parameter functions as a model regularizer. The gamma(γ) parameter controls the impact of individual training samples on the overall model, which influences the final output. The gamma(γ) parameter has a big influence on how the model behaves. If gamma is set too high, the radius of the support vectors' sphere of action is reduced to the support vector itself, and no amount

of regularization using C can avoid overfitting. To make the procedure efficient and less difficult, just the gamma(γ) parameter was adjusted for multiple SVM kernels based on these data. The parameters used during the experiment are listed with their values in Table 5.

Figure 5a–d shows the comparison between above-discussed SVMs' according to different evaluation metrics. Figure 5a compares all the kernels according to their accuracy, data are given in Table 4. Similarly, Fig. 5b–d compares specificity, sensitivity and $F1$ score, respectively. These graphs confirm that changing gamma from 0.1 to 5.0 affects the outcomes of the models. Table 4 shows that when the value of gamma(γ) is increased from 0.1 to 5.0, the accuracy reduces from 97.14 per cent to 70.29 per cent. Similarly, sensitivity and f1 score also decrease with increasing the gamma(γ) value. On a fundamental level higher value of gamma(γ) makes the model more forgiving, but trades off the accuracy and other metrics. On the other hand a lower value of gamma(γ) is more stern on the search space selection and has higher accuracy. Both higher value and lower value of gamma(γ) has their benefits and limitations. So in the next experiment we utilize the maximum potential of these different kernels by using them in cohesion with each other with the help of reliable ensemble learning and our proposed hybrid model to improve the results further.

A little discussion regarding the relevance of the features of the dataset is furthered. According to Table 6, the most essential aspects for identifying BC were bare nuclei, uniformity of cell shape, normal nucleoli, uniformity of cell size and bland chromatin, whereas the least important criteria for identifying BC were single epithelial cell size and

Table 4 Evaluation metrics of SVM

Algorithm	Accuracy	Specificity	Sensitivity	Precision	$F1$ Score	FPR	FNR	MCC
SVM 0.1	97.14	100.0	92.42	100.0	96.06	0.00	0.08	0.94
SVM 0.2	95.43	100.0	88.41	100.0	93.85	0.00	0.12	0.91
SVM 0.5	91.43	100.0	80.26	100.0	89.05	0.00	0.20	0.83
SVM 1.0	86.29	100.0	71.76	100.0	83.56	0.00	0.28	0.75
SVM 2.0	82.86	100.0	67.03	100.0	80.26	0.00	0.33	0.70
SVM 5.0	70.29	100.0	53.98	100.0	70.11	0.00	0.46	0.54

(a) SVM different rbf accuracy comparison

(b) SVM different rbf specificity comparison

(c) SVM different rbf sensitivity comparison

(d) SVM different rbf F-1 score comparison
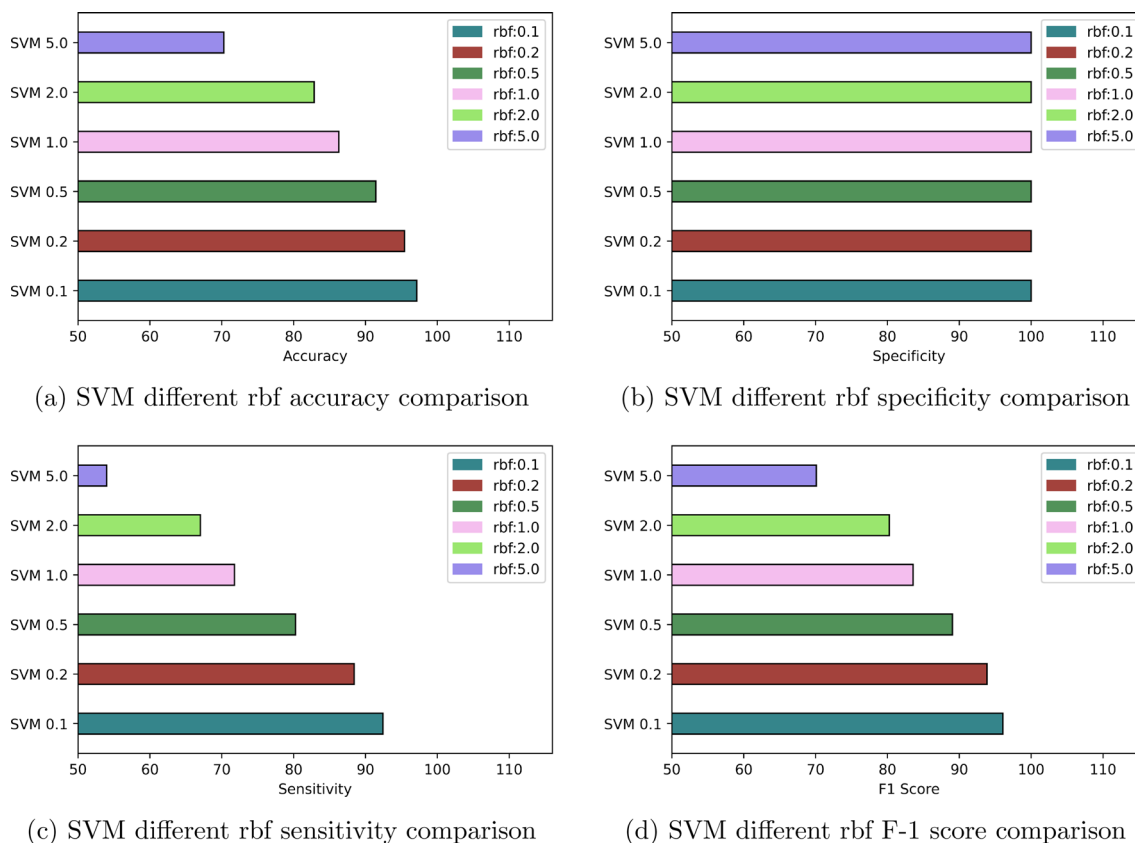
Fig. 5 SVM performance metrics

Table 6 Coefficient for each attribute

No	Features name	Coefficient for attributes
1	Bare nuclei	0.28
2	Uniformity of cell shape	0.23
3	Normal nucleoli	0.14
4	Uniformity of cell size	0.12
5	Bland chromatin	0.10
6	Mitoses	0.07
7	Clump thickness	0.05
8	Single epithelial cell size	0.02
9	Marginal adhesion	0.00

marginal adhesion. Table 6 shows the relative relevance of all aspects using the coefficient characteristics. For example, the most crucial trait for identifying breast cancer is the coefficient value of Bare nuclei, which is 0.28. With a value of 0.00, the coefficient value of marginal adhesion has the lowest relevance of all the attributes. These discoveries might be extremely beneficial to doctors who treat patients with

breast cancer. In other words, taking into account the most essential characteristics can increase the quality of breast cancer diagnosis while lowering the mistake rate.

5.3 Second experiment

The next phase of our study was to implement our proposed model to optimize the weights of an ensemble model which uses different SVM kernels discussed in experiment 1. As discussed before, SVM kernels with varying gamma (γ) values have different results and benefits. So, we propose that we use all these different kernels in a weighted voting ensemble model with weights of the model being optimized by our proposed GBHABC model.

Finally we evaluated our model on the same metrics on which individual SVM kernels were evaluated. For even more clarity, results of this experiment are on a larger pool of dataset than the previous experiment, but the general distribution of the dataset and the dataset itself were same. Our model achieved an accuracy of 99.42%, sensitivity of 98.41%, specificity of 100.00%, precision of 100.00% and F1 score of 99.20%. These results show that our proposed
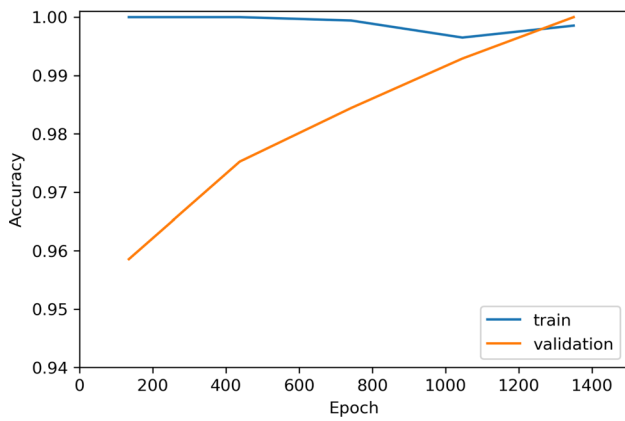
Fig. 6 Accuracy vs epochs for training and validation steps

Table 7 Dataset distribution

Dataset distribution	Accuracy	Specificity	Sensitivity	Precision	*F*1 score
50–50	98.78	98.58	98.04	97.39	98.26
60–40	98.93	98.80	98.09	97.78	98.46
75–25	99.42	100.00	98.41	100.00	99.20

Table 8 Evaluation of our model

Metric	Value (%)
Accuracy	99.42
Specificity	100.00
Sensitivity	98.41
Precision	100.00
*F*1 score	99.20
FPR	0.00
FNR	0.02
MCC	0.99

Table 9 Confusion matrix of our proposed GBHABC

		Predicted class	
		Malignant	Benign
	Malignant	124	0
Actual class	Benign	2	224

approach has improved the overall results of the model on the Wisconsin dataset.

Figure 6 shows the accuracy vs epochs comparison for training and validation steps. The y-axis represents accuracy in the range from [0, 1] which on multiplication with 100 can easily be converted to percentage measure. The x-axis represents the number of epochs. Training step is denoted with a blue line and the validation step is denoted with an orange line. As observed, the training accuracy remains almost constant throughout all the epochs, only dipping by a mere 0.57% during the entire process. Whereas the validation accuracy improves consistently from a value around 95% to 99% and during the process, it becomes equal to the training accuracy and finally settles down. This graph proves that our model handles overfitting really well and is generalized for more data points rather than just being able to classify only the training examples.

We performed our experiment on different dataset distributions. The distributions were 50–50, 60–40, 75–25 as training and test sets. We calculated accuracy, sensitivity, specificity, precision and *F*1 score for the above-mentioned dataset distributions. The results of the experiment with different dataset distributions are given in the Table 7. As observed from the table, the accuracy and other evaluation metrics were very much close to each other, and thus, it can be inferred that the model is stable for different splits of the data.

Equations (13)–(20) were used for final evaluation of our model. The final results for the proposed hybrid GBHABC on the basis of accuracy, specificity, sensitivity, precision, *F*-1 score, FPR, FNR and MCC are 99.42%, 100.00%, 98.41%, 100.00%, 99.20%, 0.00, 0.02 and 0.99, respectively, and the same is listed in Table 8. These results signify that the proposed model is superior to the classical SVM models. The confusion matrix for the given results is shown in

Table 9. Figure 7 shows the visual representation of the obtained results for better and easy understanding. All these results imply that the proposed GBHABC model had a positive impact on the detection of breast cancer by improving the ability to correctly classify all the cases.
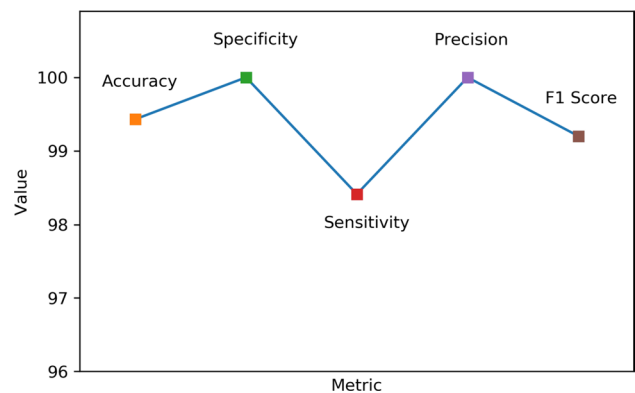

Fig. 7 Metrics representation

Table 10 Comparison of different methods applied for breast cancer prediction

S. No	Study	Year	Algorithm	Mean accuracy
1	Liu and Li [32]	2010	SVM-CPBK	93.26
2	Senapati and Dash [42]	2013	LLWNN	98.14
3	Zheng et al. [57]	2014	K-means with SVM	97.38
4	Nguyen et al. [39]	2015	GSAM wavelets	97.40
5	Murat Karabatak [23]	2015	Weighted NB	98.54
6	Sheikhpour et al. [44]	2016	PSO-KDE	98.45
7	Dora et al. [13]	2017	GNRBA	98.46
8	Zarbakhsh et al. [56]	2017	ANFIS with AR	99.26
8	Wang et. al. [51]	2018	WAUCE	97.10
9	Wuniri et al. [54]	2019	BIG-F	97.10
10	Punitha et al. [40]	2019	IABC-EMBOT	97.53
11	Abdar et al. [1]	2020	NE	98.07
12	Thawkar et al. [48]	2021	BOAALO	98.16
13	Afolayan et al. [2]	2022	PSO-DT	92.26
14	**Proposed model**	**2023**	**GBHABC**	**99.42**

Bold indicates our proposed model
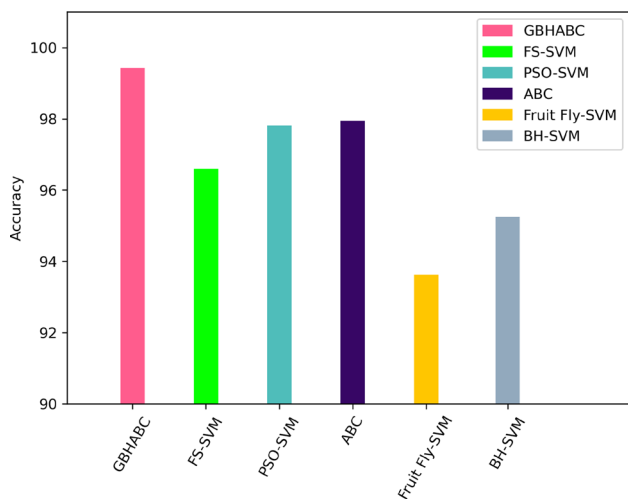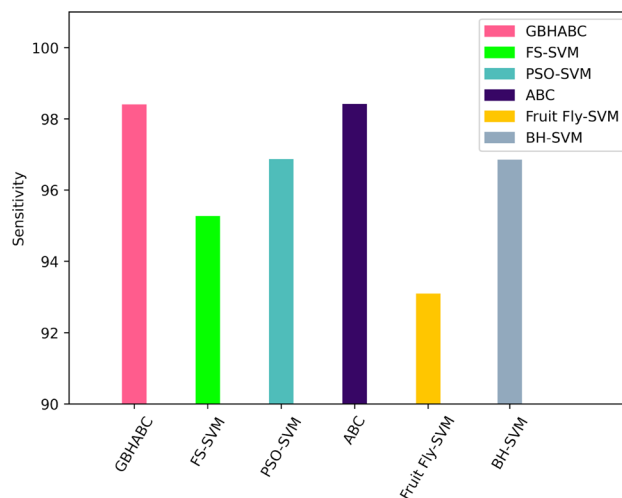

Fig. 8 Proposed vs old accuracy comparison

Fig. 9 Proposed vs old sensitivity comparison

5.4 Comparison with previous studies

We are going to expand on matters a bit by mentioning that we compared the findings obtained by utilizing the proposed methodology, with those that were already reported in the prior study. These studies had differing results, but all focused on the prediction of breast cancer incidence. Table 10 given in our study is used to represent the present levels of accuracy in our model's development, compared to those of other models built in the past.

Some of the models were able to achieve significantly higher accuracy on the breast cancer detection study.

Sheikhpour et al. [44] proposed an approach by applying particle swarm optimization (PSO) with nonparametric KDE (kernel density estimation) and achieved an accuracy of 98.45%. Abdar et al. [1] achieved a analogous accuracy of 98.07% by implementing a branched ensemble approach that utilized the stacking and voting techniques among the classifiers. Similarly, Nahato et al. [36], Zarbakhsh et al [56], Punitha et al [40] achieved 98.6%, 99.26% and 97.53% accuracy, respectively.

Figures 8, 9, 10 shows the comparison between the proposed model and some of the previous studies with the help of graphs that display a cross-sectional comparison in the form of a bar graph with different hybrid algorithms that
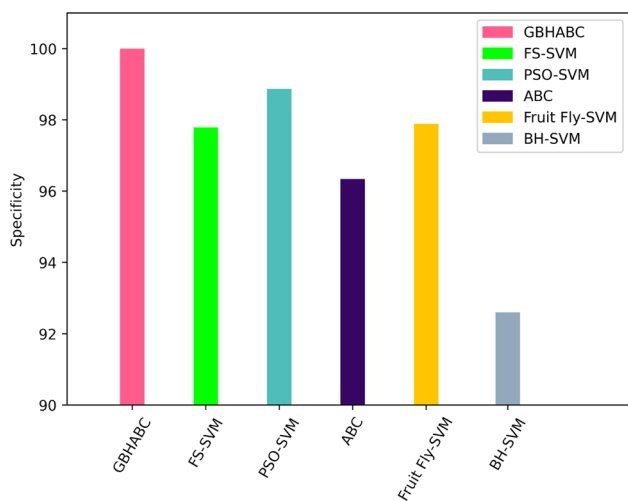
Fig. 10 Proposed vs old specificity comparison

6 Conclusion and future scope

According to the findings of the present study, we have suggested the GBHABC (artificial bee colony with black hole and genetic operators) model for the detection of breast cancer by distinguishing tumours as malignant or benign at an early stage of development. To do this, we used the well-known WBCD dataset from the University of California, Irvine dataset repository to get raw breast cancer data. The proposed GBHABC-SVM model consists of six SVM classifiers (differentiated by the RBF parameter), which are combined into a weighted voting ensemble to improve performance.

It consisted of two primary experiments. The SVM method with three major parameters, C, ϵ and γ, and was used in the first experiment. Here, a basic SVM model, a polynomial SVM (with default parameters) and multiple modified SVMs with different C, ϵ, γ and values were compared. Our findings show that optimizing these parameters improves SVM performance. We utilized the first experiment's optimal parameters of C, ϵ, γ and in subsequent studies. As part of the second experiment, we combine all of these distinct kernels into a weighted voting ensemble model, with the weights of the model being optimized using our suggested model. A new model was introduced under the name of GBHABC. Our final findings show that the suggested technique allows us to achieve an accuracy, precision, sensitivity and f1-score of 99.42%, 100%, 98.41% and 99.20% and, as a consequence, to be highly successful in terms of predicting breast cancer. In conclusion, we believe the GBHABC model should be used in medical research for breast cancer prediction.

It would be intriguing to do a preliminary investigation of the model on various kinds of medical data in the future. In future, the integration of this hybrid model into multi-class classification is extendable to other datasets and various types of cancer detection. Furthermore, alternative machine learning models can leverage diverse forms of data, such as images and X-ray signals, for cancer detection. Moreover, future research endeavours may also encompass the examination of early cancer detection and the anticipation of cancer staging and can be extended to other diseases like brain tumour, glaucoma and diabetes detection.

existed in the past. A tabular representation of the comparison on the evaluation metrics like accuracy, sensitivity and specificity of the proposed model GBHABC has been done with previously existing hybrid algorithms on the WBCD breast cancer dataset as shown in Table 11. Various hybrid algorithms like FS-SVM, PSO-SVM, ABC, BH-SVM and fruit fly-SVM were taken into account for the comparison with GBHABC. All of these algorithms were run on our preprocessed WBCD dataset for comparison with the proposed research, and the results were analysed. Among the aforementioned algorithms, FS-SVM had an accuracy of 96.60%, specificity of 95.27% and sensitivity of 97.79%. PSO-SVM had an accuracy of 97.81%, specificity of 98.87% and sensitivity of 96.87%. ABC had an accuracy of 97.95%, specificity of 96.34% and sensitivity of 98.42%. BH-SVM had an accuracy of 95.25%, specificity of 92.60% and sensitivity of 96.85%. Similarly fruit fly-SVM had an accuracy of 93.62%, specificity of 96.53% and sensitivity of 93.10%. Our model GBHABC has an accuracy of 99.42%, a sensitivity of 100.00% and a specificity of 98.41%. GBHABC definitely appears to be superior when compared to the other tests, according to its accuracy, sensitivity and specificity. Thus, we can conclude that the proposed model GBHABC is quite effective for an early diagnosis of breast cancer and better than already present methodologies.

Table 11 Model comparison

	FS-SVM	PSO-SVM	ABC	BH-SVM	Fruit fly-SVM	GBHABC (our model)
Accuracy (%)	96.60	97.81	97.95	95.25	93.62	99.42
Sensitivity (%)	95.27	96.87	98.42	96.85	93.10	98.41
Specificity (%)	97.79	98.87	96.34	92.60	96.53	100.00

Declarations

Conflict of interest The authors declare that we have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

Informed consent This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Abdar M, Zomorodi-Moghadam M, Zhou X, Gururajan R, Tao X, Barua PD, Gururajan R (2020) A new nested ensemble technique for automated diagnosis of breast cancer. Pattern Recogn Lett 132:123–131. https://doi.org/10.1016/j.patrec.2018.11.004
2. Afolayan J, Adebiyi M, Arowolo M, Chakraborty A (2022, 06). Breast cancer detection using particle swarm optimization and decision tree machine learning technique. 61-83. https://doi.org/10.1007/978-981-16-8150-94
3. Barzaman K, Karami J, Zarei Z, Hosseinzadeh A, Kazemi MH, Moradi- Kalbolandi S, Farahmand L (2020) Breast cancer: biology, biomarkers, and treatments. Int Immunopharmacol 84:106535. https://doi.org/10.1016/j.intimp.2020.106535
4. Bhardwaj A, Tiwari A (2015) Breast cancer diagnosis using genetically optimized neural network model. Expert Syst Appl 42(10):4611–4620. https://doi.org/10.1016/j.eswa.2015.01.065
5. Boser BE, Guyon IM, Vapnik VN (1992) Training algorithm for optimal margin classifiers. In: Proceedings of the fifth annual acm workshop on computational learning theory, pp 144–152. Publ by ACM. https://doi.org/10.1145/130385.130401
6. Chen HL, Yang B, Liu J, Liu DY (2011) A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. Expert Syst Appl 38(7):9014–9022. https://doi.org/10.1016/j.eswa.2011.01.120
7. Chen HL, Yang B, Wang G, Wang SJ, Liu J, Liu DY (2012) Support vector machine based diagnostic system for breast cancer using swarm intelligence. J Med Syst 36(4):2505–2519. https://doi.org/10.1016/s10916-011-9723-0
8. Chen MJ, Wu WY-Y, Yen AM-F, Fann JC-Y, Chen SL, Chiu SY-H, Chiou ST (2016) Body mass index and breast cancer: analysis of a nation-wide population-based prospective cohort study on 1 393 985 Taiwanese women. Int J Obes 40(3):524–530. https://doi.org/10.1038/ijo.2015.205
9. Cortes C, Vapnik V (1995) Supportvector networks. Mach Learn 20(3):273–297. https://doi.org/10.1007/bf00994018
10. de Boer MC, Wörner EA, Verlaan D, van Leeuwen PA (2017) The mechanisms and effects of physical activity on breast cancer. Clin Breast Cancer 17(4):272–278. https://doi.org/10.1016/j.clbc.2017.01.006
11. Deeb H, Sarangi A, Mishra D, Sarangi SK (2021) Improved Black Hole optimization algorithm for data clustering. J King Saud Univ Comput Inf Sci. https://doi.org/10.1016/j.jksuci.2020.12.013
12. Dheeba J, Albert Singh N, Tamil Selvi S (2014) Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach. J Biomed Inform 49:45–52. https://doi.org/10.1016/j.jbi.2014.01.010
13. Dora L, Agrawal S, Panda R, Abraham A (2017) Optimal breast cancer classification using Gauss-Newton representation based algorithm. Expert Syst Appl 85:134–145. https://doi.org/10.1016/j.eswa.2017.05.035
14. Dr. WIlliam H Wolberg (1992) UCI machine learning repository: breast cancer wisconsin (Original) Data Set
15. Gaitanidis A, Alevizakos M, Tsalikidis C, Tsaroucha A, Simopoulos C, Pitiakoudis M (2018) Refusal of cancer-directed surgery by breast cancer patients: risk factors and survival outcomes. Clin Breast Cancer 18(4):e469–e476. https://doi.org/10.1016/j.clbc.2017.07.010
16. Gerstl B, Sullivan E, Ives A, Saunders C, Wand H, Anazodo A (2018) Pregnancy outcomes after a breast cancer diagnosis: a systematic review and meta analysis. Clin Breast Cancer 18(1):e79–e88. https://doi.org/10.1016/j.clbc.2017.06.016
17. Goff SL, Danforth DN (2021) The role of immune cells in breast tissue and immunotherapy for the treatment of breast cancer. Clin Breast Cancer 21(1):e63–e73. https://doi.org/10.1016/j.clbc.2020.06.011
18. Gu J, Groot G, Boden C, Busch A, Holtslander L, Lim H (2018) Review of factors influencing women's choice of mastectomy versus breast conserving therapy in early stage breast cancer: a systematic review. Clin Breast Cancer 18(4):e539–e554. https://doi.org/10.1016/j.clbc.2017.12.013
19. Hatamlou A (2013) Black hole: a new heuristic optimization approach for data clustering. Inf Sci 222:175–184. https://doi.org/10.1016/j.ins.2012.08.023
20. Hatamlou A (2018) Solving travelling salesman problem using black hole algorithm. Soft Comput 22(24):8167–8175. https://doi.org/10.1007/s00500-017-2760-y
21. Jayachitra S, Prasanth A (2020) Multi-feature analysis for automated brain stroke classification using weighted Gaussian Naive Baye's classifier. J Circuits Syst Comput 30:2150178. https://doi.org/10.1142/S0218126621501784
22. Juanjuan L, Santa-Maria CA, Hongfang F, Lingcheng W, Pengcheng Z, Yuanbing X, Chuang C (2020) Patientreported outcomes of patients with breast cancer during the COVID-19 Outbreak in the epicenter of China: a cross-sectional survey study. Clin Breast Cancer 20(5):e651–e662. https://doi.org/10.1016/j.clbc.2020.06.003
23. Karabatak M (2015) A new classifier for breast cancer detection based on Naïve Bayesian. Measure J Int Measure Confed 72:32–36. https://doi.org/10.1016/j.measurement.2015.04.028
24. Karaboga D, Basturk B (2007) Artificial bee colony (ABC) optimization algorithm for solving constrained optimization problems. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics, Vol 4529 LNAI, pp 789–798. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-72950-1-7
25. Kessous R, Davidson E, Meirovitz M, Sergienko R, Sheiner E (2017) Prepregnancy obesity: a risk factor for future development of ovarian and breast cancer. Eur J Cancer Prev 26(2):151–155. https://doi.org/10.1097/CEJ.0000000000000228
26. Kiyan T, Yildirim T (2004) Breast cancer diagnosis using statistical neural networks. Istanb Univ J Electr Electron Eng 4(2):1149–1153
27. Koo MM, von Wagner C, Abel GA, McPhail S, Rubin GP, Lyratzopoulos G (2017) Typical and atypical presenting symptoms of breast cancer and their associations with diagnostic intervals: evidence from a national audit of cancer diagnosis. Cancer Epidemiol 48:140–146. https://doi.org/10.1016/j.canep.2017.04.010
28. Krawczyk B, Filipczuk P (2014) Cytological image analysis with firefly nuclei detection and hybrid one-class classification

decomposition. Eng Appl Artif Intell 31:126–135. https://doi.org/10.1016/j.engappai.2013.09.017

29. Kumar S, Datta D, Kumar Singh S, Azar AT, Vaidyanathan S (2015) Black hole algorithm and its applications. Stud Comput Intell 575:147–170. https://doi.org/10.1007/978-3-319-11017-2-7

30. Laamiri FZ, Bouayad A, Hasswane N, Ahid S, Mrabet M, Amina B (2015) Risk factors for breast cancer of different age groups: Moroccan data? Open J Obstet Gynecol 05(02):79–87. https://doi.org/10.4236/ojog.2015.52011

31. Lavanya S, Prasanth A, Jayachitra S, Shenbagarajan A (2021) A tuned classification approach for efficient heterogeneous fault diagnosis in IoT-enabled WSN applications. Measurement 183:109771

32. Li DC, Liu CW (2010) A class possibility based kernel to increase classification accuracy for small data sets using support vector machines. Expert Syst Appl 37(4):3104–3110. https://doi.org/10.1016/j.eswa.2009.09.019

33. Li Z, Yu L, Wang X, Yu H, Gao Y, Ren Y, Zhou X (2018) Diagnostic performance of mammographic texture analysis in the differential diagnosis of benign and malignant breast tumors. Clin Breast Cancer 18(4):e621–e627. https://doi.org/10.1016/j.clbc.2017.11.004

34. Littlestone N, Warmuth MK (1994) The weighted majority algorithm. Inf Comput 108(2):212–261. https://doi.org/10.1006/INCO.1994.1009

35. Luo J, Margolis KL, Wactawski-Wende J, Horn K, Messina C, Stefanick ML, Rohan TE (2011) Association of active and passivesmoking with risk of breast cancer among postmenopausal women: a prospective cohort study. BMJ 342(7796):536. https://doi.org/10.1136/bmj.d1016

36. Nahato KB, Harichandran KN, Arputharaj K (2015) Knowledge mining from clinical datasets using rough sets and backpropagation neural network. Comput Math Methods Med. https://doi.org/10.1155/2015/460189

37. Nazarian M, Dezfouli MA, Haronabadi A (2013) Classification of breast cancer samples through using the artificial bee colony algorithm. Int J Comput Appl Technol Res 2(5):522–525. https://doi.org/10.7753/ijcatr0205.1004

38. Nguyen AT, Shiao SL, McArthur HL (2021) Advances in combining radiation and immunotherapy in breast cancer. Clin Breast Cancer 21(2):143–152. https://doi.org/10.1016/j.clbc.2021.03.007

39. Nguyen T, Khosravi A, Creighton D, Nahavandi S (2015) Classification of healthcare data using genetic fuzzy logic system and wavelets. Expert Syst Appl 42(4):2184–2197. https://doi.org/10.1016/j.eswa.2014.10.027

40. Punitha S, Amuthan A, Joseph KS (2019) Enhanced monarchy butterfly optimization technique for effective breast cancer diagnosis. J Med Syst 43(7):1–14. https://doi.org/10.1007/s10916-019-1348-8

41. Sekar J, Aruchamy P, Sulaima Lebbe Abdul H, Mohammed AS, Khamuruddeen S (2022) An efficient clinical support system for heart disease prediction using TANFIS classifier. Comput Intell 38(2):610–640. https://doi.org/10.1111/coin.12487

42. Senapati MR, Mohanty AK, Dash S, Dash PK (2013) Local linear wavelet neural network for breast cancer recognition. Neural Comput Appl 22(1):125–131. https://doi.org/10.1007/s00521-011-0670-y

43. Shahbeig S, Helfroush MS, Rahideh A (2017) A fuzzy multi-objective hybrid TLBO-PSO approach to select the associated genes with breast cancer. Signal Process 131:58–65. https://doi.org/10.1016/j.sigpro.2016.07.035

44. Sheikhpour R, Sarram MA, Sheikhpour R (2016) Particle swarm optimization for bandwidth determination and feature selection of kernel density estimation based classifiers in diagnosis of breast cancer. Appl Soft Comput J 40:113–131. https://doi.org/10.1016/j.asoc.2015.10.005

45. Storn R, Price K (1997) Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. J Global Optim 11(4):341–359. https://doi.org/10.1023/A:1008202821328

46. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 71(3):209–249. https://doi.org/10.3322/caac.21660

47. Tao D, Tang X, Wu X (2006) Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. IEEE Trans Pattern Anal Mach Intell 28(7):1088–1099. https://doi.org/10.1109/TPAMI.2006.134

48. Thawkar S, Sharma S, Khanna M, Kumar Singh L (2021) Breast cancer prediction using a hybrid method based on butterfly optimization algorithm and ant lion optimizer. Comput Biol Med 139:104968. https://doi.org/10.1016/j.compbiomed.2021.104968

49. Vieira R, Sánchez Tobar JS, Dardes R, Thuler LCS (2018) Alcohol consumption as a risk factor for breast cancer development: a case-control study in Brazil. Asian Pac J Cancer Prev 19(3):703–707. https://doi.org/10.22034/APJCP.2018.19.3.703

50. Waks AG, Winer EP (2019) Breast cancer treatment: a review. JAMA 321(3):288–300. https://doi.org/10.1001/jama.2018.19323

51. Wang H, Zheng B, Yoon SW, Ko HS (2018) A support vector machine-based ensemble algorithm for breast cancer diagnosis. Eur J Oper Res 267(2):687–699. https://doi.org/10.1016/j.ejor.2017.12.001

52. Wang M, Chen H (2020) Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis. Appl Comput J 88:105946. https://doi.org/10.1016/j.asoc.2019.105946

53. Wang Y, Wang N, Xu M, Yu J, Qin C, Luo X, Ni D (2020) Deeply-supervised networks with threshold loss for cancer detection in automated breast ultrasound. IEEE Trans Med Imag 39(4):866–876. https://doi.org/10.1109/TMI.2019.2936500

54. Wuniri Q, Huangfu W, Liu Y, Lin X, Liu L, Yu Z (2019) A generic-driven wrapper embedded with feature-type-aware hybrid bayesian classifier for breast cancer classification. IEEE Access 7:119931–119942. https://doi.org/10.1109/ACCESS.2019.2932505

55. Yeh WC, Chang WW, Chung YY (2009) A new hybrid approach for mining breast cancer pattern using discrete particle swarm optimization and statistical method. Expert Syst Appl 36(4):8204–8211. https://doi.org/10.1016/j.eswa.2008.10.004

56. Zarbakhsh P, Addeh A, Demirel H (2018) Early detection of breast cancer using optimized ANFIS and features selection. In: Proceedings-9th international conference on computational intelligence and communication networks, cicn 2017, Vol 2018-Janua, pp 39–42. Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CICN.2017.8319352

57. Zheng B, Yoon SW, Lam SS (2014) Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. Expert Syst Appl 41(1):1476–1482. https://doi.org/10.1016/j.eswa.2013.08.044