



## OPEN A quantum-optimized approach for breast cancer detection using SqueezeNet-SVM

Anas Bilal<sup>1,2</sup>, Ali Alkhatlan<sup>3</sup>, Faris A. Kateb<sup>4</sup>, Alishba Tahir<sup>5</sup>, Muhammad Shafiq<sup>6</sup> & Haixia Long<sup>1,2</sup>✉

Breast cancer is one of the most aggressive types of cancer, and its early diagnosis is crucial for reducing mortality rates and ensuring timely treatment. Computer-aided diagnosis systems provide automated mammography image processing, interpretation, and grading. However, since the currently existing methods suffer from such issues as overfitting, lack of adaptability, and dependence on massive annotated datasets, the present work introduces a hybrid approach to enhance breast cancer classification accuracy. The proposed Q-BGWO-SQSVM approach utilizes an improved quantum-inspired binary Grey Wolf Optimizer and combines it with SqueezeNet and Support Vector Machines to exhibit sophisticated performance. SqueezeNet's fire modules and complex bypass mechanisms extract distinct features from mammography images. Then, these features are optimized by the Q-BGWO for determining the best SVM parameters. Since the current CAD system is more reliable, accurate, and sensitive, its application is advantageous for healthcare. The proposed Q-BGWO-SQSVM was evaluated using diverse databases: MIAS, INbreast, DDSM, and CBIS-DDSM, analyzing its performance regarding accuracy, sensitivity, specificity, precision, F1 score, and MCC. Notably, on the CBIS-DDSM dataset, the Q-BGWO-SQSVM achieved remarkable results at 99% accuracy, 98% sensitivity, and 100% specificity in 15-fold cross-validation. Finally, it can be observed that the performance of the designed Q-BGWO-SQSVM model is excellent, and its potential realization in other datasets and imaging conditions is promising. The novel Q-BGWO-SQSVM model outperforms the state-of-the-art classification methods and offers accurate and reliable early breast cancer detection, which is essential for further healthcare development.

**Keywords** Quantum computing, Breast cancer, Support vector machine, Grey wolf optimization, Medical image analysis

Cancer is a significant contributor to death globally, especially in the United States, and presents a considerable challenge to increasing life expectancy<sup>1</sup>. According to the World Health Organization's 2019 statistics, cancer is the primary or secondary cause of death in 112 out of 183 nations and the third or fourth cause in an additional 23 countries<sup>2</sup>. In 2020, the International Agency for Research on Cancer (GLOBOCAN) estimated that there were around 19.3 million new cases of cancer and 10.0 million deaths caused by cancer. Out of all the cases, there were almost 2.3 million new instances of female breast cancer, which is the most often diagnosed malignancy, exceeding lung cancer. However, lung cancer continued to be the leading cause of cancer-related mortality, resulting in almost 1.8 million deaths globally. Cancer is defined by the accelerated and unregulated proliferation of cells, which disseminate throughout the body and infiltrate tissues. Tumors are formed as a result of unregulated cell proliferation<sup>3</sup>. Normal cells undergo division to facilitate tissue repair and are ultimately eliminated from the body.

In contrast, malignant cells exhibit aberrant proliferation, infiltrating healthy tissues and inducing genetic mutations<sup>4</sup>. Tumors may be categorized into two distinct types: benign and malignant. Benign tumors are characterized by their inability to metastasize and are non-malignant. However, they might develop into masses

<sup>1</sup>College of Information Science and Technology, Hainan Normal University, Haikou 571158, China. <sup>2</sup>Key Laboratory of Data Science and Smart Education, Ministry of Education, Hainan Normal University, Haikou 571158, China.

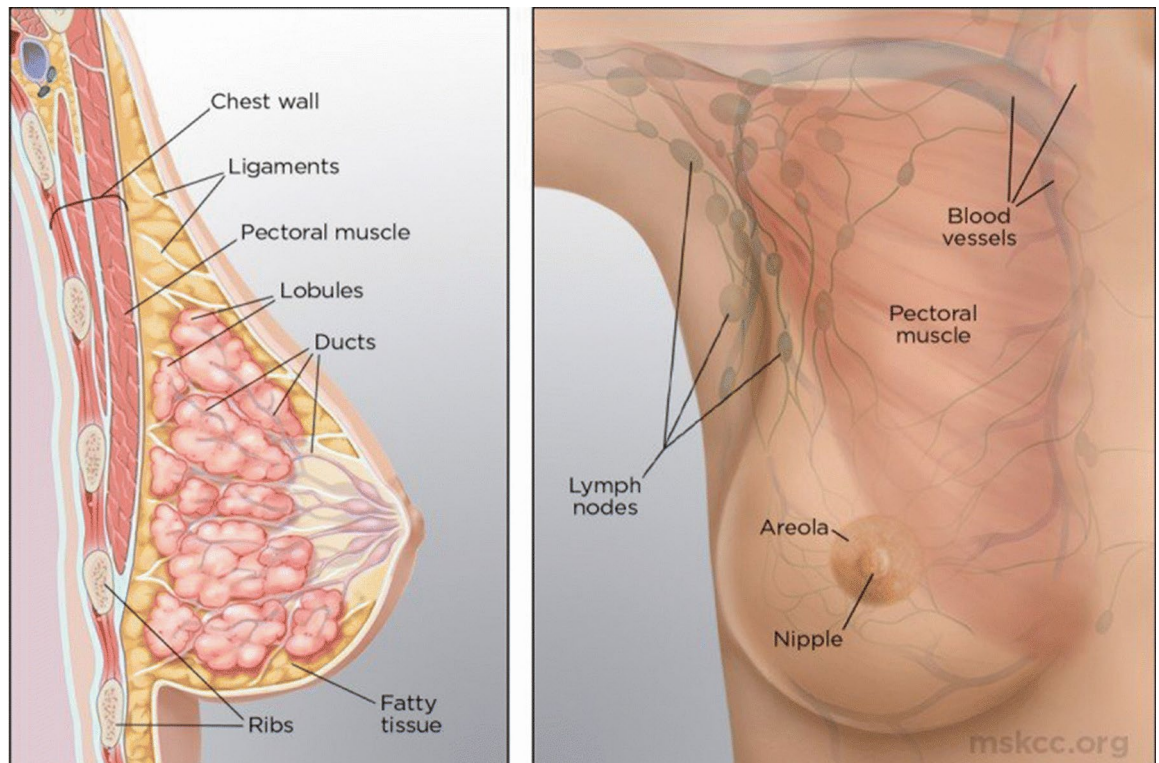
<sup>3</sup>Department of Computer Science, Faculty of Computing and Information Technology, King AbdulAziz University, Jeddah, Saudi Arabia. <sup>4</sup>Department of Information Technology, Faculty of Computing and Information Technology, King AbdulAziz University, Jeddah, Saudi Arabia. <sup>5</sup>Shifa College of Medicine, Shifa Tamere Milat University, Islamabad, Pakistan. <sup>6</sup>School of Information Engineering, Qujing Normal University, Yunnan, China. ✉email: myresearch\_hainnu@163.com

due to excessive cell proliferation. On the other hand, malignant tumors are carcinogenic and can infiltrate nearby tissues and metastasize to different areas of the body. Female breast cancer is the most common kind of cancer among numerous types, such as lung, liver, colorectal, stomach, and breast cancer (BC)<sup>5</sup>. On a global scale, around 12.5% of adult women will get a breast cancer diagnosis at some stage throughout their lives. In breast cancer, the illness may arise in several areas of the breast, such as the ducts, lobules, or the surrounding tissues, as seen in Fig. 1, which presents the anatomical structure of the breast.

BC is a major worldwide health issue, causing 10 million fatalities and 19.3 million new cases in 2020 alone. This disorder arises from genetic abnormalities that lead to aberrant cell growth, leading to non-cancerous or cancerous tumors. Breast cancer is the second most prevalent form of cancer worldwide and the fifth highest cause of mortality among women. Mammography has been crucial in the timely identification of breast cancer, leading to a considerable decrease in death rates. This is achieved by capturing breast images from two angles: Medio lateral oblique and craniocaudally. This technique classifies breast density into four categories: adipose, dispersed, heterogeneously dense, and very dense<sup>7,8</sup>. Tumors exhibit a range of forms and borders, with malignant tumors often displaying irregular and unclear boundaries, whereas benign tumors are frequently compact and well-defined. Calcifications exhibit various characteristics; non-cancerous calcifications usually manifest as sizable rods or popcorn-shaped structures, while cancerous ones are spread out or grouped<sup>9</sup>. Multiple factors contribute to the formation of breast cancer. Figure 2 depicts the variables that impact the occurrence of breast cancer, while Table 1 provides a summary of the many phases involved in the progression of breast cancer. This illness may readily propagate via a woman's thoracic region, glands, and lactiferous ducts.

Early detection and diagnosis of breast cancer greatly enhance survival rates and decrease the occurrence of false positive results, which refer to cases when patients are mistakenly identified as cancer-free. Several diagnostic techniques are used to detect breast cancer, such as mammography, CT scans that utilize X-rays, MRI, and ultrasound, which use magnetic energy and sound waves. Mammography is often the first stage in the identification of breast cancer. However, it may provide difficulties in detecting cancer in women with thick breast tissue; moreover, the radiation released during mammography hazards both radiologists and patients. Similarly, CT scans use radiation that can potentially induce genetic abnormalities in tissues. Ultrasound is the ideal modality for women with thick breast tissue; nevertheless, its reliability is not always consistent, and it is linked to a significantly elevated percentage of false positive results. In order to tackle these issues, biopsies are performed. A biopsy involves a thorough physical examination followed by the extraction of a tissue sample for microscopic analysis. Subsequently, a pathologist examines and analyses this specimen. Histopathology is the meticulous examination of biological cells and tissues using a microscope.

However, the manual inspection of samples by a pathologist is a time-consuming and laborious process, requiring high skills and knowledge. It may result in possible errors related to mental fatigue. For decades, computer-aided detection has been one of the leading systems for diagnosing cancer, successfully overcoming some associated challenges. When it comes to breast cancer examination, ordinary X-ray pictures are applied to



**Fig. 1.** Anatomy of breast<sup>6</sup>.

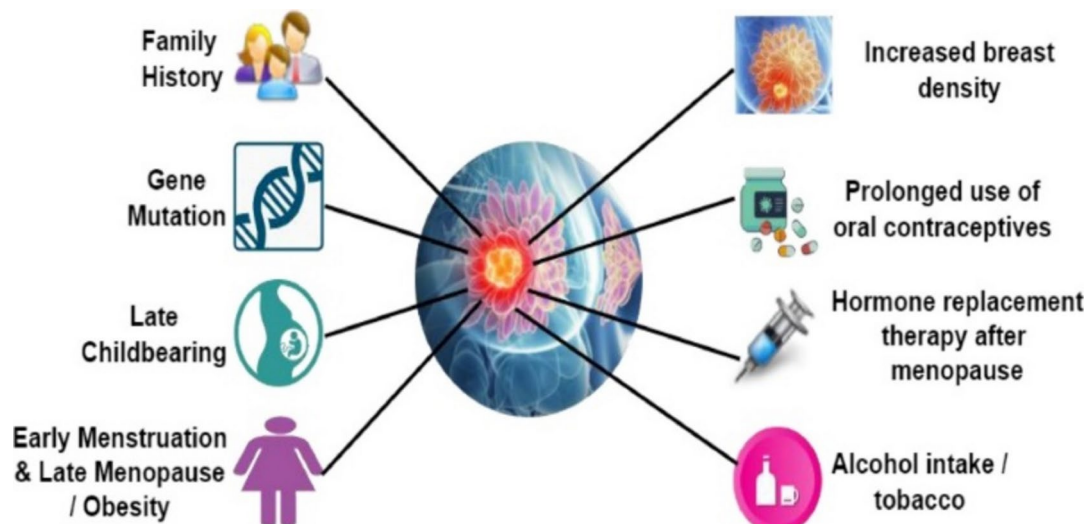


Fig. 2. Factors affecting Breast cancer disease<sup>10</sup>.

Level	Statement	Size	Recommended treatment
1	Non-invasive cancer	Small	Simple mastectomy or breast-conserving surgery
2	Invasive	2 cm or below	Radiation therapy, surgery
3	Invasive	2–5 cm	Target drugs, immunotherapy, chemotherapy, hormone therapy
4	Invasive	Above 5 cm	Chemotherapy, immunotherapy, hormone therapy
5	Invasive	Any size	Chemotherapy, immunotherapy, hormone therapy

Table 1. The stages of development of breast cancer.

distinguish and classify samples of tissues as malignant or benign<sup>11</sup>. Time-consuming manual examination and classification of breast lesions in mammography rely on the competence and experience of radiologists, which, in turn, may lead to possible errors in diagnosing. The large number of images causes the radiologists' workload and the rates of mistakes to grow. Recently, some computer-aided detection systems have been developed to improve the degree of accuracy and alleviate the work of radiologists in this field. The computer-aided detection system automatically sorts out regions of interest in mammograms and then classifies them. However, they still have not solved challenges, including the variability in shape and intensity and the problem of overlapping appearance of tissues in mammograms.

The proposed study introduces a novel method for classifying breast cancer that combines an improved Quantum-Inspired Binary Grey Wolf Optimizer (Q-BGWO) with a SqueezeNet Support Vector Machine (SQSVM), offering significant advancements over existing approaches. Unlike traditional methods, which often face challenges such as overfitting, limited adaptability to high-dimensional data, and reliance on extensively annotated datasets, this approach integrates quantum computing principles into nature-inspired optimization, addressing these limitations effectively. By combining the Q-BGWO's optimization capabilities with SqueezeNet's efficient feature extraction and SVM's robust classification, this work establishes a new standard in breast cancer diagnostics.

The novelty of this work lies in its unique hybrid framework, where quantum computational principles such as quantum rotation gates and probabilistic adjustments enhance the search space exploration and convergence speed of the Grey Wolf Optimizer. This advancement allows the model to optimize SVM hyperparameters precisely, ensuring higher classification accuracy and robustness across diverse datasets. Unlike existing nature-inspired algorithms, the Q-BGWO-SQSVM achieves superior accuracy, sensitivity, and specificity performance, outperforming state-of-the-art models across various benchmarks. Additionally, the inclusion of SqueezeNet's lightweight architecture ensures computational efficiency without compromising diagnostic precision, making it suitable for real-time clinical applications.

Following the analysis and description of quantum computing and its cooperation with nature-inspired algorithms, the study takes critical steps to use the design and mechanisms of the Grey Wolf Optimizer, which represents the strategic behaviors of various organisms in nature to optimize their means of solving complex problems. These behaviors primarily aim to overcome difficulties related to specific survival tasks. The integration of quantum computational capacities enables the optimizer to refine its operation by incorporating quantum characteristics, such as superposition and probabilistic adjustments, to address the challenges faced by traditional methods. Tools such as Grover's and Shor's brief factorization algorithms are applied, concretely improving the speed of search and calculation. The study effectively extends the application of classical heuristics

to resolving significant engineering and computing tasks through a combination with quantum improvements. This innovative approach not only advances the field of medical image classification but also sets a new benchmark for automated detection of breast cancer, making it a transformative tool in clinical diagnostics.

The innovative contributions of the study include:

- Developing a novel hybrid framework that combines the Q-BGWO with SqueezeNet-SVM significantly advances the capabilities of automated breast cancer detection systems.
- Integrating quantum-inspired optimization principles into the Grey Wolf Optimizer enhances its ability to handle high-dimensional, complex data and avoids premature convergence common in traditional meta-heuristic algorithms.
- An efficient feature extraction mechanism using SqueezeNet's fire modules and bypass mechanisms, ensuring computational efficiency while retaining diagnostic precision.
- State-of-the-art performance metrics, as demonstrated on diverse datasets such as MIAS, INbreast, DDSM, and CBIS-DDSM, achieve outperformance over existing models in accuracy, sensitivity, and specificity.
- A practical and scalable solution for clinical diagnostics, reducing reliance on large annotated datasets while maintaining high accuracy and robustness.

Key contributions of the study are:

- The study addressed issues concerning the external diagnosis of breast cancer through the careful dissection of resource/manual tracking as insufficient for the verification of such matters. Subsequently, a system based on computer vision tools and artificial intelligence for the automated diagnosis of issues, such as the detection of the individual modality, was founded.
- The study has developed a novel method for classifying breast cancer by integrating a Q-BGWO and a SQSVM. The proposed strategy contributes to improving accuracy at the point of search application.
- In the first phase of this study, a computer vision algorithm based on pixel-wise segmentation was implemented to accurately extract regions of architectural distortion regions of interest from digital mammogram images.
- In the next phase, the SqueezeNet algorithm was used to automatically extract features of the abovementioned ROIs and classify them as malignant or benign. Therefore, an optimized support vector machine was used to classify the same effectively, resulting in a comprehensive diagnostic tool.
- This novel method has been tested with different datasets, such as MIAS, INbreast, DDSM, and CBIS-DDSM, and greatly improved performance metrics can distinguish the results from existing ones. Notably, the Q-BGWO-SQSVM achieved 99% accuracy, 98% sensitivity, and 100% specificity on the CBIS-DDSM dataset, setting a new benchmark for automated breast cancer detection.
- Achieved outperformance over existing models, delivering higher accuracy and specificity while maintaining sensitivity, thus providing a reliable and efficient tool for early breast cancer detection.

The paper is organized as follows: Sect. “Preliminaries” presents a comprehensive introduction to machine learning algorithms, quantum computing (QC), and nature-inspired algorithms, which serve as the foundation for the subsequent sections. Sect. “Related work” examines a literature review. Sects. “Material and methods” and “Experimental results” discuss the suggested approach and provide the findings from the experiments. Sect. “Conclusions” serves as the final part of the report, concisely summarizing the key insights and conclusions derived from the research.

## Preliminaries

### Support vector machine (SVM)

SVM aims to identify an  $N$ -dimensional hyperplane that classifies the data vectors with minimal error. SVM uses convex quadratic programming to avoid local minima<sup>12</sup>. Considering a binary classification problem with a training dataset containing class labels:  $(x_1, y_1), \dots, (x_n, y_n)$   $x_i \in R_d$  and  $y_i \in (-1, +1)$ ,  $x_i$  represents the feature vector, and  $y_i$  represents the class label; the optimal hyperplane can be described as follows:

$$wx^T + b = 0 \quad (1)$$

$W$ ,  $x$ , and  $b$  represent the weight, input vector, and bias, respectively. The parameters  $w$  and  $b$  must satisfy the following conditions:

$$wx_i^T + b \geq +1 \text{ if } y_i = 1 \quad (2)$$

$$wx_i^T + b \leq -1 \text{ if } y_i = -1 \quad (3)$$

The primary objective in training an SVM model is to identify values for  $w$  and  $b$  that maximize the margin  $\frac{1}{\|w\|^2}$ . Challenges often arise with problems that are not linearly separable. To address this, the input space is transformed into a higher-dimensional space, making the problem linearly manageable.

Kernel functions<sup>13</sup> are used to increase the dimensionality of the data, making the issue easier to solve linearly. Kernel functions also make it possible to calculate more efficiently in spaces of higher dimensions. For example, a linear kernel might compute the dot product of two sets of features in this numerical feature space. Radial Basis Function and polynomial are two popular SVM kernels. The following is an explanation of these kernels.

$$K(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (4)$$

$$K(x_i, x_j) = (1 + x_i^T x_j)^p \quad (5)$$

Parameters such as  $\gamma$  and  $p$  are significant in the given SVM framework. In particular, while  $\gamma$  denotes the width of the Gaussian kernel,  $p$  is the order of the polynomial. According to the outcomes of prior studies, properly configuring these model parameters considerably enhances the accuracy of SVM classification<sup>14</sup>. Thus, it is essential to gradually fine-tune such significant parameters as  $C$  and  $\gamma$  and the selection of the SVM Gaussian kernel function. In turn, such adjustments contribute to increasing the input space's dimensionality to allow non-linear issues to be construed as linear.

### Quantum computing

In quantum computing, the equivalent of a bit is a qubit, which can exist in a superposition of 0 and 1 states. In other words, a qubit  $|\varphi\rangle$  can be equal to  $|0\rangle$  and  $|1\rangle$  at the same time, with specific probabilities of such states. Most importantly, the superposition principle is ensured since a qubit  $|\varphi\rangle$  can be represented as a linear combination of  $|0\rangle$  and  $|1\rangle$  states. At the same time,  $|0\rangle$  is the ground state of a qubit, while  $|1\rangle$  is the excited one they could be compared to the two binary states of a conventional bit. Therefore, the probability of measuring  $|0\rangle$  or  $|1\rangle$  is associated with their coefficients and can be calculated by squaring them within the linear combination after normalizing these states. The normalization condition establishes that the sum of the squared probabilities of any particle is always 1; thereby, the probability conservation is provided. Thus, qubits are used to develop quantum algorithms and help achieve exponentially faster calculations of such activities as factoring large numbers or scanning unsorted databases<sup>15</sup>. The mathematical qubit is expressed as:

$$|\varphi\rangle = x|0\rangle + y|1\rangle \quad (6)$$

where  $x$  and  $y$  are complex numbers that satisfy the normalization condition  $|x|^2 + |y|^2 = 1$ . In this case,  $|x|^2$  and  $|y|^2$  are the probabilities of finding the qubit as a  $|0\rangle$  and  $|1\rangle$  state. Quantum operations on the qubits are mostly done through quantum gates or operators, representing such operators as matrices. Thus, operators enable qubit state manipulation, allowing for the move from one state to another. The primary quantum operators are the Pauli operators: P, X, Y, and Z, as shown in Table 2. Examples of complex gates include Toffoli, Feynman, Fredkin, Swap, and Peres gates that perform vital roles in quantum circuits. Rotation gates such as x-gate, y-gate, and z-gate rotate qubits along the X, Y, or Z axes, and their transformations in matrices form is Mathematically shown<sup>16</sup>.

Equations 7, 8 introduced in the article show how rotation gates are applied to change the update phases of GWO. To continue the development of algorithms that are influenced by quantum theories, this study also applies the principles of quantum rotation matrices to traditional algorithms. This approach remains novel, and this is the first time quantum mechanics and nature-inspired algorithms have been applied to analyze the outcomes of the interaction and integration of the two discussed approaches in feature selection.

$$x_{gate(\theta)} = \begin{pmatrix} \cos(\theta) & -i\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad (7)$$

$$y_{gate(\theta)} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \quad (8)$$

$$z_{gate(\theta)}(\theta) = \begin{pmatrix} e^{-i\theta} & 0 \\ 0 & e^{i\theta} \end{pmatrix} \quad (9)$$

### Binary grey wolf optimization

The recent improvements have stimulated the considerable growth of nature-inspired algorithms in many technical domains. The Artificial Bee Colony is the approach developed by Derviş Karaboa that is based on the foraging behavior of honeybees to solve numerical optimization problems. Another solution is the Grey

Gate name	Matrix
I-gate	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
X-gate	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Y-gate	$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$
Z-gate	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

**Table 2.** Q-Gates along with respective matrices.

Wolf Optimizer introduced by Seyed Mirjalili<sup>17</sup>. The latter approach employs the hunting strategies of wolves to improve the found solutions. Meanwhile, the Elephant Search Algorithm relies on the unique search behaviors of male and female groups of elephants. The Artificial Algae Algorithm<sup>18</sup> is an innovative platform that leverages the evolutionary processes of microalgae. The Fish Swarm Algorithm<sup>19</sup> has also been developed based on how fish schools forage. The GWO algorithm was conducted by Mirjalili et al. in 2014<sup>17</sup>. It exploits the search power of grey wolves with a complex social organization. According to the process, the GWO algorithm involves three best solutions during each iteration. In this case,  $\alpha$ ,  $\beta$ , and  $\delta$  start leading and moving the pack to promising areas. The latter is assisted by so-called omegas who participate in exploring and optimizing the search landscape. This well-organized structure allows investigating the search surface and the identification all possible optimal results. The structure of this process is depicted in Fig. 3.

The encirclement behavior of Grey Wolves (GW) can be analytically described using the equations below:

$$\vec{K} = \left| \vec{C} * \vec{X}_{prey}(t) - \vec{X}(t) \right| \tag{10}$$

$$\vec{X}(t+1) = \vec{X}_{prey}(t) - \vec{A} * \vec{K}$$

In these equations,  $t$  denotes the current iteration,  $\vec{X}_{prey}$  Is the positional vector of the prey,  $\vec{X}$  represents the positional vector of the GW,  $\vec{A}$  linearly decreases from 2 to 0, and  $\vec{r}_1, \vec{r}_2$  are random vectors uniformly distributed within the range [0,1]. These equations are crucial in modeling the foraging tactics of grey wolves, as illustrated in Fig. 4.

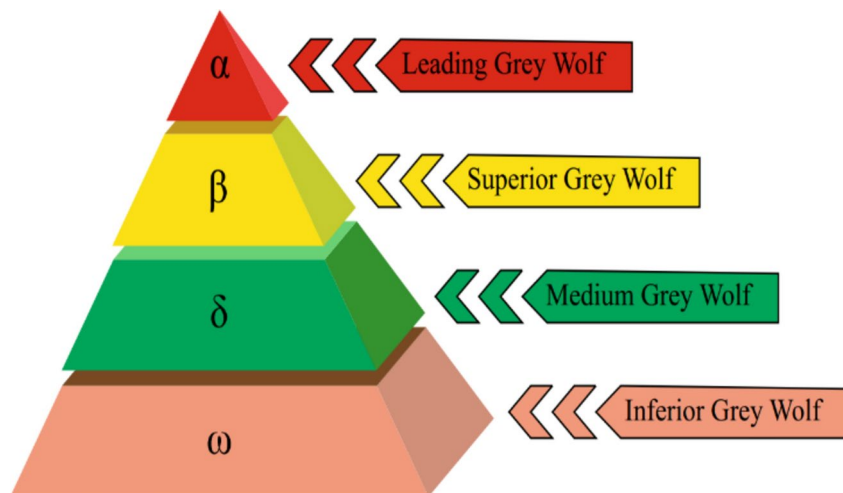
Figure 5 presents the Binary Grey Wolf Optimizer (BGWO) as a flowchart. In the BGWO algorithm, each Grey Wolf possesses a flag vector, the length of which matches that of the dataset. The positions of the Grey Wolves are updated according to Eqs. (10).

$$\text{flag}_{i,j} = \begin{cases} 1 & X_{i,j} > 0.5 \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

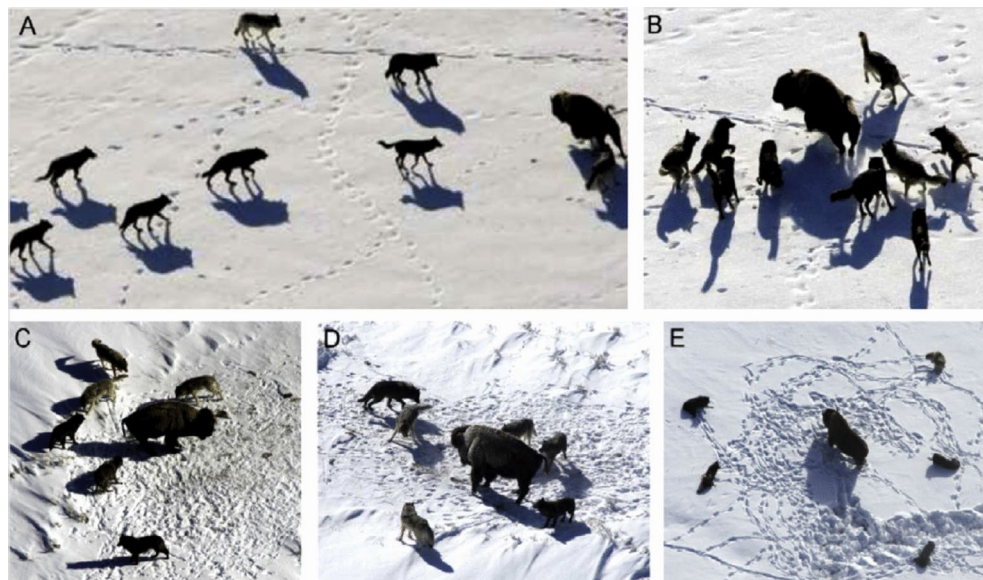
where  $X_{i,j}$  Designates the  $i$ th GW's  $j$ th position.

**Related work**

Recent advancements have highlighted the potential of expert systems and artificial intelligence in aiding the detection of breast cancer, helping medical professionals avoid costly errors. These systems efficiently process medical data and critically support less experienced doctors. Breast cancer detection has seen notable successes through various artificial intelligence strategies, including<sup>21</sup> GWO-based feature selection combined with deep learning for microarray cancer datasets<sup>22</sup> ensemble models for predicting breast cancer relapse from histopathological images<sup>23</sup> hybrid optimization techniques like Binary Grey Wolf with Jaya Optimizer for biomarker selection<sup>24</sup> fog-empowered transfer learning frameworks such as CanDiag for cancer diagnosis<sup>25</sup> chaotic firefly algorithms for global optimization with dropout regularization<sup>26</sup> GA-based hierarchical feature selection for handwritten word recognition<sup>27</sup> hybrid firefly algorithms for XGBoost tuning in intrusion detection<sup>28</sup> improved sine cosine metaheuristics for IoT security using hybrid machine learning models and<sup>29</sup> metaheuristic-tuned extreme learning machines tackling IoT security challenges<sup>30</sup>.



**Fig. 3.** Hierarchy of the grey wolves(GW)<sup>17</sup>.



**Fig. 4.** Grey wolves go through many distinct phases in their hunting activity. The first phase 'A involves the wolf's pursuit, approach, and tracking of its prey. When the wolves have located their prey, they will begin harassing and pursuing it in stage B-D. At this point, the pack may form a circle around the victim to keep it from running away. The last step E is to assault the prey while the wolves remain positioned<sup>20</sup>.

These innovative methods collectively underline the effectiveness of hybrid and metaheuristic approaches in improving diagnostic accuracy and addressing complex challenges in medical and security domains. Despite their success, these approaches often face limitations such as dependency on large datasets, lack of adaptability across varied imaging conditions, and challenges in achieving consistent optimization of hyperparameters in high-dimensional data.

The AMMLP method has recorded an impressive accuracy of 99.26%<sup>31</sup>. Diagnoses have achieved peak and average accuracies of 100% and 96.87% using an RS-SVM classifier<sup>32</sup>. A PSO-SVM approach has yielded a 99.3% accuracy rate<sup>33</sup>. A genetically optimized neural network (GONN) specifically designed for breast cancer detection has used a genetic approach to enhance its structure, resulting in an astounding accuracy rate of 97.73%<sup>34</sup>. In addition, using genetic algorithms for feature selection has effectively identified the most efficient characteristics of breast cancer for SVM classification, leading to a 95.8% accuracy<sup>35</sup>. Advanced applications of genetic algorithms have improved breast cancer prediction. These include a genetic algorithm based on tribal competition and an online gradient-boosting strategy for feature selection. These methods have achieved accuracies of 98.32, 94.28, and 95.75% respectively<sup>36–38</sup>. In addition, a differential-evolution algorithm with multi-objective elitism and a skill acquisition technique based on graphs achieved accuracies of 96.86 and 75.05%, respectively, as reported in references<sup>33</sup>. Ensemble learning algorithms have also been applied, achieving a 97.4% accuracy rate<sup>41</sup>. An adjusted bat algorithm has optimized SVM parameters significantly, improving classifier accuracy to 96.49% from 96.31% for the WDBC dataset<sup>42</sup>. A feature selection strategy using GA reached a 96.9% accuracy<sup>43</sup>. Comparisons among different cancer classification models like naïve Bayes, logistic regression, and decision tree identified the logistic regression classifier as the most effective, with the highest accuracy of 97.9%<sup>44</sup>. Particle swarm optimization has been utilized to refine the C4.5 algorithm, improving its accuracy to 96.49% from 95.61% for the WBC dataset<sup>45</sup>. An improved cost-sensitive SVM, considering the unequal misclassification costs of breast cancer diagnosis, achieved a 98.83% accuracy on the WDBC dataset<sup>46</sup>. Additional techniques, such as the sparse pseudoinverse incremental-ELM, likelihood-based fuzzy analysis, and a dual-stage SVM-driven BC classification, have achieved accuracies of 95.26, 97.28, and 98% correspondingly<sup>47–49</sup>. In addition, preprocessing methods such as resampling, discretization, and handling missing data have been used. This has resulted in classifications using Sequential-Minimal-Optimization, naïve Bayes, and J48 algorithms, with claimed average accuracies of 97.5, 98.73, and 98% respectively<sup>47</sup>.

While these approaches have made significant strides, their dependence on traditional optimization methods and general machine learning frameworks limits their ability to achieve high performance consistently across diverse imaging datasets. These methods cannot often balance computational efficiency with the complexity of medical imaging data.

The combination of Support Vector Machine (SVM) with Radial Basis Function (RBF) kernel and AdaBoost algorithm classifiers, using nature-inspired algorithms and applying the maximum likelihood principle, has improved the stability of classification. The accuracy rates achieved are 96% for AdaBoost, 97.37% for Particle Swarm Optimisation (PSO), 97.19% for Genetic Algorithm (GA), and 95.96% for Ant Colony Optimisation (ACO), as determined through tenfold cross-validation<sup>50,51</sup>. Furthermore, the integration of Support Vector Machines (SVM) with K-means clustering has successfully distinguished between cancerous and non-cancerous tumours, with an impressive accuracy rate of 97.38%<sup>52</sup>. PSO has been successfully applied to the MIAS, WDBC,

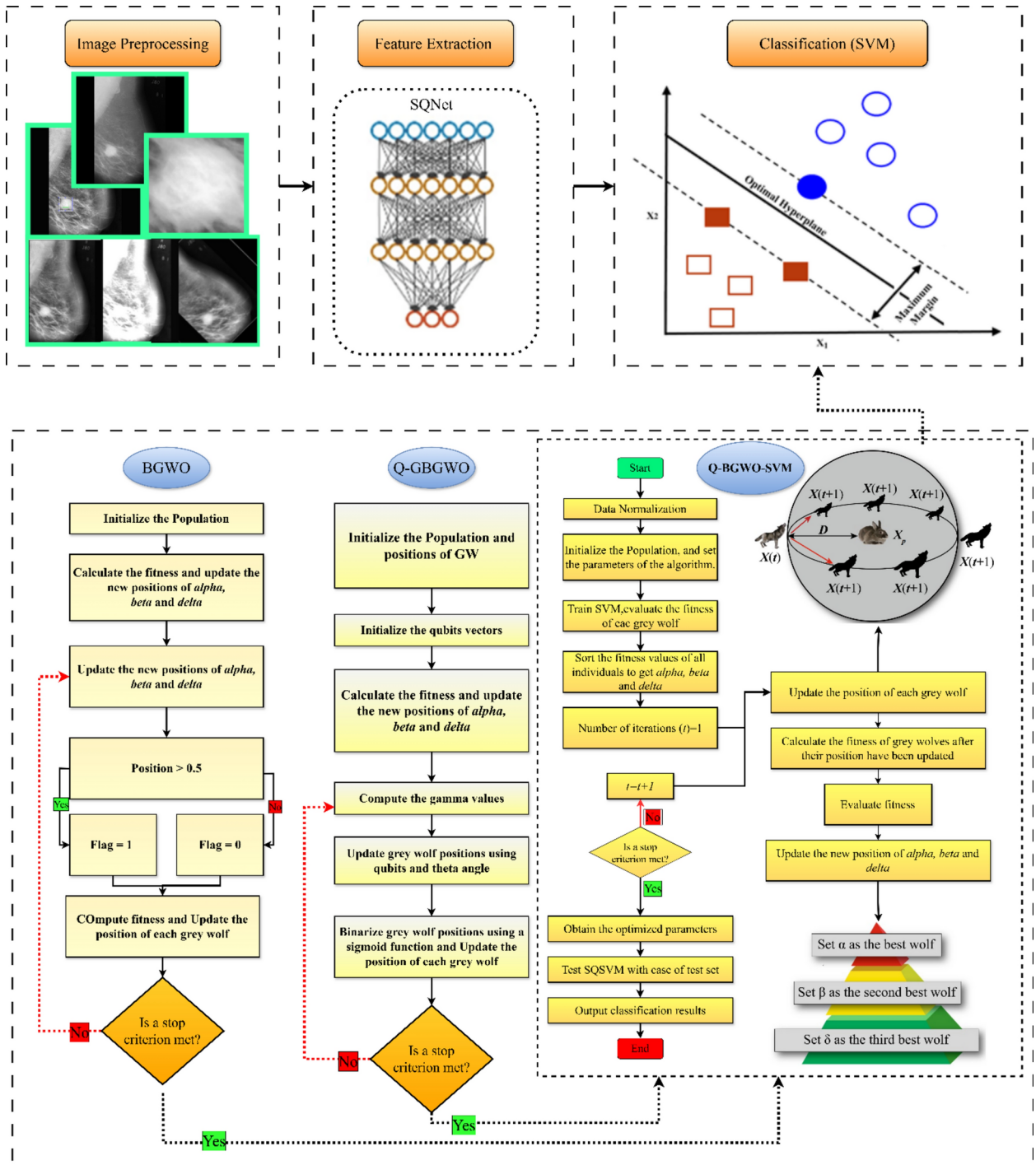


Fig. 5. The framework of the proposed methodology.

and WBCD datasets, as reported in datasets<sup>53</sup>. A study comparing six machine learning approaches achieved a 99.04% accuracy rate<sup>54</sup>. Augmented by the Levy flying strategy, the fruit fly optimization technique successfully optimized the Support Vector Machine (SVM) parameters and yielded significant outcomes in breast cancer diagnostics<sup>55</sup>. A novel approach has been proposed<sup>56</sup>, which combines Support Vector Machines (SVM) with Radial Basis Function (RBF) and polynomial kernel functions, together with the use of the dragonfly method. Two machine learning methods, a decision tree and K-nearest neighbors, were used to build models. These models were trained on data that underwent feature selection using principal component analysis (PCA). The results showed that the K-nearest neighbors classifier performed better than the decision tree<sup>57</sup>.

Building on these advancements, this study introduces a novel hybrid approach, the Quantum-Optimized Binary Grey Wolf Optimizer-SqueezeNet Support Vector Machine (Q-BGWO-SQSVM), which addresses these limitations by combining the strengths of quantum-inspired optimization with efficient deep learning

architectures. Unlike traditional approaches, Q-BGWO-SQSVM leverages quantum mechanics principles for superior hyperparameter optimization, enhancing feature selection and classification accuracy across diverse datasets.

Furthermore, genetic programming and machine learning techniques have resulted in a classification accuracy of 98.24%<sup>58</sup>. Researchers have developed optimization algorithms for categorizing the Iris and Breast Cancer datasets using Artificial Neural Networks (ANN), showing notable results<sup>59</sup>. The efficacy of several machine learning models in forecasting breast cancer was assessed using the WDBC dataset. The Support Vector Machine (SVM) model achieved a precision of 96.25%<sup>60</sup>. In addition, a novel diagnostic model that integrates the batting approach, gravitational search algorithm, and a feed-forward neural network achieved testing and training accuracies of 94.28% and 92.10%, respectively, on the same dataset<sup>61</sup>.

The unique contribution of Q-BGWO-SQSVM lies in its ability to overcome challenges such as high-dimensional data complexity, overfitting, and dataset variability. By integrating SqueezeNet's efficient feature extraction with Q-BGWO's robust optimization capabilities, the proposed model achieves state-of-the-art performance and ensures computational efficiency and adaptability to clinical applications.

A hybrid metaheuristic-swarm-intelligence-based SVM classifier, optimized using grey wolf and whale optimization methods, was evaluated on the WDBC dataset. The classifier achieved a classification accuracy of 97.721%<sup>62</sup>. Recent studies using the grey wolf optimizer have improved the performance of support vector machines (SVM) in diagnosing breast cancer. These studies have shown that the optimizer achieves rapid convergence and a high accuracy rate of 99.30%<sup>63</sup>. Efficient scaling solutions for optimizing Support Vector Machines (SVM) are successful for linear programming, as evidenced by references<sup>64–67</sup>. A novel parallel swarm approach has been created to address two-sided balancing issues.

Additionally, techniques have been devised to minimize computation time and solution area in conventional processes. These advancements have resulted in notable enhancements in efficiency and accuracy, as shown in previous studies<sup>68,69</sup>. Metaheuristic algorithms, including ABC, FPA, BA, PSO, and MFO, have optimized SVMs and extreme learning machines, successfully tackling problems related to local minima and overfitting<sup>70,71</sup>. Numerous research papers have explored a variety of ML algorithms for diagnosing breast cancer, particularly in datasets affected by impulsive noise. A multilayer extreme learning machine method based on full correntropy has been highlighted for its effectiveness<sup>72</sup>. Efforts to boost ANN performance while simultaneously cutting down on misclassification costs have been encapsulated in the LS-SOED method<sup>73</sup>. Moreover, logistic regression has been used to select features, in combination with the Data Handling Group Method and a smooth group L1/2 regularisation approach. This strategy seeks to remove unnecessary nodes from the inputs of feedforward neural networks, leading to substantial improvements in accuracy and precision<sup>74,75</sup>.

A fuzzy inference system that leverages an adaptive network and decision tree methodologies has shown considerable classification accuracies<sup>76</sup>. Recent advancements in hybrid ML models utilize diverse metaheuristic optimization techniques, including BBO, GWO, PSO, SCA, ChOA, SSA, WOA, AGPSO, and DA, to solve a broad spectrum of problems. Despite the No Free Lunch theorem suggesting no single metaheuristic algorithm universally outperforms others, specific algorithms have demonstrated greater efficacy for particular applications.

The proposed study seeks to determine how the Q-BGWO-SQSVM architecture categorizes mammographic pictures in the MIAS, Inbreast, DDSM, and CBIS-DDSM datasets. Here, the fundamental image processing methods are meant to eliminate background noise, consequently making the pictures more vivid and defined. Secondly, benign and malignant areas of interest are gathered. For Abnormal, the included aberrations presented in datasets along with their central coordinates indicate, simplify the process of taking square regions that perfectly fit the circumstances. Since Normal is not provided with specific spatial coordinates, the region of the place mentioned before is randomly picked<sup>77</sup>. These representative samples of both normality and deviancy are relatively distinct. However, their malignant and benign simulated Region of Interest are not easily distinguishable. Therefore, it can be suggested that to make the 'discernable patterns' more distinct from the two buckets—normal and abnormal categories BGWO and Q-BGWO could be applied as separate models.

Moreover, the combination of BGWO and Q-BGWO, using the QSVM, is supposed to give a highly accurate classifier that can tell the difference between malignant and benign breast cancer. This algorithm aimed to automate breast cancer classification, which promises high precision. Figure 5 reflects the entire schematic of our study.

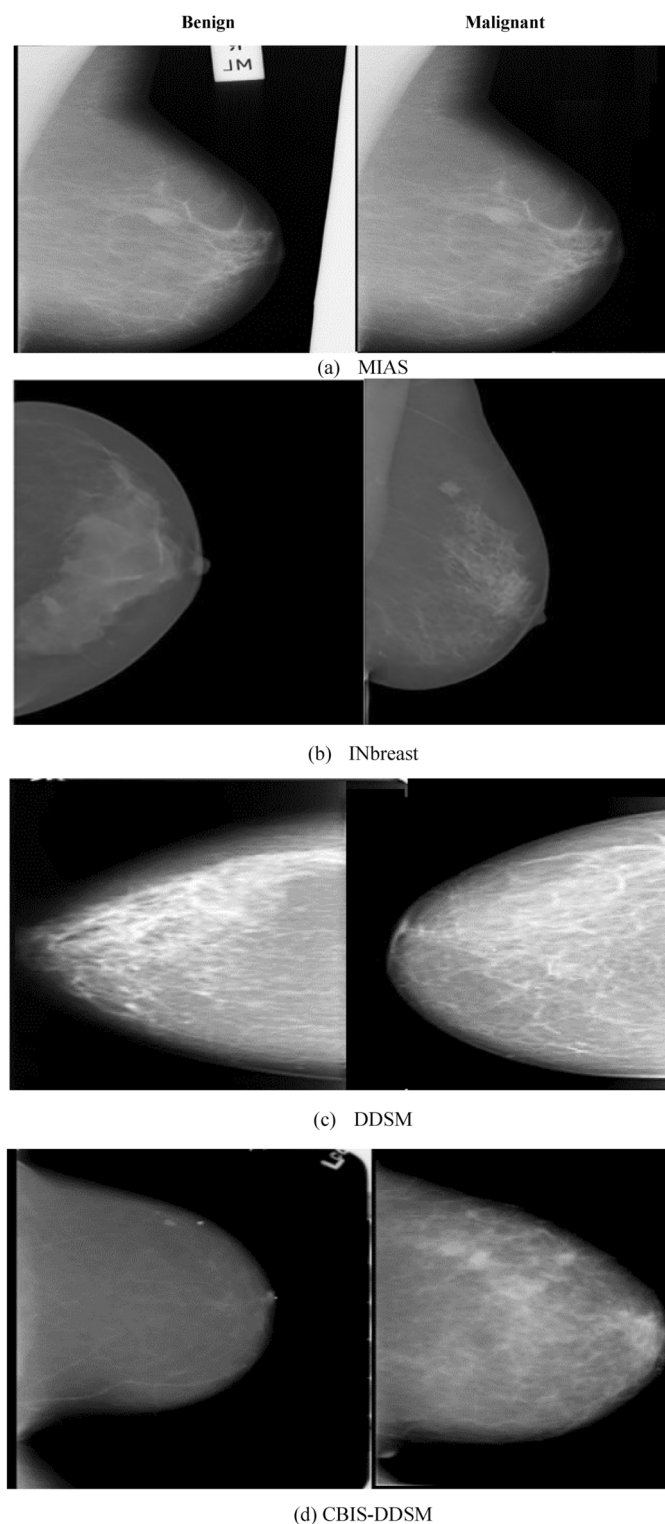
## Material and methods

### Dataset

This study evaluates the effectiveness of proposed algorithms by utilizing mammography images from four benchmark datasets: Mammographic Image Analysis Society (MIAS)<sup>78</sup>, INbreast<sup>79</sup>, Digital Database for Screening Mammography (DDSM)<sup>80</sup>, and Curated Breast Imaging Subset of DDSM (CBIS-DDSM)<sup>81</sup>. Figure 6 displays examples of normal and malignant mammograms from these datasets.

Rationale for Dataset Selection: These datasets were selected to ensure a comprehensive evaluation of the Q-BGWO-SQSVM model across varied imaging conditions, including differences in resolution, imaging formats, lesion characteristics, and annotation precision.

- MIAS represents a compact dataset offering annotated mammograms with diverse lesion types and associated prognostic data, making it suitable for validating the model's ability to handle small-scale, annotated datasets.
- INbreast provides high-resolution images in DICOM format, incorporating expert-annotated benign and malignant lesions, which test the model's capability to process high-dimensional medical data.



**Fig. 6.** Breast mammogram images from the datasets.

- DDSM contributes a large dataset of mammograms with detailed patient-specific information, enabling the evaluation of the model's generalization on extensive datasets with variable patient demographics and density ratings.
- CBIS-DDSM is an improved version of DDSM, offering binary masks and pixel-level annotations, which allow for precise validation of the model's feature extraction and classification accuracy.

The MIAS dataset<sup>78</sup> comprises 322 mammograms from 161 individuals, including 54 malignant, 61 benign, and 207 normal images. Each image, rendered in  $1024 \times 1024$  pixel grayscale, offers a resolution of 200 microns and is stored in PGM format. This dataset includes diverse mammogram stances and detailed annotations. It features various lesion types and related pathological prognostic data. The dataset also details the ROI, cropped based on the provided coordinates for each anomaly's center and radius. Only the benign (BIRADS 2, 3) and malignant (BIRADS 4, 5, 6) annotations are considered for this research.

The INbreast dataset<sup>79</sup> contains 410 mammography images, incorporating benign, malignant, and normal cases. This collection, saved in DICOM format, includes 192 benign and 107 malignant breast lesions, each delineated with expert precision. Image dimensions range from  $2560 \times 3328$  to  $3328 \times 4084$  pixels, adjusted according to breast size.

The DDSM dataset<sup>80</sup>, contributed to by institutions like Massachusetts General Hospital and Wake Forest University School of Medicine, includes comprehensive mammography data from 2500 benign and 3000 malignant cases. Each record in this dataset provides detailed patient-specific information such as age, ACR breast density rating, and subtlety assessment of anomalies. The ground truth data in the DDSM offers detailed descriptions of lesion locations and types at the pixel level, with 200 ROIs associated with both benign and malignant images.

The CBIS-DDSM<sup>81</sup>, an improved version of the DDSM curated by the University of Florida, provides a public repository of mammography images in DICOM format. This dataset includes 3,568 mammograms, showcasing 1740 benign and 1828 malignant cases. Each image offers dual views (MLO and CC) and is accompanied by binary mask pictures that delineate the anomalies for precise ROI identification, allowing for exact pixel-level analysis of the irregular areas.

### Image preprocessing

Breast lesions within the breast tissue present a high risk of misdiagnosis<sup>9</sup>. These lesions are typically recognized by their shape, size, and boundary characteristics, and they appear lighter than the surrounding tissue<sup>82</sup>. This study's preprocessing steps include noise reduction using adaptive median filtering, morphological operations for boundary refinement, and Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance image clarity. Parameters for noise reduction include a window size of  $3 \times 3$  for initial filtering, which expands to  $7 \times 7$  if outliers persist. Morphological operations, specifically erosion and dilation, use a  $3 \times 3$  structuring element to remove non-breast regions while preserving critical features effectively. To monitor the development of significant lesions, radiologists often compare mammograms of the same breast taken over time. Identifying suspicious lesions over time involves spotting key control points like the edges of shapes caused by tissues, arteries, and ducts and analyzing the connections involving these points in sequential images. Adapting deep learning techniques to analyze such complex environments is crucial in medical image processing. Preprocessing of images is crucial to increase the contrast between normal and pathological images in mammograms by reducing the noise and improving the image clarity. Such preprocessing includes noise reduction, enhancements, scaling, and cropping of images, which is essential before implementing a learning-based algorithm to avoid overfitting. This paper does image enhancements such as averaging, median filtering, morphological operations, and CLAHE to increase the readability of mammogram images.

The region between the breast area is identified, affecting the adjacent tissue, using the global thresholding method. Probably, this refers to the application of a threshold value  $T$  on the mammography image  $I(x, y)$ , and the pixels with values greater than the threshold value are considered to be a breast and the other a non-breast, where  $I(x, y) > T$  represent the breast region. As an experienced radiologist, the threshold level here is about 30.

### Image acquisition

Compiling a vast assortment of mammography pictures presents substantial logistical challenges. Mammograms in their original state often include patient information, inherent interference, and fluctuations in illumination, which might result in a greater occurrence of incorrect positive results. Typically, these photos are transformed to grayscale in order to preserve important information throughout other modifications. The inherent limitations of mammography may have a negative impact on the effectiveness and education of deep learning algorithms. Moreover, a lack of breast photos sometimes results in overfitting during model training. Deep learning systems need a large number of annotated mammograms, but the task of annotating these pictures is both time-consuming and expensive<sup>83</sup>.

During data acquisition, morphological operations such as opening and closing eliminate labels and annotation tags. Erosion and dilation, fundamental morphological techniques, reduce and enlarge the elements of an image, respectively. Erosion is defined as the compilation of all points  $P$  such that the structuring element  $S$  fits within the set  $E$  when translated by  $P$ :

$$E \ominus S = \{P \mid (S)P \subseteq E\} \quad (12)$$

Dilation involves expanding  $E$  with the structuring element  $\hat{S}$ , organizing all displacements  $P$  so that  $SSS$  and  $EEE$  intersect at least once:

$$E \oplus S = \{P \mid (P) \cap E \neq \phi\} \quad (13)$$

These operations are crucial for preserving image features:

$$E \circ S = (E \ominus S) \oplus S \quad (14)$$

$$E \bullet S = (E \oplus S) \ominus S \quad (15)$$

Equations 17, 18 remove labeling and annotation markers from mammography images.

#### Image Noise Reduction and contrast enhancement

Post-acquisition, mammography images undergo processing to remove noise. Various noise types such as Gaussian, salt and pepper, speckle, and Poisson are reduced using an adaptive median filtering technique, resulting in a noise-free image  $I_{nf}$ . The process begins by determining the maximum and minimum pixel intensities within a window, denoted as  $MaxP_{\omega}$  and  $MinP_{\omega}$ , respectively. If the median  $Med_{\omega}$  is not an outlier, the process moves to the next verification step; otherwise, the window size is increased until the median is not an outlier or the maximum window size is reached. If  $I_{nf}$  falls within the range  $MinP_{\omega} < I_{nf} < MaxP_{\omega}$ , the image resolution remains unchanged; if not, the median value from step 2 is used.

Following noise reduction, image contrast enhancement is necessary as boundaries in noise-free mammography may fade. Edge-stopping functions prevent diffusion in areas of low image gradient, preserving sharp edges and enhancing detail. The edge-stopping function and the diffusion constant  $K$  are given by:

$$Ed_s(f) = \exp\left(\frac{1}{\alpha K^2} \times (1 - s^2)\right) \quad (16)$$

This function also helps maintain the object's edges while reducing noise. Histogram equalization is applied to boost global contrast, particularly enhancing relative contrast values crucial for imaging data. A histogram represents the distribution of digital image pixels across grey levels:

$$P(gl) = \frac{Px_n}{Pt_n} \times E_s(f) \quad (17)$$

This equation demonstrates anisotropic diffusion histogram equalization, estimating the likelihood of each grey level  $gl$  within the image.

#### Data-augmentation

CAD systems are capable of detecting and categorizing breast lesions. However, they frequently face high false-positive rates<sup>84</sup>. Conventional techniques that depend on traits created by humans are not sufficient for the precise identification of abnormalities<sup>85</sup>. The limited availability of annotated pictures for deep learning models results in overfitting<sup>86</sup> and difficulties in training owing to the significant expense and time needed for human annotation<sup>87</sup>. Previous methodologies had trouble handling the data's intricacy<sup>88</sup>. Figure 7 illustrates the suggested method for enhancing the detection and classification of breast cancer. This work utilizes data-augmentation methods such as geometric adjustments and patch extraction to enhance picture datasets, as shown in Table 3. Each augmentation technique was applied sequentially with the following parameters: sharpening intensity ranged from 0.5 to 2, Gaussian blur was applied with variances of 0.25, 0.5, 1, and 2, and rotations were set at intervals of 45° (e.g., 45°, 90°, 135°, and 180°). Skewing was performed in four directions: left, right, forward, and backward, with offsets of 10 pixels. Shearing transformations were limited to  $\pm 10^\circ$  along the X-axis and Y-axis. Flip transformations included left–right and top–bottom inversions to maximize dataset

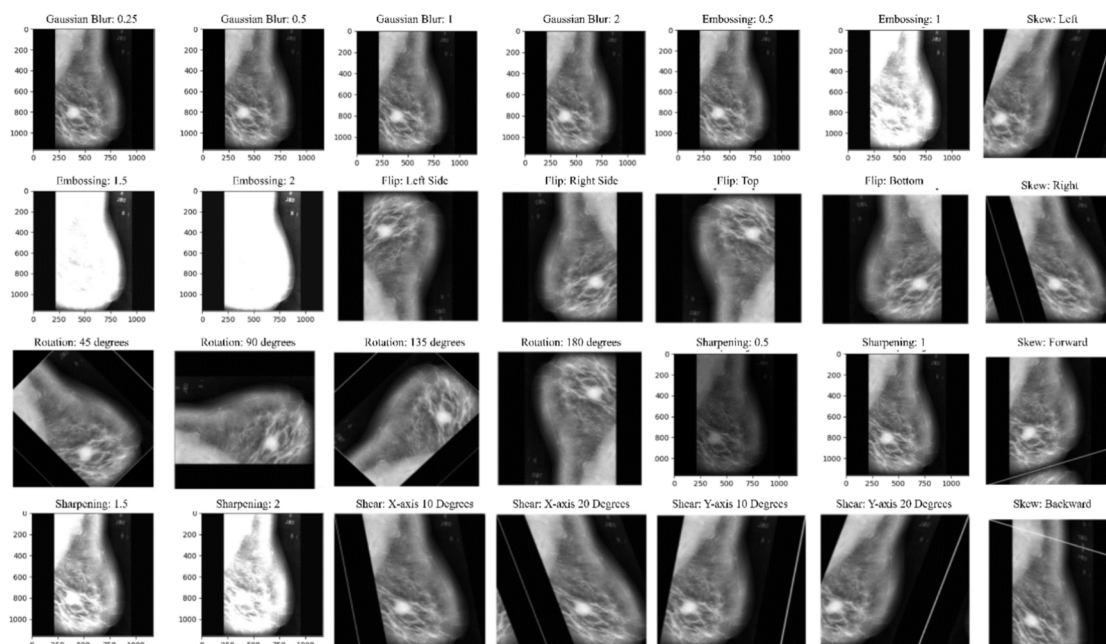
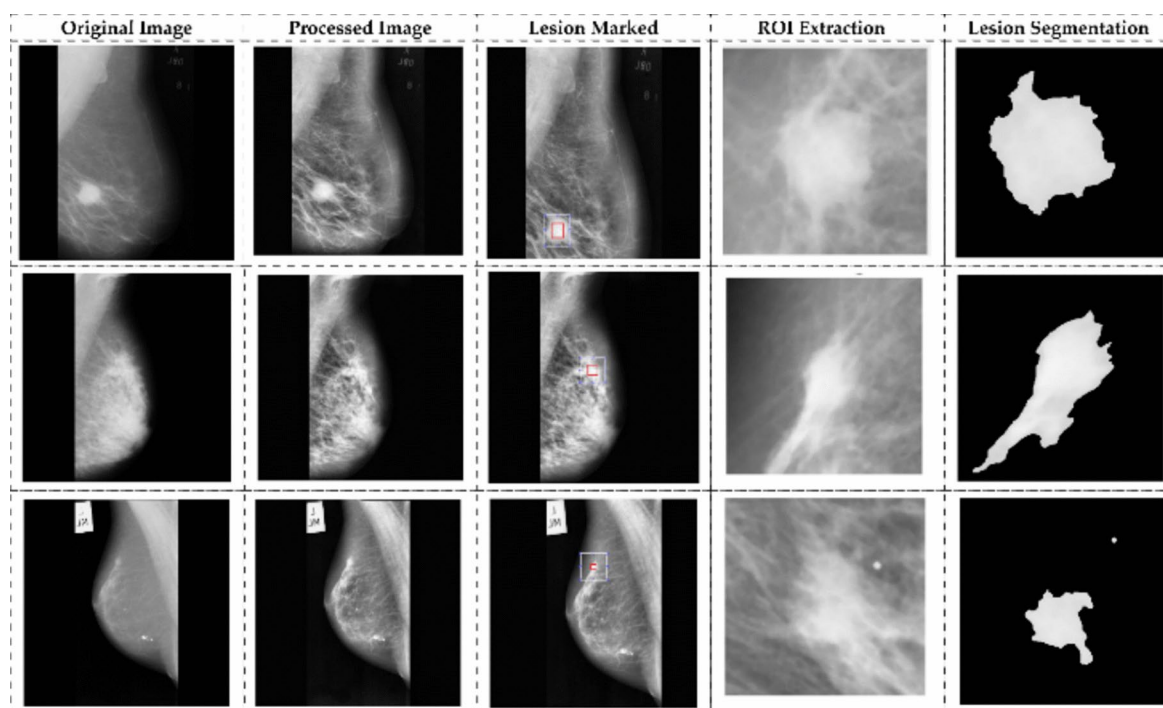


Fig. 7. Data augmentation.

Sr	Augmenting approaches	Invariance features
1	Sharpen	0.5, 1, 1.5, 2
3	Gaussian blur	0.25, 0.5, 1, 2
4	Rotation	45°, 90°, 135°, 180°
5	Skew	Left, Right, Forward, Backward,
6	Flipping	Left side, Right side, Top, Bottom,
7	Shear	X-axis 10° and Y-axis 10°

**Table 3.** Several methods for enhancing mammography by including invariance properties.



**Fig. 8.** Segmentation results.

variance. Patch extraction involved splitting images into non-overlapping tiles of  $256 \times 256$  pixels to improve spatial diversity during training.

#### *Lesion segmentation*

Breast segmentation is critical in mammography analysis because it removes artifacts, pectoral muscles, and background noise that could hinder anomaly detection. This process divides the mammogram into regions of interest (ROIs) for more effective feature extraction from breast anomalies. However, inaccurate segmentation can lead to poor feature selection, subsequently causing errors in classification. The segmentation of breast masses is typically performed based on attributes like intensity, texture, shape, and presence of anomalies. A method was introduced<sup>89</sup> to classify lesions by these characteristics since clinically precise mass boundaries are often unattainable. Another study<sup>90</sup> applied a thresholding technique to enhance edge detection in mammograms, identifying the Otsu method as the most effective for mammogram segmentation. Threshold segmentation is the prevalent method for isolating breast lesions; it categorizes lesion pixels based on a set threshold value. This approach is convenient for images exhibiting diverse grayscale levels. Due to its simplicity, rapid computation, and ease of implementation, threshold segmentation is a fundamental preliminary step before extracting image features and conducting further analysis. This research utilized the Otsu thresholding model<sup>91</sup> to delineate potential mass areas within mammograms. The Otsu method primarily utilizes the intensity data of the image, disregarding its shape and geometric properties, which facilitates the segmentation of the relevant regions for detailed pixel-level investigation. The method effectively binarizes the image by splitting its pixels into the object (foreground) and backdrop (background) based on the optimal threshold set by the adopted thresholding technique, as depicted in Fig. 8. This study set thresholds when performing segmentation on an image. The purpose is to divide pixels according to different thresholds to obtain several sets of pixels. Each set may provide information about the image's specific structure or texture, and each region's pixels share the same features while those of neighboring regions varied. The algorithm's precise process is as follows:

- The Otsu technique needs the threshold T to divide a mammogram with a grey value range of 298 [0, 1 – L] into two halves, C<sub>1</sub> and C<sub>2</sub>, with C<sub>1</sub> corresponding to the grey value range [0, T] and C<sub>2</sub> corresponding to the grey value range [T + 1, L – 1].
- Calculate the pixel average grey values μ<sub>1</sub> and μ<sub>2</sub> for areas C<sub>1</sub> and C<sub>2</sub>, as well as their respective proportions ω<sub>1</sub> and ω<sub>2</sub> to the whole image pixel, and then calculate the grey average value μ<sub>r</sub> for the entire image pixel described in Eq. 18.

$$\mu_r = \omega_1\mu_1 + \omega_2\mu_2 \tag{18}$$

- Compute the total variance σ<sub>r</sub><sup>2</sup> of regions C<sub>1</sub> and C<sub>2</sub>, described in Eq. 19.

$$\sigma_r^2 = \omega_1(\mu_1 - \mu_r)^2 + \omega_2(\mu_2 - \mu_r)^2 = \omega_1\omega_2(\mu_1 - \mu_2)^2 \tag{19}$$

T takes on consecutive values between [0, 1 – L]; hence, the value of σ<sub>r</sub><sup>2</sup> fluctuates continually. Larger the value of σ<sub>r</sub><sup>2</sup>, the greater the contrast between the foreground and background. The calculation stops when σ<sub>r</sub><sup>2</sup> reaches its maximum value. Suppose T<sub>0</sub> is the threshold that corresponds to the highest variance. In that case, this threshold is the optimal segmentation threshold because it effectively distinguishes the foreground from the background and has the lowest misclassification probability.

### Quantum-inspired binary grey wolf optimization

While the conventional GWO uses continuous values in the [0,1] range to guide the wolves' positions, the BGWO transforms each wolf's position into a binary format using a sigmoid function on the former<sup>92</sup>. To solve the unit commitment challenge, a BGWO version referring to the quantum approaches was launched and called Q-BGWO. The novel approach takes advantage of the quantum-inspired mechanism by which each wolf's position is binary and affected by a specific qubit vector that was combined with a quantum rotation gate. This transformation is made possible owing to an adjustment specified with the help of the γ-gate (θ) in Equation. The new solution is fundamentally different from the conventional GWO because while the latter adjusts wolves' positions following Equations A and C, the quantum approach makes the required adjustments based on the relevant qubit and the rotation angle θ that can be calculated with the help of two random probabilistic values, γ and ζ.

$$(\theta)_\alpha(t + 1) = (\zeta)_\alpha(\gamma)_\alpha \sum \left( \vec{X} \right)_\alpha \cdot (t) - \left( \vec{X} \right)_\alpha \cdot (t) \cdot 2\pi \tag{20}$$

$$(\theta)_\beta(t + 1) = (\zeta)_\beta(\gamma)_\beta \sum \left( \vec{X} \right)_\beta \cdot (t) - \left( \vec{X} \right)_\beta \cdot (t) \cdot 2\pi \tag{21}$$

$$(\theta)_\delta(t + 1) = (\zeta)_\delta(\gamma)_\delta \sum \left( \vec{X} \right)_\delta \cdot (t) - \left( \vec{X} \right)_\delta \cdot (t) \cdot 2\pi \tag{22}$$

$$(\zeta)_\alpha = (\lambda)_1 \cdot \pi \tag{23}$$

$$(\zeta)_\beta = (\lambda)_2 \cdot \pi \tag{24}$$

$$(\zeta)_\delta = (\lambda)_3 \cdot \pi \tag{25}$$

The α, β, and δ wolves are allocated random coefficients λ<sub>1</sub>, λ<sub>2</sub>, and λ<sub>3</sub>, respectively. The theta magnitude for the α wolf is represented by ζ<sub>α</sub>. These coefficients determine the rotation angles for each wolf's qubit vector, designated as Q. The rotations are performed based on Eqs. (26), (27), and (28), which specify the adjustments made to the location of each wolf following their respective rotation degrees.

$$(Q)_\alpha(t + 1) = (R)_\alpha \cdot ((\theta)_\alpha \cdot (t + 1)) \cdot (Q)_\alpha(t) \tag{26}$$

$$(Q)_\beta(t + 1) = (R)_\beta \cdot ((\theta)_\beta \cdot (t + 1)) \cdot (Q)_\beta(t) \tag{27}$$

$$(Q)_\delta(t + 1) = (R)_\delta \cdot ((\theta)_\delta \cdot (t + 1)) \cdot (Q)_\delta(t) \tag{28}$$

Q is a quantum state vector that forms a single qubit.

$$|(Q)_\alpha\rangle = (x)_\alpha|0\rangle + (y)_\alpha|1\rangle \tag{29}$$

$$|(Q)_\beta\rangle = (x)_\beta|0\rangle + (y)_\beta|1\rangle \tag{30}$$

$$|(Q)_\delta\rangle = (x)_\delta|0\rangle + (y)_\delta|1\rangle \tag{31}$$

Initially, the coordinates (x)<sub>α</sub>, (y)<sub>α</sub>, (x)<sub>β</sub>, (y)<sub>β</sub>, (x)<sub>δ</sub>, and (y)<sub>δ</sub> are all set to 0.5. Updates to the wolves' positions are then governed by the likelihood of their corresponding qubit vectors being in the state |1>. This method

aligns with the quantum behavior of the system, where each position change reflects the quantum probability calculations.

$$(X)_{\alpha}(t+1) = (X)_{\alpha}(t) \cdot (y_{\alpha}^2)(t+1) \tag{32}$$

$$(X)_{\beta}(t+1) = (X)_{\beta}(t) \cdot (y_{\beta}^2)(t+1) \tag{33}$$

$$(X)_{\delta}(t+1) = (X)_{\delta}(t) \cdot (y_{\delta}^2)(t+1) \tag{34}$$

The binary conversion of the wolves' positions in the BGWO involves straightforward thresholding operations. The position update equation is represented as:

$$\left(\vec{X}\right)_{\alpha}(t+1) = \begin{cases} 1 & \text{if } \left(\vec{X}\right)_{\alpha}(t) \geq y_{\alpha}^2(t+1) \\ 0 & \text{otherwise} \end{cases} \tag{35}$$

The following process converts probabilistic values into binary outcomes. The process compares each value with a threshold, which is squared. The threshold is derived from the probability of the qubit being in the state. After thresholding, a binary value of each feature is determined by a majority voting process between the alpha (α), beta (β), and delta (δ) wolves. The feature is chosen if most of these wolves assign a value of '1' to it. Otherwise,

the feature is rejected. Thus, a binary feature vector  $\left(\vec{X}\right)$  is created in order to represent the chosen features

that are essential for solving the problem. The binary feature vector  $\left(\vec{X}\right)$ , is then passed on through another process which involves finding the sigmoid function:

$$sigm(p) = \frac{1}{1 + e^{(-p)}} \tag{36}$$

where p ranges from 0 to 1. This function helps soften the thresholding process, providing a smoother transition between binary states.

The resultant values are then compared to a randomized threshold s, ranging from 0 to 1, to finalize the binary states of the features:

$$\left(\vec{X}\right) = \begin{cases} 1 & \text{if } F\left(\vec{X}\right) \geq s \\ 0 & \text{if } F\left(\vec{X}\right) < s \end{cases}$$

This step is crucial for ensuring that the feature selection process is robust and aligns with the probabilistic nature of quantum-inspired computations.

**Pseudo Code: Quantum Binary Grey Wolf Optimization (QBGWO)**

Input:
`N` —Total number of wolves (Population size) `D` —Dimensions of the search space `T` —Maximum number of iterations `f` —Fitness evaluation function
Output:

' $\alpha$ '—Optimal solution discovered
Procedure:
<pre> 1. Initialize Population: Create population 'p' with 'N' wolves, each with a D-dimensional binary vector For each wolf 'F_i' in 'p': For each dimension 'd' from 1 to 'D': 'F_i[d] = GenerateBinary()' 2. Assign Roles: Order wolves by their fitness using 'f' Designate roles: <math>\alpha</math> (Alpha), <math>\beta</math> (Beta), <math>\delta</math> (Delta), <math>\omega</math> (Omega) based on fitness order 3. Initialize Qubits: For each wolf 'F_i' in 'p': Initialize a qubit vector 'Q_i' with values ' 0&gt;' and ' 1&gt;' having equal probabilities 4. Update Positions Using Quantum Gates: For iterations 't' from 1 to 'T': Adjust control parameter 'a': 'a = 2 * (1 - t / T)' For each wolf 'F_i' in 'p': For top roles 'F_p' in {<math>\alpha</math>, <math>\beta</math>, <math>\delta</math>}: Set vectors 'A' and 'C': 'A = 2 * a * GenerateVector() - a' 'C = 2 * GenerateVector()' Compute distance 'K' from 'F_p' to 'F_i': 'K =  C * position of F_p at t - position of F_i at t ' Update position for the next iteration: 'F_i[t+1] = position of F_p at t - A * K' Update qubit vector 'Q_i' using rotation gates: For each dimension 'd' from 1 to 'D': Apply rotation gate 'R(<math>\theta</math>)' to 'Q_i[d]' Update 'Q_i[d]' based on the new angle '<math>\theta</math>' Use a sigmoid function for binary transitions: For each dimension 'd' from 1 to 'D': 'probability = Sigmoid((Q_<math>\alpha</math>[d] + Q_<math>\beta</math>[d] + Q_<math>\delta</math>[d]) / 3)' 'F_i[d][t+1] = 1 if GenerateRandom() &lt; probability else 0' Reassess fitness and readjust roles: For each wolf 'F_i' in 'p': Update <math>\alpha</math>, <math>\beta</math>, <math>\delta</math> based on new fitness evaluations 5. Finalize: Return the <math>\alpha</math> with the best fitness </pre>

### SqueezeNet- SVM

SqueezeNet<sup>93</sup> was selected as the foundational architecture for this study due to its compact size and low number of parameters, making it highly efficient for training on limited computational resources. SqueezeNet operates with significantly fewer parameters than larger models: AlexNet has 61 million, VGG-16 features 138 million, Inception-v3 holds 23 million, ResNet-50 has 25 million, and MobileNet comprises 4.2 million parameters. In contrast, SqueezeNet has only 1.2 million, enabling it to efficiently handle large datasets, such as mammography images, which require substantial computational power due to their high resolution and complexity. Additionally, its architecture is particularly effective in detecting subtle patterns, such as those found in breast cancer features. This streamlined architecture does not compromise on performance, maintaining comparable accuracy to its larger counterparts in image classification tasks<sup>94</sup>.

The fire module is the centremost part of a SqueezeNet, which is a small and flat convolutional network. As illustrated in Fig. 9, its purpose is to analyze the information to be retained efficiently. Specifically, it performs compression to reduce the number of input channels and expand to develop them. This mechanism plays a crucial role in processing mammography images by efficiently reducing redundant data while emphasizing critical features that are indicative of breast cancer. As the corresponding source describes, the compression feature reductions rely on  $1 \times 1$  convolution, and the expansion development comes with both  $1 \times 1$  and  $3 \times 3$  convolutions. In this way, the channel count is increased with  $1 \times 1$  and  $3 \times 3$  convolutions, which means bypass links can capture the finer details and changes in the input. Otherwise, very little useful information can be gained from it.

In addition, the SqueezeNet is equipped with sophisticated bypass links designed to improve the transmission of information across the network, as indicated in Fig. 10. These bypass mechanisms allow for improved retention and transmission of intricate details found in mammography images, ensuring that fine-grained patterns relevant to breast cancer detection are preserved throughout the layers. Specifically, these links allow for direct data transport to profiles from one level of detail to another, bypassing many layers. The corresponding bypass links can be defined in the following way: Let  $X$  be an input feature map and  $F$  be the corresponding transform form. The traditional bypass link can be formally expressed as  $y = x + F(x)$ ,  $x$  is the output of such a bypass link, while the operator “+” signifies that the two profiles are combined element-wise. In this way, any bypass link's input and output can be combined to provide the same fundamental details to the upgraded profile—including the differences. This property is critical for mammography, where subtle variations between normal and abnormal tissues are essential for accurate classification. Finally, SqueezeNet allows efficient and effective feature retention functions thanks to the Fire Module and several bypass link parameters. These unique design elements make SqueezeNet particularly suited for extracting distinct features from mammography images, as it balances computational efficiency with the ability to detect subtle patterns characteristic of breast cancer. In turn, the SGS model can identify fine-grained patterns and improve the accuracy and reliability of normal path charge detection.

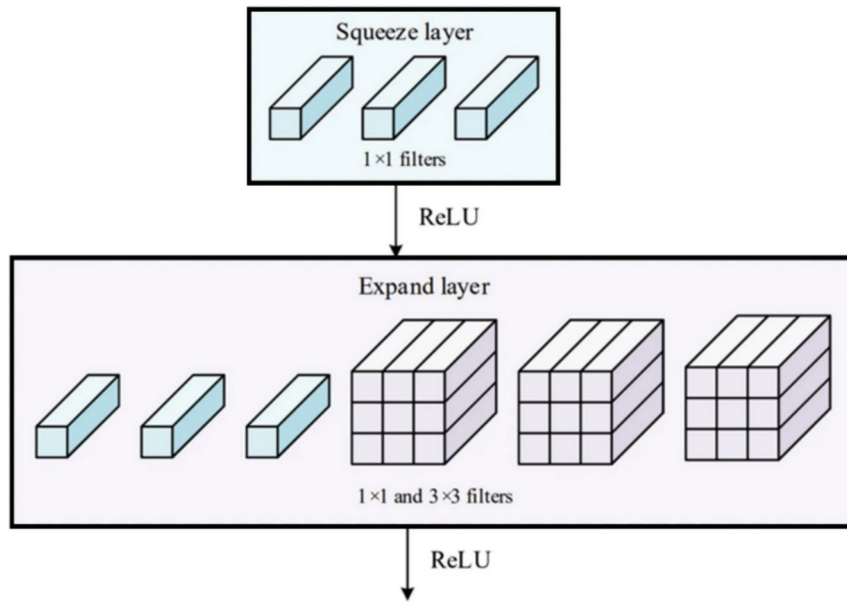


Fig. 9. Fire module structure.

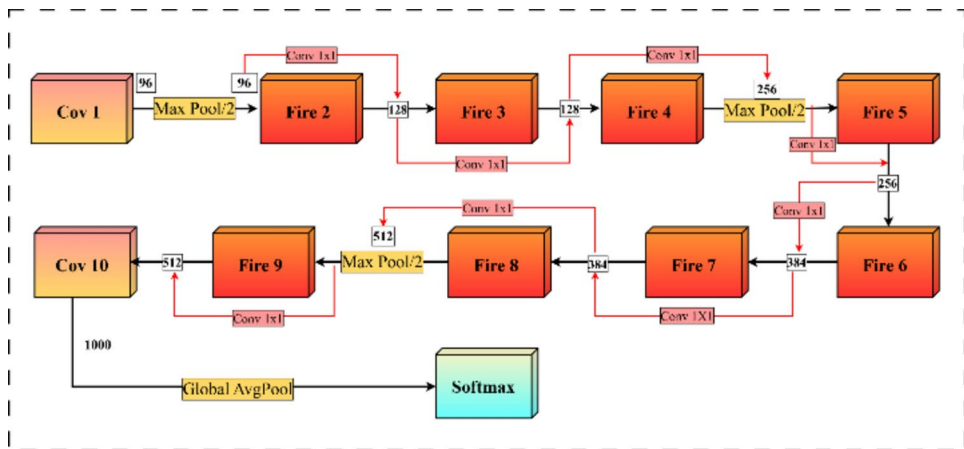


Fig. 10. SqueezeNet with complex bypass.

This study seeks to classify the pictures from breast cancer datasets. The images will be trained to characterize the traits generated by SqueezeNet. The Support Vector Machine is one of the most popular binary classifiers and is well known for its maximum precision of linear and non-linear classification issues. Additionally, it might be pretty efficient in multi-class conditions. The fundamental principle used in applying Support Vector Machines is to create one or more hyperplanes in the input space that might identify multiple classes appropriately. The linear determination hyperplane is expressed as: where  $w$  is the weight vector,  $b$  is the bias,  $w^T$  denotes the transpose of the weight vector  $x$  is the input vector that contains the parameters of an input feature. The prime purpose of the classification duties in SVM is to choose the perfect hyperplane so that the data points lead the class. It should also enhance the capacity for separation.

The margin is the gap between the nearest data points and the hyperplanes for all classes. It was noted that a larger margin indicates that newer data points have less impact on a classifier. There is no real variance due to details as the points are defined as optimal separation<sup>95</sup>. Thus, SVMs work to minimize the margin exactly and to ensure good classification between class and data transition. Optimal changes in the value of  $w$  and  $b$  normally do this.

$$\begin{cases} \min_{w,b} \frac{1}{2} w^T w + C \sum_{i=1}^n \sigma_i \\ \text{subject to } y_i (w^T x_i + b) \geq 1 - \sigma_i \end{cases} \quad (37)$$

The input vector is denoted as  $x_i$ , and the corresponding output label is denoted as  $y_i$ .

In Support Vector Machines,  $\sigma_i$  is the slack variable, allowing certain training samples to be mishandled to a limited extent. It includes the C parameter that measures the weight assigned to punishing inappropriate classification. It also supports finding the best balance between incorporating the highest margin within two classes and the ability to classify each of the training samples appropriately. It serves to give the penalty value for misclassification. The variable n is the number of all training samples.  $\frac{1}{2}w^T w$  is the first component of the target function, and its role is to maximize the margin, while the second one,  $C \sum_{i=1}^n \beta_i \sigma_i$ , is responsible for handling the punishment value for the misclassified training samples. The Lagrangian dealing with this optimisation task can be presented in the following manner:

$$L(w, b, \sigma, \beta) = \frac{1}{2}w^t w + C \sum_{i=1}^n \sigma_i - \sum_{i=1}^n a_i [y_i (w^t x_i + b) - 1 + \sigma_i] - \sum_{i=1}^n \beta_i \sigma_i \quad (38)$$

In order to find optimal values for w, b, and  $\zeta$  in an SVM formulation by Lagrange multipliers  $\alpha = (\alpha_1, \dots, \alpha_n)$  and  $\beta = (\beta_1, \dots, \beta_n)$  have to set up and solve a system of equations by taking partial derivatives of Lagrangian L concerning these variables and setting them to zero, this calculative process can be described in the following way:

$$w = \sum_{i=1}^n \alpha_i y_i x_i \quad (39)$$

The dual Lagrangian may be obtained by using the Karush–Kuhn–Tucker condition<sup>49</sup>:

$$\begin{cases} \max_a L(a) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i^T x_j \\ \text{Subject to } \sum_{i=1}^n \alpha_i y_i = 0 \end{cases} \quad (40)$$

In non-linear SVM, kernel functions transform input data into a higher-dimensional space, enabling the algorithm to identify a separating hyperplane in this new space, even if the data cannot be separated linearly in the original space. The optimization issue in non-linear SVM is similar to that of linear SVM, but it involves using a kernel function. The optimization problem may be expressed as follows:

$$\begin{cases} \max_a L(a) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j G(x_i, x_j) \\ \text{Subject to } \sum_{i=1}^n \alpha_i y_i = 0, 0 < \alpha_i < C \end{cases} \quad (41)$$

$G(x_i, x_j)$  represents the kernel function in this context. Among the most commonly utilized kernel functions is the Gaussian kernel:

$$G_{gau}(x_i, x_j) = \exp\left(\frac{\|x_i - x_j\|^2}{2\varphi^2}\right) \quad (42)$$

$\varphi$  is a hyperparameter that regulates the level of smoothness in the decision boundary. A polynomial kernel is a function used in machine learning and support vector machines. It is a mathematical function that calculates the similarity between two data points in a higher-dimensional space:

$$G_{pol}(x_i, x_j) = (x_i^T + h)^d \quad (43)$$

d represents the degree of the polynomial. The symbol “h” represents a variable that determines the extent of the polynomial expansion.

### SQSVM- optimization

It is crucial to discretely optimize the SVM parameters for optimal performance of SVM classifiers, particularly in complex image classification tasks. The optimization process is based on Grey Wolf Optimizer, which is a metaheuristic algorithm based on the social hierarchy and hunting behavior of Grey Wolves. The process involves several step-by-step results, which are very technical, to ensure the efficiency of the SVM operation.

First, the initialization step involves the creation of a randomly organized population of Grey Wolves—where each Grey Wolf is indicative of a specific array containing a 2-dimensional representation of a particular SVM parameter. This array contains the penalty factor, C, which is the most significant SVM parameter to optimize, and kernel parameters, when applicable, present critical information to the working of the classifier, which is conserved within the dimension. A separate step is applied to define the respective operational parameters of the Quantum-enhanced Binary Grey Wolf Optimizer (Q-BGWO), the improved version of the Grey Wolf Optimizer incorporating principles of quantum mechanics. The Q-BGWO introduces quantum behaviors, such as superposition and probabilistic adjustments, allowing the optimizer to explore the search space more comprehensively and avoid premature convergence. This approach significantly enhances the algorithm’s ability to handle high-dimensional, complex data, such as mammography images in breast cancer classification.

The utilization of Q-BGWO in this context is motivated by its ability to balance exploration and exploitation during the optimization process effectively. By integrating quantum mechanics principles, such as quantum rotation gates and qubit state vectors, Q-BGWO achieves a broader search capacity and faster convergence. This ensures the selection of optimal SVM parameters, such as CCC and kernel parameters, with higher accuracy and robustness, which is crucial for improving breast cancer classification performance. Compared to traditional metaheuristic algorithms, Q-BGWO excels in navigating complex optimization landscapes, reducing the risk of local optima, and delivering precise feature optimization.

Next, the respective operational parameters of the Quantum-enhanced Binary Grey Wolf Optimizer are comprised of the population size, the number of iterations, and the preannouncement requirement to commence. The operational parameters of the SqueezeNet-SVM, which is the SVM model, are initialized. After initialization, the fitness function is used to determine the effectiveness of the operation of each configuration based on the classification accuracy of the SVM model achieved by the cross-validation test. Cross-validation testing is meant to evaluate the ability of the model to generalize, divided into the folds present in K-fold scales of 5, 10, and 15. This mitigates overfitting and is meant to test the applicability of the SVM on datasets. Correspondingly, feature selection operations are optimized by applying the constant  $q = 1 - e$ , where  $q$  varies from 0, increasing to 1 gradually, which controls the exponential decay of the number of features to ensure optimal computational efficiency and accuracy. The final configuration of the number of features that participates in the modeling, denoted as  $R$ , and the total number of features for selecting the data subsets are selected maximally to focus the model on the most probable, most informative attributes.

Subsequently, the fitness of the classified operations of each population of Grey Wolves is tested based on the SVM classification accuracy on the validation folded datasets. At the end of each iteration, the Grey Wolves classification is conducted and evaluated – with the top three best performers ranked as the alpha, beta, and delta associated with the most optimal sets. Using this hierarchical ranking to change the dynamic space of all the wolves in the population ensures the optimal parameters are identified. The iterative process is continued until the optimal parameter configurations are identified or the operation of parameter learning ability on the classifier, which can be evaluated, has been completed. The final chosen configurations are applied in the QSVM, and the testing of the operational efficiency of the classifier on the testing set confirms the analytical applications in real-time tasks.

### Performance evaluation matrices

Various performance metrics typical of many preceding studies are applied to provide a verifiably accurate assessment of the Q-BGWO-QSVM model and theBGWO-QSVM model. It should be mentioned that evaluation of a model's effectiveness includes testing it not only based on its training scores but also through its ability to provide accurate performance on the new data in the future. A set of evaluative indicators is selected to assess the performance whenever a new model is developed.

Sensitivity  $\left(\frac{TP}{TP+FN} \times 100\%\right)$  : is an evaluation of a model's ability to identify the positive cases of predicting a target class attribute or variable in such a way, measuring the model's ability to recognize all of the positive instances.

Specificity  $\left(\frac{TN}{TN+FP} \times 100\%\right)$  : is defined by. This method is prescript to sensitivity, as it measures a model's ability to identify non-positive or negative instances. The method is applied to assess how much a model's positive and negative predictions are accurate.

Accuracy  $\left(\frac{TP+TN}{TP+TN+FP+FN} \times 100\%\right)$ —Specified as evaluating the ability to make correct predictions and outcomes on both positive and negative instances. The method relies on performance evaluation in terms of precision and recall.

Precision  $\left(\frac{TP}{TP+FP} \times 100\%\right)$  : is defined as being focused on measuring exactness in positive predictions.

F1-Score  $\left(\frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\%\right)$  : Defines a harmonic mean of precision and recall evaluation and is useful when a class distribution imbalance is visible. In this way, whenever either precision or recall is strangely high and the other—low, the harmonious mean provides a balanced metric that indicates the advantages and disadvantages of both methods.

Matthews correlation coefficient  $\left(\frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \times 100\%\right)$  : is defined by the formula. The method is particularly applicable when class sizes are imbalanced and unequal. It offers a balanced measure that includes sensitivity, precision, true positives, false positives, and false negatives.

### Experimental results

This section presents a detailed examination of the experimental results of studying the datasets utilizing mammography data. This research assesses the proposed technique's effectiveness in the breast tumor categorization field.

### Implementation details

This study used the MATLAB platform and GPU acceleration to analyze the most efficient classification algorithms: the Q-BGWO based on the SVM. The experimental settings were related to a high-performance fusion: the Intel Core i7 13th generation central processing unit, 32 GB RAM, and a robust NVIDIA GeForce RTX 3090 Ti graphics processing unit. At the same time, the Q-BGWO was altered in terms of the SVM parameters, which was explained by a higher appropriateness for the classification task.

Cross-validation through 5, 10, and 15 folds ensured the model assessment's reliability using different datasets. In this regard, the optimizer, functioning within  $[0,1]$  as the initial domain and with a consistent population size of 10, can be described as the one dynamically adjusting the SVM's penalty parameter  $C$  for obtaining a coherent level of complexity and the trade-off concerns related to the differences in training errors. Various types of kernel applications, such as Linear, RBF, Polynomial, and Sigmoid, were applied to provide opportunities for both types of data separations, linear and non-linear. Moreover, the weighting parameters, such as  $\alpha = 0.97$  and  $\beta = 0.03$ , were tuned for the optimal balance between exploration and exploitation and to increase the algorithm's convergence rates using all the best-fitted conditions specified. The performance was measured by the best-fitted optimal iterations varying from 7 to 60 and the corresponding fitness values ranging from 0.0135 to 0.1077—the critical pictures of the classifier's highest precision. Finally, the qubit vector set was created around  $[0.5, 0.5]$  for  $\alpha$ -the wolf to implement the most optimal approach for balanced quantum computing and allow the classifier to obtain high classification accuracy as a ready solution with robust specificity problems.

### Assessment of Q-GBGWO compared to traditional bio-inspired optimization methods

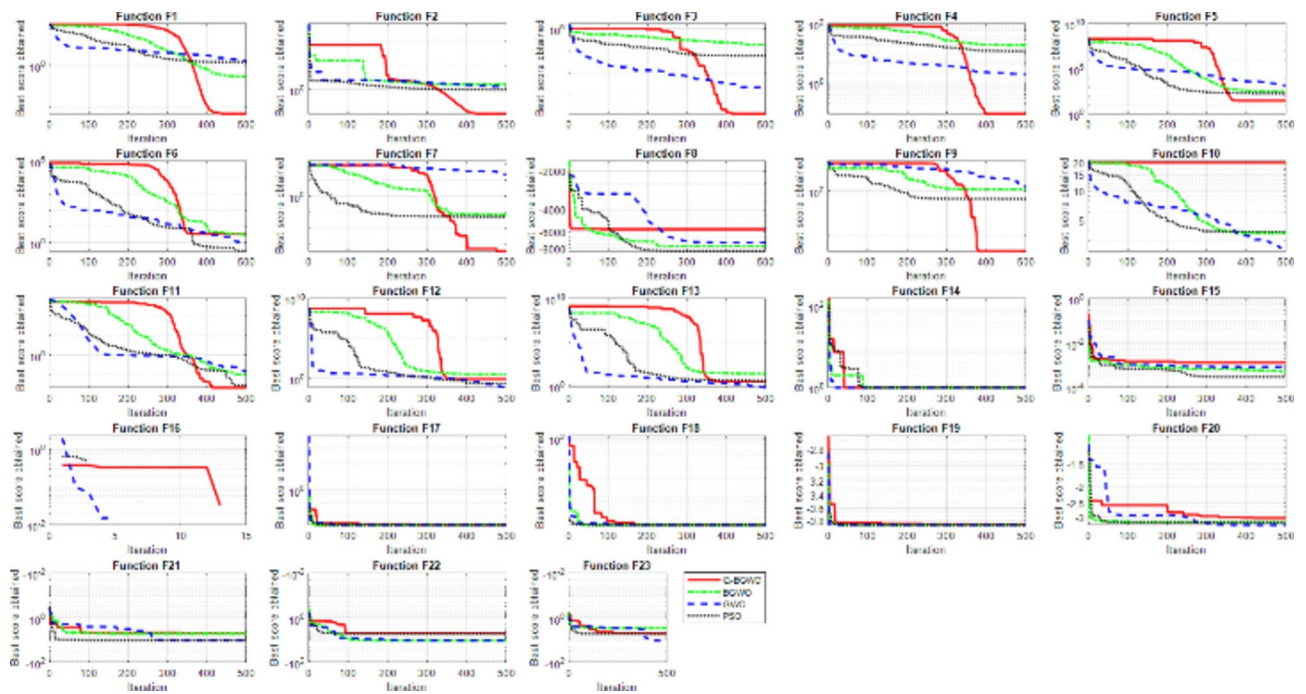
This research assesses the efficacy of the Quantum Grey Binary Grey Wolf Optimizer (Q-GBGWO) in comparison to conventional bio-inspired algorithms such as the Binary Grey Wolf Optimizer (BGWO), Grey Wolf Optimizer (GWO), and Particle Swarm Optimisation (PSO). For this comparison, we performed testing using the benchmark functions<sup>17,96</sup>. Every algorithm was executed following the requirements given in their respective foundational texts. The findings in Table 4 demonstrate that Q-GBGWO consistently outperforms other methods, exhibiting greater performance across numerous tests. Although the specific specifics of how each function is carried out align with the processes and results shown in Table 4, Q-GBGWO has demonstrated significant efficacy across several benchmarks. While BGWO, GWO, and PSO have shown competence in certain activities, Q-GBGWO has exhibited comparable or higher effectiveness in most assessed functions. The forthcoming figures, such as Fig. 11, will illustrate the convergence trends of Q-GBGWO. These trends will demonstrate the adaptability and effectiveness of Q-GBGWO during both the initial and later phases of the iterations. They will emphasize the well-maintained balance between exploration and exploitation in optimization.

### Cross-validation results

This study used several cross-validation (CV) methods for optimal outcomes and reliable validation. A tenfold cross-validation was used due to its continuous delivery of excellent outcomes. The whole dataset was randomly partitioned into 10 sections of equal size. Throughout the training phase, nine subsets were used, with the final subset designated as the validation set. The technique was repeated 10 times, with each subset utilized for validation once, in order to achieve comprehensive validation across all subgroups. Furthermore, for accurate validation, we conducted fivefold and 15-fold cross-validation and documented their respective results in Fig. 12.

Function	Q-BGWO	BGWO	GWO	PSO
F1	5.62E-06	0.169128234	1.183301971	0.603106942
F2	1.24E-05	10.37589831	3.856437111	0.326454917
F3	30.0342881	3878.161525	156.6945295	300.625596
F4	0.062406547	24.2152155	1.812399255	9.875028467
F5	28.93312121	112.934761	994.4792144	198.3568905
F6	3.941060642	0.628962914	2.099735534	0.020609621
F7	0.003583321	0.142225758	27.40623994	0.128575963
F8	-5753.495034	-9596.391775	-5834.328807	-8260.117265
F9	16.55299471	103.5559957	194.1249357	57.74076835
F10	19.96397213	2.590710518	2.874960771	1.488599648
F11	0.03556385	0.345515711	0.157819171	0.6376282
F12	0.308273215	3.61868785	0.050510421	0.295120471
F13	2.807176382	6.174534011	0.301430388	0.230653105
F14	0.998050756	0.998003838	8.840835963	0.998003838
F15	0.001303212	0.000376244	0.000834597	0.001223173
F16	-1.031622045	-1.031628453	-1.031628453	-1.031628453
F17	0.400527015	0.397887358	0.397887358	0.397887358
F18	3.000000026	3	3	3
F19	-3.854720743	-3.862782148	-3.862782148	-3.862782148
F20	-3.090069499	-3.321995172	-3.321995167	-3.20310205
F21	-4.979597891	-5.055197729	-5.100772139	-5.10077214
F22	-4.991267477	-5.128822797	-10.40294057	-10.40294057
F23	-5.101072966	-10.53640982	-10.53640982	-10.53640982

**Table 4.** Q-GBGWO Comparative Analysis on Bench-Mark Functions.



**Fig. 11.** Examination of convergence behaviour for Q-BGWO on standard test functions.

The comparative analysis focusing on the Q-BGWO-SQSVM model across three different cross-validation schemes—fivefold, tenfold, and 15-fold—as revealed through the study of diverse medical image datasets such as MIAS, INbreast, DDSM, and CBIS-DDSM, highlights its exceptional performance. This proposed model has demonstrated top-tier results, consistently outperforming other evaluated models in every key metric: Accuracy (ACC), Sensitivity (SEN), Specificity (SPC), Precision, F1 Score, and Matthews Correlation Coefficient (MCC).

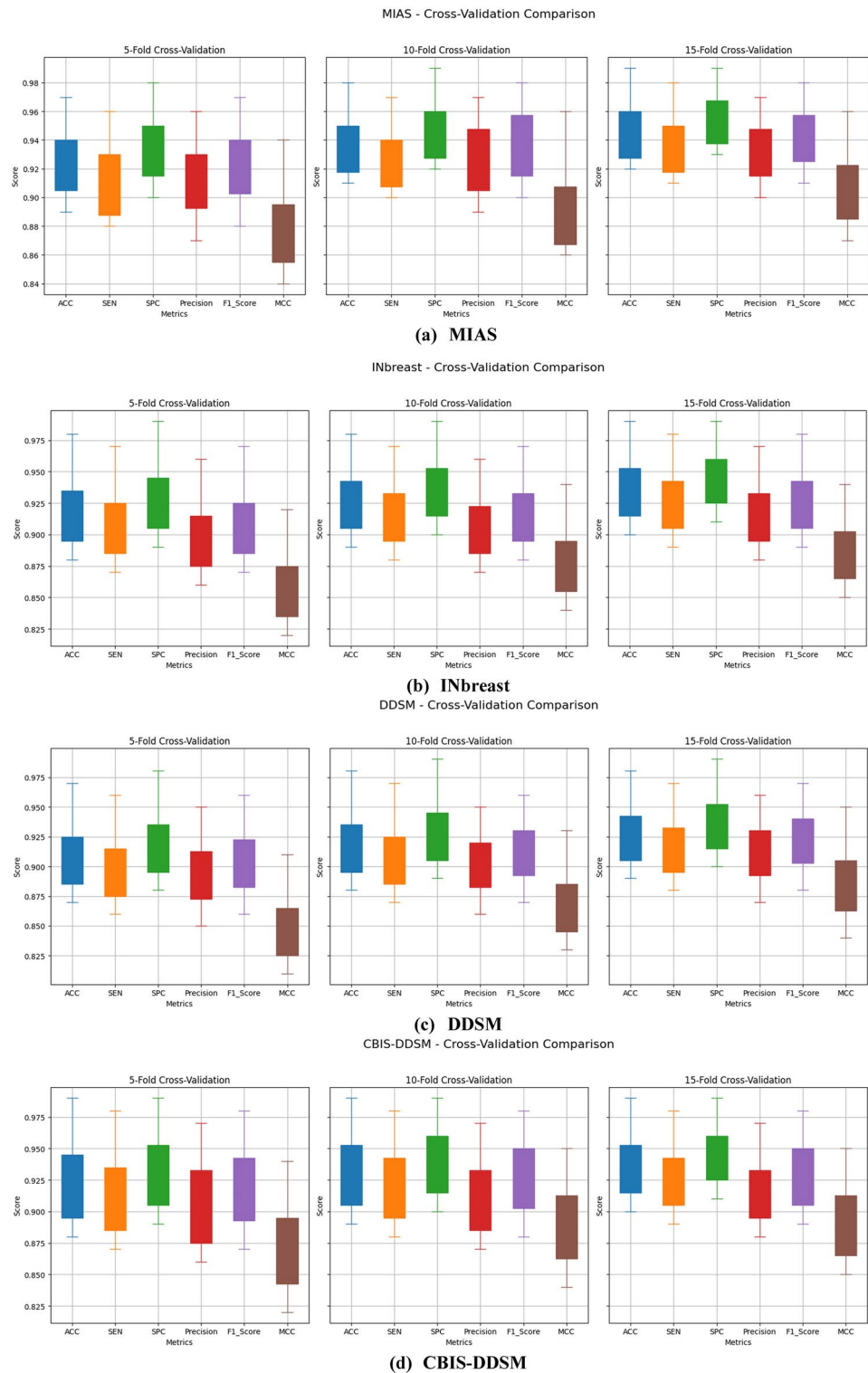
The Q-BGWO-SQSVM model achieved the highest values for ACC, SEN, SPC, Precision, F1 Score, and MCC and showed remarkable stability and reliability across all datasets and metrics. Its performance underscores its suitability for handling intricate medical image analysis tasks, making it an ideal candidate for applications requiring high precision and reliability in clinical and diagnostic settings.

Throughout the evaluation process over various cross-validation folds, the Q-BGWO-SQSVM model exhibited robustness and adaptability, effectively addressing dataset-specific nuances and maintaining consistent performance. This adaptability is crucial for applications in medical image analysis, where varying data characteristics can significantly impact the efficacy of machine learning models.

The insights gleaned from this study suggest that the Q-BGWO-SQSVM model holds significant promise for enhancing diagnostic accuracy and clinical decision-making in healthcare. Further validation on independent datasets and continuous fine-tuning are recommended to fully capitalize on its capabilities and ensure its reliability and applicability in real-world medical imaging scenarios. This approach reaffirms the importance of advanced machine learning techniques in driving forward the fields of medical diagnostics and personalized medicine.

### Performance comparison of proposed algorithm with ML models across different cross-validation schemes

For the MIAS dataset, the 5-Fold Cross-Validation results showcase varying performance levels among different models. The Q-BGWO-SQSVM model exhibits the highest performance, achieving an accuracy of 0.97, sensitivity of 0.96, and specificity of 0.98. These results underscore its superior capability in effectively diagnosing medical images compared to traditional models like SVM and newer variants such as SQSVM and BGWO-SQSVM, which also perform robustly but do not reach the peak metrics of the Q-BGWO-SQSVM. Expanding the analysis to 10-Fold Cross-Validation, there is a noticeable improvement across all models, indicative of the increased data utilization enhancing model training and testing robustness. The Q-BGWO-SQSVM continues to lead with exemplary metrics, including a near-perfect accuracy of 0.98. This enhancement across the board suggests that models can generalize better with more comprehensive cross-validation, with less variance in performance due to the increased data exposure during training. With the 15-Fold Cross-Validation, the trends observed in the 10-Fold validation continue, with even slight metric increments for models like SVM, SQSVM, and BGWO-SQSVM. The Q-BGWO-SQSVM reaches peak performance metrics, maintaining top scores with an accuracy of 0.99, reinforcing its robustness and consistency across different validation folds. This highest level of validation provides a more stringent test of the models' abilities to maintain performance, suggesting the Q-BGWO-SQSVM's potential utility in clinical settings where high reliability is paramount.



**Fig. 12.** Cross-validation results.

Across Table 5, the progression from 5-Fold to 15-Fold Cross-Validation demonstrates a general trend of increasing performance, highlighting the importance of model training and validation on larger subsets for stability and accuracy. The Q-BGWO-SQSVM not only consistently outperforms other models but also shows minimal performance degradation as the complexity of the validation scheme increases. This suggests that the model's architecture and underlying algorithms are well-suited for tackling complex patterns in medical imaging data, potentially due to superior feature extraction capabilities and optimization strategies that mitigate overfitting more effectively than its counterparts. These insights into the model performances on the MIA

Model	ACC	SEN	SPC	PRE	F1-S	MCC
5-CV						
RF	0.85	0.84	0.86	0.85	0.85	0.83
KNN	0.87	0.88	0.89	0.86	0.87	0.85
DT	0.83	0.81	0.85	0.84	0.82	0.80
NB	0.86	0.87	0.88	0.85	0.86	0.84
LR	0.88	0.87	0.90	0.89	0.88	0.86
AB	0.84	0.83	0.87	0.86	0.85	0.82
GB	0.89	0.90	0.91	0.88	0.89	0.87
SVM	0.89	0.88	0.90	0.87	0.88	0.85
BGWO-SQSVM	0.92	0.91	0.94	0.93	0.92	0.90
Q-BGWO-SQSVM	0.97	0.96	0.98	0.97	0.97	0.95
10-CV						
RF	0.87	0.86	0.89	0.88	0.87	0.85
KNN	0.88	0.89	0.90	0.87	0.88	0.86
DT	0.84	0.82	0.86	0.85	0.83	0.81
NB	0.87	0.88	0.89	0.86	0.87	0.85
LR	0.90	0.89	0.92	0.91	0.90	0.88
AB	0.85	0.84	0.88	0.87	0.86	0.83
GB	0.91	0.92	0.93	0.90	0.91	0.89
SVM	0.92	0.91	0.93	0.92	0.92	0.90
BGWO-SQSVM	0.94	0.93	0.96	0.95	0.94	0.92
Q-BGWO-SQSVM	0.98	0.97	0.99	0.98	0.98	0.96
15-CV						
RF	0.88	0.87	0.90	0.89	0.88	0.86
KNN	0.89	0.90	0.91	0.88	0.89	0.87
DT	0.85	0.83	0.87	0.86	0.84	0.82
NB	0.88	0.89	0.90	0.87	0.88	0.86
LR	0.91	0.90	0.93	0.92	0.91	0.89
AB	0.86	0.85	0.89	0.88	0.87	0.84
GB	0.92	0.93	0.94	0.91	0.92	0.90
SVM	0.93	0.92	0.94	0.93	0.93	0.91
BGWO-SQSVM	0.95	0.94	0.97	0.96	0.95	0.93
Q-BGWO-SQSVM	0.98	0.96	0.99	0.97	0.98	0.97

**Table 5.** MIAS Dataset Cross-Validation Performance Metrics.

dataset across varying cross-validation schemes illustrate the significant impact of sophisticated modeling techniques and rigorous validation on enhancing diagnostic accuracies in medical image analysis.

In the INbreast dataset under 5-Fold Cross-Validation, the performance metrics vary across models, with the Q-BGWO-SQSVM again showing the highest performance, achieving nearly perfect metrics that emphasize its robust feature extraction and classification capabilities. This model achieves remarkable accuracy, sensitivity, and specificity, demonstrating its effectiveness in handling INbreast's imaging complexities compared to more conventional models such as SVM and RF, which exhibit competitive but slightly lower performance metrics. The 10-Fold Cross-Validation results for the INbreast dataset indicate improvements in model performance, which can be attributed to a more comprehensive utilization of the dataset during the training phase. The Q-BGWO-SQSVM continues to excel with the highest scores, illustrating its consistent performance and reliability. This trend highlights the model's capacity to effectively leverage more extensive training data, enhancing its predictive accuracy and generalizability. Moving to 15-Fold Cross-Validation, there is a further enhancement in performance metrics across all models, particularly noted in the Q-BGWO-SQSVM, which maintains its superior metrics amidst more rigorous validation. This level of validation tests the models under stricter conditions, confirming the Q-BGWO-SQSVM's superior adaptability and robustness in clinical diagnostic settings.

Throughout the INbreast dataset analysis in Table 6, increasing the cross-validation folds consistently improves all models' performance, with the Q-BGWO-SQSVM model maintaining the lead. This pattern supports the inference that higher cross-validation folds reduce overfitting and provide a more reliable estimate of model performance, especially evident in the advanced modeling and optimization techniques employed by the Q-BGWO-SQSVM.

For the DDSM dataset, models exhibit an upward trend in performance metrics as the number of folds increases from 5 to 15 as presented in Table 7. The Q-BGWO-SQSVM showcases exemplary performance across all metrics, with other models like SVM and GB also showing significant improvements. The consistent

Model	ACC	SEN	SPC	PRE	F1-S	MCC
5-CV						
RF	0.84	0.83	0.86	0.85	0.84	0.81
KNN	0.86	0.85	0.87	0.86	0.86	0.84
DT	0.82	0.80	0.85	0.83	0.81	0.78
NB	0.85	0.83	0.88	0.86	0.84	0.82
LR	0.87	0.86	0.89	0.88	0.87	0.85
AB	0.85	0.84	0.87	0.86	0.85	0.82
GB	0.88	0.87	0.90	0.89	0.88	0.86
SVM	0.89	0.88	0.91	0.90	0.89	0.87
BGWO-SQSVM	0.93	0.92	0.94	0.93	0.92	0.91
Q-BGWO-SQSVM	0.96	0.95	0.97	0.96	0.96	0.94
10-CV						
RF	0.85	0.84	0.87	0.86	0.85	0.83
KNN	0.87	0.86	0.89	0.88	0.87	0.85
DT	0.83	0.81	0.86	0.84	0.82	0.80
NB	0.86	0.85	0.88	0.87	0.86	0.84
LR	0.88	0.87	0.90	0.89	0.88	0.86
AB	0.86	0.85	0.88	0.87	0.86	0.84
GB	0.90	0.89	0.92	0.91	0.90	0.88
SVM	0.91	0.90	0.93	0.92	0.91	0.89
BGWO-SQSVM	0.95	0.94	0.96	0.95	0.95	0.93
Q-BGWO-SQSVM	0.98	0.97	0.99	0.98	0.98	0.96
15-CV						
RF	0.86	0.85	0.88	0.87	0.86	0.84
KNN	0.88	0.87	0.90	0.89	0.88	0.86
DT	0.84	0.82	0.87	0.85	0.83	0.81
NB	0.87	0.86	0.89	0.88	0.87	0.85
LR	0.89	0.88	0.91	0.90	0.89	0.87
AB	0.87	0.86	0.89	0.88	0.87	0.85
GB	0.91	0.90	0.93	0.92	0.91	0.89
SVM	0.92	0.91	0.94	0.93	0.92	0.90
BGWO-SQSVM	0.96	0.95	0.97	0.96	0.95	0.93
Q-BGWO-SQSVM	0.99	0.98	0.99	0.98	0.98	0.97

**Table 6.** INbreast Dataset Cross-Validation Performance Metrics.

enhancement in sensitivity and specificity indicates these models' effectiveness in distinguishing between normal and pathological findings in DDSM's complex imaging data.

In the CBIS-DDSM dataset, the Q-BGWO-SQSVM's dominance continues, with peak performance metrics in the 15-Fold Cross-Validation setting, demonstrating near-perfect scores as presented in Table 8. This model's effectiveness is particularly noteworthy in a dataset known for its challenging diagnostic images, proving its potential in high-stakes medical environments where diagnostic accuracy is critical.

For both DDSM and CBIS-DDSM datasets, increasing the validation folds improves the metrics across models and reinforces the benefits of robust cross-validation in achieving high generalizability and reliability. The consistent performance of the Q-BGWO-SQSVM across these datasets underscores its superior architectural and algorithmic framework, which is ideal for complex and high-dimensional medical imaging tasks. These evaluations reveal how advanced machine learning models, especially the Q-BGWO-SQSVM, adapt and perform under rigorous testing conditions, offering insights into their practical applications in medical imaging and diagnostics.

### AUC comparison proposed algorithm with traditional ML algorithms

The ROC curve is a critical evaluation tool in assessing the effectiveness of classifiers in distinguishing between patients with a disease and those without it. The Area Under the Curve (AUC) is a crucial metric derived from the ROC curve, with higher AUC values indicating better discrimination capabilities of the model. In our study, we present ROC curves for four datasets: MIAS (Fig. 13A), INbreast (Fig. 13B), DDSM (Fig. 13C), and CBIS-DDSM (Fig. 13D), evaluated using fivefold, tenfold, and 15-fold cross-validation (CV).

On the ROC curve, the True Positive Rate (TPR) is depicted on the y-axis, while False Positive Rate (FPR) is represented on the x-axis. AUC values close to 1 signify that the classifier can accurately differentiate between positive and negative class instances, achieving perfect discrimination. Conversely, an AUC of 0 would indicate that the classifier incorrectly classified all cases. AUC values between 0.5 and 1 suggest that the classifier can

Model	ACC	SEN	SPC	PRE	F1-S	MCC
5-CV						
RF	0.86	0.84	0.87	0.85	0.85	0.83
KNN	0.87	0.85	0.88	0.86	0.86	0.84
DT	0.83	0.82	0.84	0.83	0.82	0.81
NB	0.84	0.83	0.85	0.84	0.84	0.82
LR	0.88	0.87	0.89	0.87	0.88	0.86
AB	0.85	0.83	0.86	0.85	0.84	0.83
GB	0.89	0.87	0.91	0.90	0.89	0.88
SVM	0.90	0.89	0.92	0.91	0.90	0.89
BGWO-SQSVM	0.94	0.93	0.95	0.94	0.94	0.93
Q-BGWO-SQSVM	0.97	0.96	0.98	0.97	0.97	0.96
10-CV						
RF	0.87	0.86	0.89	0.88	0.87	0.85
KNN	0.89	0.87	0.90	0.88	0.89	0.87
DT	0.85	0.84	0.86	0.85	0.85	0.83
NB	0.86	0.85	0.88	0.87	0.86	0.84
LR	0.90	0.89	0.91	0.90	0.90	0.88
AB	0.88	0.87	0.89	0.88	0.88	0.86
GB	0.92	0.91	0.94	0.93	0.92	0.91
SVM	0.93	0.92	0.94	0.93	0.93	0.92
BGWO-SQSVM	0.96	0.95	0.97	0.96	0.96	0.95
Q-BGWO-SQSVM	0.98	0.97	0.99	0.98	0.98	0.97
15-CV						
RF	0.88	0.87	0.90	0.89	0.88	0.87
KNN	0.90	0.89	0.91	0.90	0.90	0.89
DT	0.86	0.85	0.88	0.87	0.86	0.84
NB	0.87	0.86	0.89	0.88	0.87	0.85
LR	0.91	0.90	0.93	0.92	0.91	0.90
AB	0.89	0.88	0.91	0.90	0.89	0.87
GB	0.93	0.92	0.95	0.94	0.93	0.92
SVM	0.94	0.93	0.96	0.95	0.94	0.93
BGWO-SQSVM	0.97	0.96	0.98	0.97	0.97	0.96
Q-BGWO-SQSVM	0.99	0.98	0.99	0.99	0.99	0.98

**Table 7.** DDSM Dataset Cross-Validation Performance Metrics.

effectively distinguish between positive and negative classes, with a higher proportion of True Positives and True Negatives than False Positives and False Negatives. When AUC is below 0.5, the classifier's ability to differentiate diminishes, and it may perform no better than random guessing or consistently assigning a fixed class to all data points.

The ROC curves and AUC values in the figures provide insights into the classifiers' ability to differentiate between disease and non-disease instances across different datasets and CV schemes. The use of cross-validation helps in selecting the best-performing model, contributing to robust and reliable classification outcomes.

### Comparative analysis with the state-of-the-art models

The efficacy of the proposed models was evaluated by comparing them to other approaches for breast cancer detection, as presented in Table 9. The comparative results are systematically detailed in the table presented earlier. The Q-BGWO-SQSVM algorithm demonstrated significant effectiveness across key metrics, including achieving 99% accuracy, 98% sensitivity, and 100% specificity in 15-fold cross-validation, which is the highest compared methods. These benchmarks underline its ability to provide near-perfect classification. For instance, compared to the SVM-AR method (98% accuracy, 98.03% sensitivity, and 97.93% specificity), the Q-BGWO-SQSVM achieves a notable improvement in specificity, reaching 100%, which is critical for minimizing false positives in clinical diagnostics. Similarly, the Q-BGWO-SQSVM surpasses GWO-SVM in accuracy (97.33%), sensitivity (95.96%), and specificity, reflecting the advantages of the proposed hybrid optimization framework. This model outperforms traditional methods and sets a new standard in breast cancer diagnostics, establishing its dominance over other cutting-edge methodologies.

Diagnosing cancerous breast tissue is critical for reducing the disease's complications and treating it in its early stages. Due to the dense breast tissue, the detection and localization of cancerous tissues have become more complex and error-prone. However, mammography misinterpretation may cause a needless biopsy of the FalseP outcomes, increasing the patient's anxiety. So, methods that learn to recognize breast lesions can mitigate

Model	ACC	SEN	SPC	PRE	F1-S	MCC
5-CV						
RF	0.85	0.83	0.86	0.85	0.84	0.82
KNN	0.86	0.85	0.87	0.86	0.86	0.84
DT	0.82	0.80	0.84	0.83	0.81	0.79
NB	0.84	0.82	0.86	0.85	0.83	0.81
LR	0.87	0.86	0.89	0.88	0.87	0.85
AB	0.84	0.83	0.85	0.84	0.84	0.82
GB	0.88	0.87	0.90	0.89	0.88	0.86
SVM	0.89	0.88	0.91	0.90	0.89	0.87
BGWO-SQSVM	0.93	0.92	0.95	0.94	0.93	0.91
Q-BGWO-SQSVM	0.96	0.95	0.98	0.97	0.96	0.94
10-CV						
RF	0.86	0.85	0.88	0.87	0.86	0.84
KNN	0.87	0.86	0.89	0.88	0.87	0.85
DT	0.83	0.81	0.85	0.84	0.82	0.80
NB	0.85	0.84	0.87	0.86	0.85	0.83
LR	0.88	0.87	0.90	0.89	0.88	0.86
AB	0.85	0.84	0.87	0.86	0.85	0.83
GB	0.90	0.89	0.92	0.91	0.90	0.88
SVM	0.92	0.91	0.94	0.93	0.92	0.90
BGWO-SQSVM	0.95	0.94	0.97	0.96	0.95	0.93
Q-BGWO-SQSVM	0.98	0.97	0.99	0.98	0.98	0.96
15-CV						
RF	0.87	0.86	0.89	0.88	0.87	0.85
KNN	0.89	0.88	0.91	0.90	0.89	0.87
DT	0.84	0.83	0.86	0.85	0.84	0.82
NB	0.86	0.85	0.88	0.87	0.86	0.84
LR	0.90	0.89	0.92	0.91	0.90	0.88
AB	0.87	0.86	0.89	0.88	0.87	0.85
GB	0.91	0.90	0.93	0.92	0.91	0.89
SVM	0.93	0.92	0.95	0.94	0.93	0.91
BGWO-SQSVM	0.97	0.96	0.98	0.97	0.96	0.94
Q-BGWO-SQSVM	0.99	0.98	1.00	0.99	0.99	0.98

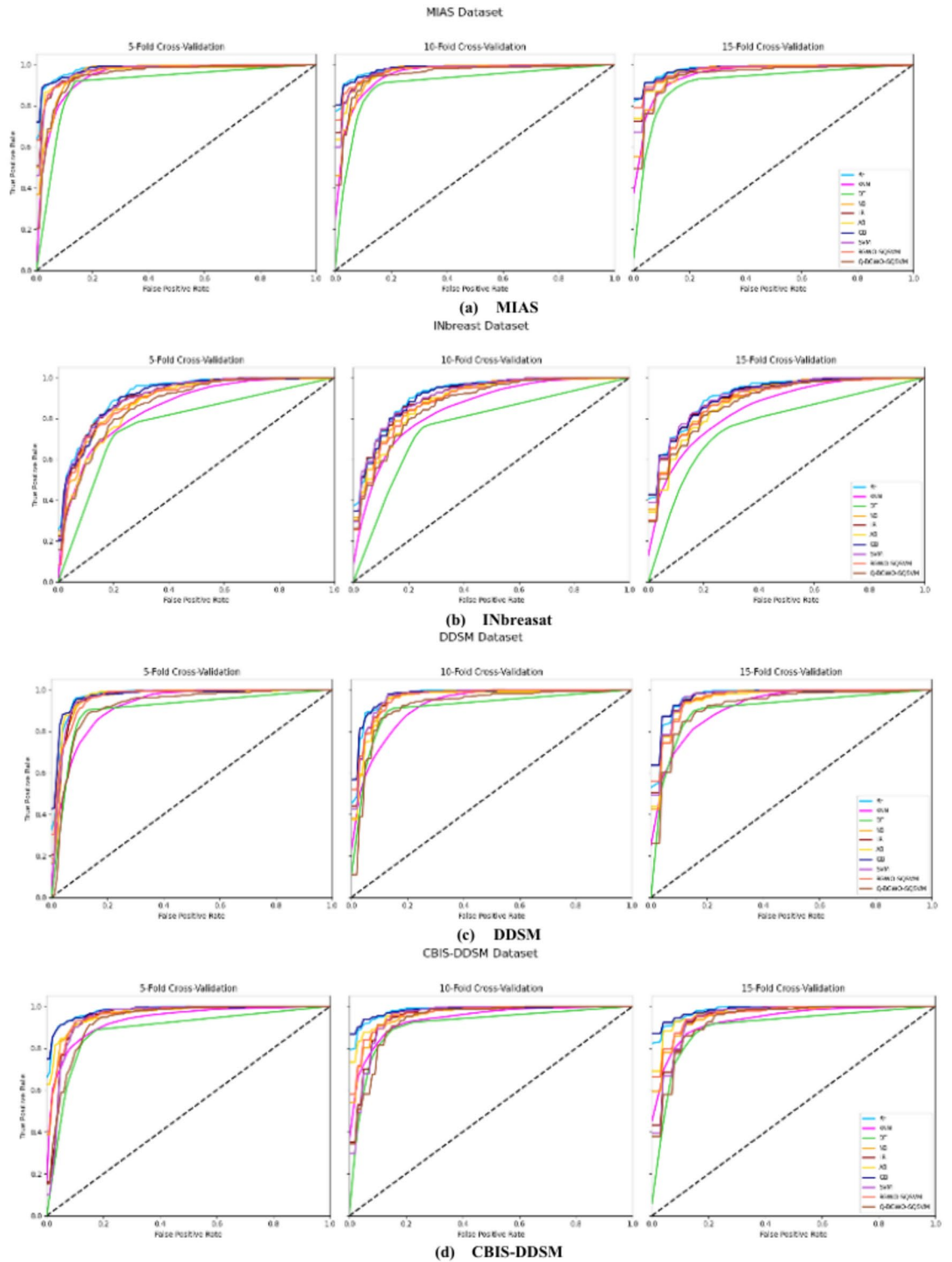
**Table 8.** CBIS-DDSM Dataset Cross-Validation Performance Metrics.

the number of ambiguities and inaccurate predictions. These errors may be reduced if computer-based detection technologies aid radiologists in recognizing subtle and probably curable breast cancer. Computer-based detection technologies' accuracy may enable more selective biopsies and decrease the number of biopsies. This study aims to improve current computerized schemes using innovative image recognition techniques to help radiologists diagnose microcalcifications and masses from the mammography image and cropped ROI. This research analyses the potential, constraints, and performance of modern deep learning and machine learning-based breast cancer detection systems by reducing computational and time complexity. Machine learning algorithms are insufficient for accurately detecting, localizing, and classifying breast densities. However, deep learning algorithms have reduced the false-positive rate (FPR) in mass detection and localization. Precise mammogram grading using deep learning-based models has a substantial advantage, including lower annotation, improved contextual information use, minor call-back rates, and needless examinations without losing the algorithm's sensitivity. The advent of deep learning techniques has assisted in detecting, segmenting, and classifying breast abnormalities, which benefits radiologists and researchers.

Researchers have proposed various methodologies over the years to enhance the accuracy and efficiency of breast cancer diagnosis. The table below presents a comparative analysis of different methods, highlighting their performance in terms of accuracy (ACC), sensitivity (Sen), specificity (SpC), and cross-validation (CV) techniques used.

In 2018, the L-FA method achieved an accuracy of 97.28% with fivefold cross-validation, while the ELM-SPI method reached 96.43%. The PML-Net method from the same year reported a 94.0% accuracy. Another approach mentioned, Kmean-SVM, demonstrated a similarly high accuracy of 97.28%.

The following year, 2019, saw the introduction of GA-OGB, which achieved 94.28% accuracy, 93.20% sensitivity, and 93.11% specificity. In 2020, LF-ANN achieved 97% accuracy with a sensitivity of 94%. The SVM-AR approach reported notable high performance with 98% accuracy, 98.03% sensitivity, and 97.93% specificity. Other 2020 methodologies like N-B and SMO achieved 98% and 98.73% accuracy, respectively.



**Fig.13.** AUC Comparison.

In subsequent years, various other methodologies were proposed. Disep-Net, introduced in 2021, achieved 95.60% accuracy. In 2023, a combination of LR + SVM reached 98.77% accuracy. By 2024, methods such as BC-NET and GWO-SVM demonstrated 98.01 and 97.33% accuracy, respectively.

Particularly notable are the methods that utilized more robust cross-validation techniques. The proposed Q-BGWO-SQSVN method, detailed in this study, demonstrated exceptional performance across various cross-validation techniques. Specifically, it showed remarkable outcomes with an accuracy of 99.00%, sensitivity of 98.00%, and specificity of 100% in 15-fold cross-validation. This high level of performance indicates near-perfect

Ref #	Methodology	ACC%	Sen%	Sp%
97	D-T	95.8	–	–
40	GSL	75.05	–	–
72	FCM-ELM	96.46	–	–
74	GL-SGL1/2	92.42	–	–
48	L-FA	97.28	–	–
47	ELM-SPI	96.43	–	–
98	PML-Net	94.0	–	–
34	Kmean-SVM	97.28	–	–
36	GA-OGB	94.28	93.20	93.11
99	LF-ANN	97	94.00	–
49	SVM-AR	98.00	98.03	97.93
100	N-B	98	–	–
	j-48	97.54	–	–
	SMO	98.73	–	–
101	CNN	92.54	–	–
39	FA-EMODE	96.86	–	–
102	DBN-SFO	91.5	94.1	72.4
103	Disep-Net	95.60	–	–
76	DT-AFIS	95.91	–	–
10	LR + SVM	98.77	–	–
65	BC-NET	98.01	97.75	98.16
104	GWO-SVM	97.33	95.96	–
Proposed	Q-BGWO-SQSVM	99.00	98.00	1.00

**Table 9.** State-of-the-art Comparison.

classification capability, significantly outperforming other methods and setting a new benchmark in breast cancer diagnostics.

The proposed hybrid models, particularly the Q-BGWO-SQSVM, demonstrate superior performance compared to existing methods, providing a robust and efficient approach to breast cancer diagnosis. The use of varying cross-validation techniques further validates the reliability and generalizability of these models.

## Conclusions

This paper presents a novel approach to classifying breast cancer by combining a quantum-inspired binary Grey Wolf Optimizer (Q-BGWO) with sophisticated computational models such as SqueezeNet and Support Vector Machines (SVM). The procedure consists of two separate stages. Firstly, it involves extracting important characteristics from mammographic pictures, identifying significant indications such as calcifications and masses. This extraction establishes the foundation for the subsequent phase, in which an advanced hybrid classification algorithm accurately differentiates between benign and malignant tumours.

Applying the Q-BGWO method to optimize SVM parameters significantly improves classification accuracy. Thorough testing on the CBIS-DDSM dataset has shown that the Q-BGWO-SQSVM model achieves remarkable accuracy rates: 98% in tenfold cross-validation, 96% in fivefold cross-validation, and an exceptional 99% accuracy with 100% specificity in 15-fold cross-validation situations. These results demonstrate the model's effectiveness and highlight this method's resilience across different cross-validation settings. The BGWO-SVM version shown excellent performance, highlighting the strength of this method.

This method's effectiveness stems from its accuracy and innovative combination of quantum physics and machine learning technologies, establishing a new benchmark in breast cancer diagnosis. Although the findings show promise, the research demonstrates the foundation for future improvements. Expanding the validation to include a wide range of datasets will verify the method's suitability worldwide. Future advancements may consist of using more sophisticated deep learning methods and investigating novel optimization strategies to enhance diagnosis accuracy while reducing processing requirements.

Improving the method for selecting areas of interest might enhance the precision of diagnostic accuracy. As we progress further, it is crucial to translate these improvements into clinical settings by extensively validating them with bigger real-world datasets and collaborating with medical specialists. This is necessary to transform these scientific discoveries into useful diagnostic tools. This research not only pushes the boundaries of technology in detecting early breast cancer but also makes a significant impact in reducing mortality rates and improving the effectiveness of computer-aided diagnosis systems in medical imaging. This paves the way for more dependable and efficient healthcare solutions in oncology.

## Data availability

The datasets used in this study are publicly available. And can be download using the links below. MIAS: <http://peipa.essex.ac.uk/info/mias.html> INbreast: <https://www.kaggle.com/datasets/martholi/inbreast> DDSM: <http://www.eng.usf.edu/cvprg/Mammography/Database.html> CBIS-DDSM: <https://www.cancerimagingarchive.net/collection/cbis-ddsm/>

Received: 3 November 2024; Accepted: 13 January 2025

Published online: 25 January 2025

## References

- Sung, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA. Cancer J. Clin.* <https://doi.org/10.3322/caac.21660> (2021).
- World Cancer Report. *World Heal. Organ.* (2023).
- Sharma, A. & Ives, C. Management of breast cancer: Basic principles. *Surgery (United Kingdom)* **40**, 113–120 (2022).
- Moo, T. A., Sanford, R., Dang, C. & Morrow, M. Overview of breast cancer therapy. *PET Clin.* **13**, 339–354 (2018).
- Organization, W.H. WHO Cancer factsheet. *Who* (2017).
- Sahu, Y. et al. A CNN-SVM based computer aided diagnosis of breast cancer using histogram K-means segmentation technique. *Multimed. Tools Appl.* <https://doi.org/10.1007/s11042-022-13807-x> (2023).
- Momenimovahed, Z. & Salehiniya, H. Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer Targets Ther.* <https://doi.org/10.2147/BCTT.S176070> (2019).
- Lopez, M.E.; Olutoye, O.O. Breast embryology, anatomy, and physiology. In *Endocrine Surgery in Children*. (2017).
- Kretz, T., Mueller, K. R., Schaeffter, T. & Elster, C. Mammography image quality assurance using deep learning. *IEEE Trans. Biomed. Eng.* <https://doi.org/10.1109/TBME.2020.2983539> (2020).
- Uddin, K. M. M., Biswas, N., Rikta, S. T. & Dey, S. K. Machine learning-based diagnosis of breast cancer utilizing feature optimization technique. *Comput. Methods Programs Biomed. Updat.* <https://doi.org/10.1016/j.cmpbup.2023.100098> (2023).
- Aswathy, M. A. & Jagannath, M. Detection of breast cancer on digital histopathology images: Present status and future possibilities. *Inform. Med. Unlocked* <https://doi.org/10.1016/j.imu.2016.11.001> (2017).
- Sain, S. R. & Vapnik, V. N. The nature of statistical learning theory. *Technometrics* <https://doi.org/10.2307/1271324> (1996).
- Cristianini, N.; Shawe-Taylor, J. *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. (2000).
- Boser, B.E.; Guyon, I.M.; Vapnik, V.N. Training algorithm for optimal margin classifiers. In *Proceedings of the Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*. (1992).
- McMahon, D. *Quantum Computing Explained*. (2007).
- Ferry, D. An introduction to quantum computing. In *Quantum Mechanics*. (2020).
- Mirjalili, S., Mirjalili, S. M. & Lewis, A. Grey wolf optimizer. *Adv. Eng. Softw.* **69**, 46–61 (2014).
- Uymaz, S. A., Tezel, G. & Yel, E. Artificial algae algorithm (AAA) for nonlinear global optimization. *Appl. Soft Comput. J.* <https://doi.org/10.1016/j.asoc.2015.03.003> (2015).
- Li, X. L., Shao, Z. J. & Qian, J. X. Optimizing method based on autonomous animals: Fish-swarm algorithm. *Xitong Gongcheng Lilun yu Shijian/System Eng. Theory Pract.* **22**, 32–38 (2002).
- Muro, C., Escobedo, R., Spector, L. & Coppinger, R. P. Wolf-Pack (Canis Lupus) hunting strategies emerge from simple rules in computational simulations. *Behav. Process.* <https://doi.org/10.1016/j.beproc.2011.09.006> (2011).
- Sahu, B.; Dash, S. BIBHU: Biomarker identification using bio-inspired evolutionary hybrid unique machine learning model. In *Proceedings of the 2023 World Conference on Communication and Computing*. (2023).
- Sahu, B.; Dash, S. Hybrid multifilter ensemble based feature selection model from microarray cancer datasets using GWO with deep learning. In *Proceedings of the 2023 3rd International Conference on Intelligent Technologies*. (2023).
- Sahoo, G. et al. Predicting breast cancer relapse from histopathological images with ensemble machine learning models. *Curr. Oncol.* **31**, 6577–6597 (2024).
- Sahu, B.; Dash, S. Hybrid binary grey wolf with jaya optimizer for biomarker selection from cancer datasets. In *Proceedings of the 2023 International Conference on Emerging Smart Computing and Informatics*. (2023).
- Pati, A., Parhi, M., Pattanayak, B. K., Sahu, B. & Khasim, S. CanDiag: Fog empowered transfer deep learning based approach for cancer diagnosis. *Designs* <https://doi.org/10.3390/designs7030057> (2023).
- Bacanin, N. et al. Performance of a novel chaotic firefly algorithm with enhanced exploration for tackling global optimization problems: Application for dropout regularization. *Mathematics* <https://doi.org/10.3390/math9212705> (2021).
- Malakar, S., Ghosh, M., Bhowmik, S., Sarkar, R. & Nasipuri, M. A GA based hierarchical feature selection approach for handwritten word recognition. *Neural Comput. Appl.* <https://doi.org/10.1007/s00521-018-3937-8> (2020).
- Zivkovic, M. et al. Novel hybrid firefly algorithm: An application to enhance XGBoost tuning for intrusion detection classification. *PeerJ Comput. Sci.* <https://doi.org/10.7717/peerj-cs.956> (2022).
- Dobrojevic, M. et al. Addressing internet of things security by enhanced sine cosine metaheuristics tuned hybrid machine learning model and results interpretation based on SHAP approach. *PeerJ Comput. Sci.* <https://doi.org/10.7717/peerj-cs.1405> (2023).
- Jovanovic, L.; Gajevic, M.; Dobrojevic, M.; Budimirovic, N.; Bacanin, N.; Zivkovic, M. Tackling IoT security challenge by metaheuristics tuned extreme learning machine. In *Proceedings of the Lecture Notes in Networks and Systems*. (2023).
- Marcano-Cedeño, A., Quintanilla-Domínguez, J. & Andina, D. WBCD breast cancer database classification applying artificial metaplasticity neural network. *Expert Syst. Appl.* <https://doi.org/10.1016/j.eswa.2011.01.167> (2011).
- Chen, H. L., Yang, B., Liu, J. & Liu, D. Y. A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. *Expert Syst. Appl.* <https://doi.org/10.1016/j.eswa.2011.01.120> (2011).
- Chen, H. L. et al. Support vector machine based diagnostic system for breast cancer using swarm intelligence. *J. Med. Syst.* <https://doi.org/10.1007/s10916-011-9723-0> (2012).
- Kumari, M.; Singh, V. Breast cancer prediction system. In *Proceedings of the Procedia Computer Science*. (2018).
- Kompalli, V.S.; Kuruba, U.R. Combined effect of soft computing methods in classification. In *Proceedings of the Advances in Intelligent Systems and Computing*. (2017).
- Lu, H., Wang, H. & Yoon, S. W. A dynamic gradient boosting machine using genetic optimizer for practical breast cancer prognosis. *Expert Syst. Appl.* <https://doi.org/10.1016/j.eswa.2018.08.040> (2019).
- Ma, B. & Xia, Y. A tribe competition-based genetic algorithm for feature selection in pattern classification. *Appl. Soft Comput. J.* <https://doi.org/10.1016/j.asoc.2017.04.042> (2017).
- Pota, M., Esposito, M. & De Pietro, G. Designing rule-based fuzzy systems for classification in medicine. *Knowl. Based Syst.* <https://doi.org/10.1016/j.knosys.2017.03.006> (2017).
- Nayak, S. K., Rout, P. K., Jagadev, A. K. & Swarnkar, T. Elitism based multi-objective differential evolution for feature selection: A filter approach with an efficient redundancy measure. *J. King Saud Univ. Comput. Inf. Sci.* <https://doi.org/10.1016/j.jksuci.2017.08.001> (2020).

40. Shoeleh, F. & Asadpour, M. Graph based skill acquisition and transfer learning for continuous reinforcement learning domains. *Pattern Recognit. Lett.* <https://doi.org/10.1016/j.patrec.2016.08.009> (2017).
41. Bashir, S., Qamar, U. & Khan, F. H. Heterogeneous classifiers fusion for dynamic breast cancer diagnosis using weighted vote based ensemble. *Qual. Quant.* <https://doi.org/10.1007/s11135-014-0090-z> (2015).
42. Tuba, E., Tuba, M.; Simian, D. Adjusted bat algorithm for tuning of support vector machine parameters. In *Proceedings of the 2016 IEEE Congress on Evolutionary Computation.* (2016).
43. Aalaei, S., Shahraki, H., Rowhanimanesh, A. & Eslami, S. Feature selection using genetic algorithm for breast cancer diagnosis: Experiment on three different datasets. *Iran. J. Basic Med. Sci.* **19**, 476 (2016).
44. Kumar Mandal, S. Performance analysis of data mining algorithms for breast cancer cell detection using naïve Bayes, logistic regression and decision tree. *Int. J. Eng. Comput. Sci.* **6**, 20388–20391 (2017).
45. Muslim, M.A.; Rukmana, S.H.; Sugiharti, E.; Prasetyo, B.; Alimah, S. Optimization of C4.5 algorithm-based particle swarm optimization for breast cancer diagnosis. In *Proceedings of the Journal of Physics: Conference Series.* (2018).
46. Liu, N. et al. Improved cost-sensitive support vector machine classifier for breast cancer diagnosis. *Math. Probl. Eng.* <https://doi.org/10.1155/2018/3875082> (2018).
47. Kassani, P. H., Teoh, A. B. J. & Kim, E. Sparse pseunverse incremental extreme learning machine. *Neurocomputing* <https://doi.org/10.1016/j.neucom.2018.01.087> (2018).
48. Pota, M., Esposito, M. & De Pietro, G. Likelihood-fuzzy analysis: from data, through statistics, to interpretable fuzzy classifiers. *Int. J. Approx. Reason.* <https://doi.org/10.1016/j.ijar.2017.10.022> (2018).
49. Maalmi, K. Breast cancer classification with reduced feature set using association rules and support vector machine. *Netw. Model. Anal. Heal. Inform. Bioinform.* <https://doi.org/10.1007/s13721-020-00237-8> (2020).
50. Fu, Z.; Zhang, D.; Zhao, X.; Li, X. Adaboost algorithm with floating threshold. In *Proceedings of the IET Conference Publications.* (2012).
51. Prasad, Y.; Biswas, K.K.; Jain, C.K. SVM Classifier based feature selection using GA, ACO and PSO for SiRNA design. In *Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics).* (2010).
52. Zheng, B., Yoon, S. W. & Lam, S. S. Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. *Expert Syst. Appl.* <https://doi.org/10.1016/j.eswa.2013.08.044> (2014).
53. De Falco, L., Della Cioppa, A. & Tarantino, E. Facing classification problems with particle swarm optimization. *Appl. Soft Comput. J.* <https://doi.org/10.1016/j.asoc.2005.09.004> (2007).
54. Agarap, A.F.M. On breast cancer detection: An application of machine learning algorithms on the wisconsin diagnostic dataset. In: *Proceedings of the ACM International Conference Proceeding Series.* (2018).
55. Huang, H. et al. A new fruit fly optimization algorithm enhanced support vector machine for diagnosis of breast cancer based on high-level features. *BMC Bioinform.* <https://doi.org/10.1186/s12859-019-2771-z> (2019).
56. Xie, T., Yao, J. & Zhou, Z. DA-based parameter optimization of combined kernel support vector machine for cancer diagnosis. *Processes* <https://doi.org/10.3390/pr7050263> (2019).
57. Rajaguru, H. & Sannasi Chakravarthy, S. R. Analysis of decision tree and K-nearest neighbor algorithm in the classification of breast cancer. *Asian Pacific J. Cancer Prev.* <https://doi.org/10.31557/APJCP.2019.20.12.3777> (2019).
58. Dhahri, H., Al Maghayreh, E., Mahmood, A., Elkilani, W. & Faisal Nagi, M. Automated breast cancer diagnosis based on machine learning algorithms. *J. Healthc. Eng.* <https://doi.org/10.1155/2019/4253641> (2019).
59. Hemeida, A. M. et al. Implementation of nature-inspired optimization algorithms in some data mining tasks. *Ain Shams Eng. J.* <https://doi.org/10.1016/j.asej.2019.10.003> (2020).
60. Telsang, V.A.; Hegde, K. Breast cancer prediction analysis using machine learning algorithms. In *Proceedings of the Proceedings of the 2020 IEEE International Conference on Communication, Computing and Industry 4.0.* (2020).
61. Umme Salma, M. & Doreswamy, D. Hybrid BATGSA: A metaheuristic model for classification of breast cancer data. *Int. J. Adv. Intell. Paradig.* <https://doi.org/10.1504/IJAIP.2020.105144> (2020).
62. Singh, I.; Bansal, R.; Gupta, A.; Singh, A. A hybrid grey wolf-whale optimization algorithm for optimizing SVM in breast cancer diagnosis. In *Proceedings of the PDGC 2020 - 2020 6th International Conference on Parallel, Distributed and Grid Computing.* (2020).
63. Badr, E., Almotairi, S., Salam, M. A. & Ahmed, H. New sequential and parallel support vector machine with grey wolf optimizer for breast cancer diagnosis. *Alexandria Eng. J.* <https://doi.org/10.1016/j.aej.2021.07.024> (2022).
64. Badr, E., Salam, M. A., Almotairi, S. & Ahmed, H. From linear programming approach to metaheuristic approach: Scaling techniques. *Complexity* <https://doi.org/10.1155/2021/9384318> (2021).
65. Bilal, A. et al. BC-QNet: A quantum-infused ELM model for breast cancer diagnosis. *Comput. Biol. Med.* <https://doi.org/10.1016/J.COMPBIOMED.2024.108483> (2024).
66. Bilal, A., Liu, X., Shafiq, M., Ahmed, Z. & Long, H. NIMEQ-SACNet: A novel self-attention precision medicine model for vision-threatening diabetic retinopathy using image data. *Comput. Biol. Med.* **171**, 108099. <https://doi.org/10.1016/J.COMPBIOMED.2024.108099> (2024).
67. Bilal, Anas et al. Improved support vector machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification. *PLoS ONE* **19**, e0295951 (2024).
68. Tapkan, Pi., Özbakir, L. & Baykasoğlu, A. Bee algorithms for parallel two-sided assembly line balancing problem with walking times. *Appl. Soft Comput. J.* <https://doi.org/10.1016/j.asoc.2015.11.017> (2016).
69. Polap, D., Kesik, K., Woźniak, M. & Damaševičius, R. Parallel technique for the metaheuristic algorithms using devoted local search and manipulating the solutions space. *Appl. Sci.* <https://doi.org/10.3390/app8020293> (2018).
70. Jiao, S., Gao, Y., Feng, J., Lei, T. & Yuan, X. Does deep learning always outperform simple linear regression in optical imaging?. *Opt. Express* <https://doi.org/10.1364/oe.382319> (2020).
71. Chauhan, D. et al. Comparison of machine learning and deep learning for view identification from cardiac magnetic resonance images. *Clin. Imaging* <https://doi.org/10.1016/j.clinimag.2021.11.013> (2022).
72. Liangjun, C., Honeine, P., Hua, Q., Jihong, Z. & Xia, S. Correntropy-based robust multilayer extreme learning machines. *Pattern Recognit.* <https://doi.org/10.1016/j.patcog.2018.07.011> (2018).
73. Jafari-Marandi, R., Davarzani, S., Soltanpour Gharibdousti, M. & Smith, B. K. An optimum ANN-based breast cancer diagnosis: Bridging gaps between ANN learning and decision-making goals. *Appl. Soft Comput. J.* <https://doi.org/10.1016/j.asoc.2018.07.060> (2018).
74. Li, F., Zurada, J. M. & Wu, W. Smooth group L1/2 regularization for input layer of feedforward neural networks. *Neurocomputing* <https://doi.org/10.1016/j.neucom.2018.06.046> (2018).
75. Taghizadeh, E., Heydarheydari, S., Saberi, A., JafarpourNesheli, S. & Rezaei, S. M. Breast cancer prediction with transcriptome profiling using feature selection and machine learning methods. *BMC Bioinform.* <https://doi.org/10.1186/s12859-022-04965-8> (2022).
76. Chen, T. et al. A decision tree-initialised neuro-fuzzy approach for clinical decision support. *Artif. Intell. Med.* <https://doi.org/10.1016/j.artmed.2020.101986> (2021).
77. Van Der Maaten, L. & Hinton, G. Visualizing data using T-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
78. Suckling, J. et al. The mammographic image analysis society digital mammogram database. *Expert. Medica Int. Congr. Ser.* **1069**, 375–378 (1994).

79. Moreira, I. C. et al. INbreast: Toward a full-field digital mammographic database. *Acad. Radiol.* <https://doi.org/10.1016/j.acra.2011.09.014> (2012).
80. Florida, U. of S. Digital database for screening mammography (DDSM). (2021).
81. Archive, C.I. Curated breast imaging subset of digital database for screening mammography (CBIS-DDSM). (2023).
82. Kerlikowske, K. et al. Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Ann. Intern. Med.* <https://doi.org/10.7326/0003-4819-133-11-200012050-00009> (2000).
83. Suri, J.S.; Sun, Y.; Janer, R. Method for Breast Screening in Fused Mammography. *US Pat.* 10,363,010. (2013).
84. Khamparia, A. et al. Diagnosis of breast cancer based on modern mammography using hybrid transfer learning. *Multidimens. Syst. Signal Process.* <https://doi.org/10.1007/s11045-020-00756-7> (2021).
85. Dhungel, N.; Carneiro, G.; Bradley, A.P. Automated mass detection in mammograms using cascaded deep learning and random forests. In *Proceedings of the 2015 International Conference on Digital Image Computing: Techniques and Applications, DICTA*. 2015; (2015).
86. Ting, F. F., Tan, Y. J. & Sim, K. S. Convolutional neural network improvement for breast cancer classification. *Expert Syst. Appl.* <https://doi.org/10.1016/j.eswa.2018.11.008> (2019).
87. Mahmood, T., Li, J., Pei, Y. & Akhtar, F. An automated in-depth feature learning algorithm for breast abnormality prognosis and robust characterization from mammography images using deep transfer learning. *Biology (Basel)* <https://doi.org/10.3390/biology10090859> (2021).
88. Wu, E.; Wu, K.; Cox, D.; Lotter, W. Conditional Infilling GANs for Data Augmentation in Mammogram Classification. In *Proceedings of the Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11040 (2018).
89. Sérvulo, Soares et al. Classification of breast regions as mass and non-mass based on digital mammograms using taxonomic indexes and SVM. *Comput. Biol. Med.* <https://doi.org/10.1016/j.combiomed.2014.11.016> (2015).
90. Medina-Carnicer, R., Madrid-Cuevas, F. J., Fernández-García, N. L. & Carmona-Poyato, A. Evaluation of global thresholding techniques in non-contextual edge detection. *Pattern Recognit. Lett.* <https://doi.org/10.1016/j.patrec.2004.11.024> (2005).
91. Otsu, N. Threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern* <https://doi.org/10.1109/tsmc.1979.4310076> (1979).
92. Srikanth, K. et al. Meta-heuristic framework: Quantum inspired binary grey wolf optimizer for unit commitment problem. *Comput. Electr. Eng.* <https://doi.org/10.1016/j.compeleceng.2017.07.023> (2018).
93. Başaran, E. Classification of white blood cells with SVM by selecting SqueezeNet and LIME properties by MRMR method. *Signal Image Video Process.* <https://doi.org/10.1007/s11760-022-02141-2> (2022).
94. Bilal, A., Sun, G., Mazhar, S., Imran, A. & Latif, J. A transfer learning and U-Net-based automatic detection of diabetic retinopathy from fundus images. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **10**, 663–674 (2022).
95. Kandil, H., Soliman, A., Alghamdi, N. S., Jennings, J. R. & El-Baz, A. Using mean arterial pressure in hypertension diagnosis versus using either systolic or diastolic blood pressure measurements. *Biomedicines* <https://doi.org/10.3390/biomedicines11030849> (2023).
96. Derrac, J., García, S., Molina, D. & Herrera, F. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm Evol. Comput.* <https://doi.org/10.1016/j.swevo.2011.02.002> (2011).
97. Hayashi, Y. & Nakano, S. Use of a recursive-rule extraction algorithm with J48graft to achieve highly accurate and concise rule extraction from a large breast cancer dataset. *Inform. Med. Unlocked* <https://doi.org/10.1016/j.imu.2015.12.002> (2015).
98. Jiao, Z., Gao, X., Wang, Y. & Li, J. A parasitic metric learning net for breast mass classification based on mammography. *Pattern Recognit.* <https://doi.org/10.1016/j.patcog.2017.07.008> (2018).
99. Hajiabadi, H., Babaiyan, V., Zabihzadeh, D. & Hajiabadi, M. Combination of loss functions for robust breast cancer prediction. *Comput. Electr. Eng.* <https://doi.org/10.1016/j.compeleceng.2020.106624> (2020).
100. Mohammed, S.A.; Darrab, S.; Noaman, S.A.; Saake, G. Analysis of Breast Cancer Detection Using Different Machine Learning Techniques. In *Proceedings of the Communications in Computer and Information Science*. 1234, (2020).
101. Gnanasekaran, V. S., Joypaul, S., Sundaram, P. M. & Chairman, D. D. Deep learning algorithm for breast masses classification in mammograms. *IET Image Process* <https://doi.org/10.1049/iet-ipr.2020.0070> (2020).
102. Shen, L. et al. Optimal breast tumor diagnosis using discrete wavelet transform and deep belief network based on improved sunflower optimization method. *Biomed. Signal Process. Control* <https://doi.org/10.1016/j.bspc.2020.101953> (2020).
103. Yu, X., Xia, K. & Zhang, Y. D. DisepNet for breast abnormality recognition. *Comput. Electr. Eng.* <https://doi.org/10.1016/j.compeleceng.2020.106961> (2021).
104. Bilal, Anas, Imran, Azhar, Baig, Talha Imtiaz & Xiaowen Liu, E. A. N. Breast cancer diagnosis using support vector machine optimized by improved quantum inspired grey wolf optimization. *Sci. Rep.* **14**, 10714. <https://doi.org/10.1038/s41598-024-61322-w> (2024).

## Author contributions

Conceptualization, A.B. and A.A.; methodology, A.B. F.K. and A.T.; software, A.B. and F.K.; validation, A.A., A.T. and M.S.; formal analysis, A.A., M.S. and H.L.; investigation, A.A., M.S.; resources, H.L.; data curation, A.B., A.A., F.K.; writing—original draft preparation, A.B.; writing—review and editing, A.T. and M.S.; visualization, A.B., A.A., F.K. and H.L.; All authors have read and agreed to the published version of the manuscript.

## Funding

This research was funded by the National Natural Science Foundation of China (No.62262019) and the Hainan Provincial Natural Science Foundation of China (823RC488).

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-86671-y>.

**Correspondence** and requests for materials should be addressed to H.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025