Enhancing Breast Cancer Classification based on BPSO Feature Selection and Machine Learning Techniques

Osama I. Ramadan
Oral Surgery and Diagnostic Sciences Department, Faculty of Dentistry, Applied Science Private University, Amman, Jordan
o_bayyoumi@asu.edu.jo

Lashin S. Ali
Department of Basic Medical Science, Faculty of Dentistry, Al-Ahliyya Amman University, Amman, Jordan | Department of Medical Physiology, Faculty of Medicine, Mansoura University, Egypt
lashin.saad@yahoo.com

Yasser Ramadan
Department of Computer Science, Faculty of Computers and Information, Suez University, Suez, Egypt
yasserfrb@gmail.com (corresponding author)

Randa M. Abobaker
Maternal and Child Health Nursing, North Private College of Nursing, Arar, Saudi Arabia
randaabobakre@gmail.com

Hoda M. Flifel
Nursing Education, North Private College of Nursing, Arar, Saudi Arabia
hoda85378@gmail.com

Mohamed A. Elkholy
Department of Basic Medical Science, Faculty of Dentistry, Al-Ahliyya Amman University, Amman, Jordan
m.alkholi@ammanu.edu.jo

Hadaiea I. Abobaker
Medical Surgical Nursing, North Private College of Nursing, Arar, Saudi Arabia
hadaiea.abobaker@alexu.edu.eg

Eman M. M. Gabr
Medical Surgical Nursing, North Private College of Nursing, Arar, Saudi Arabia
dr_emangabr@yahoo.com

Ibrahim I. Hemdan
Computer Sciences and Intelligent Systems, Basic Sciences Department, Faculty of Physical Therapy, Horus University, Egypt
ihemdan@horus.edu.eg

Samah A. Z. Hassan
Department of Information Systems, Faculty of Computers and Information, Suez University, Suez, Egypt
s.hassan@suezuni.edu.eg

Received: 10 March 2025 | Revised: 6 April 2025 and 9 April 2025 | Accepted: 19 April 2025

 DOI: https://doi.org/10.48084/etasr.10900

ABSTRACT

Breast cancer remains one of the most prevalent and life-threatening diseases among women worldwide. Early and accurate diagnosis have been shown to enhance treatment effectiveness and patient survival rates. This study presents an enhanced breast cancer classification framework by leveraging Machine Learning (ML) techniques and feature selection methods. The methodology involves data preprocessing, feature selection using the Binary Particle Swarm Optimization (BPSO), and classification through advanced ML models, including Random Forest (RF), Logistic Regression (LR), Gradient Boosting (GB), Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Naive Bayes (NB). The proposed approach is rigorously evaluated using key performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. By reducing the feature set from 30 to 13, BPSO enhances both model efficiency and predictive performance. Among the classifiers evaluated, RF achieved the highest accuracy of 99.2%, accompanied by a perfect ROC-AUC score of 1.0. The results demonstrate the potential of ML-driven breast cancer classification in revolutionizing healthcare by enabling more accurate, efficient, and personalized treatment strategies.

Keywords-breast cancer; breast cancer classification; breast cancer diagnosing; BPSO

I. INTRODUCTION

Breast cancer is the most prevalent and most often diagnosed malignancy affecting females [1, 2]. Early identification of breast cancer has been demonstrated to improve survival and treatment outcomes. A variety of techniques have been employed to detect breast cancer, including self-examination, clinical assessment, and screening tools such as mammography [3].

Machine Learning (ML) algorithms and optimization techniques have been employed to identify patterns and correlations in data associated with the biological processes underlying cancer development [4]. These methodologies establish a robust foundation for the development of prediction models that accurately classify breast cancer patients with high efficiency and reliability [5, 6]. ML algorithms leverage historical medical data, including clinical, histopathological, and imaging data, to identify complex patterns and relationships that are often indiscernible to traditional methods [7, 8]. When paired with optimization techniques, these models can be further refined to enhance performance by selecting the most relevant features, optimizing hyperparameters, and minimizing classification errors. The integration of ML with optimization techniques has emerged as a synergistic approach that addresses the challenges of breast cancer classification, including data imbalance, feature redundancy, and interpretability [9]. Due to tumor complexity, medical dataset dimensionality, and diagnostic data imbalances, breast cancer categorization remains a challenging task. Traditional diagnostic methods often generate an excessive number of false positives and negatives, thereby delaying therapy and compromising patient care. Innovative approaches that utilize sophisticated ML and optimization methods are necessary to construct robust and efficient diagnostic models [4].

The objective of the present study is to develop a comprehensive breast cancer classification framework that employs sophisticated ML models and optimization techniques. The proposed framework utilizes robust feature selection approaches, such as the Binary Particle Swarm Optimization (BPSO) method, to identify significant features and assess model performance. Despite the presence of data asymmetry and feature redundancy, the proposed approach can reduce false positives and negatives in breast cancer diagnosis [10].

Extensive research has been conducted on the classification of breast cancer based on optimization algorithms. Authors in [11] addressed the challenge of accurately classifying breast cancer by developing a novel ML model that integrates optimization techniques to enhance diagnostic precision. The authors employ a hybrid methodology, combining feature selection algorithms with advanced classifiers to optimize the model's performance. This optimization is achieved through the implementation of techniques such as Genetic Algorithms (GAs) and cross-validation. Their research plan involves the training and testing of the model on publicly available breast cancer datasets, followed by a comparative analysis against existing classification methods. The findings demonstrate that the proposed approach outperforms traditional models, achieving higher accuracy and reduced false positive and negative rates. This contributes to more reliable breast cancer diagnosis and potentially minimizes unnecessary biopsies.

Authors in [12] propose a technique that utilizes ML to classify breast cancers in mammograms for the purpose of early diagnosis and therapy. The researchers extracted 1,792 feature vectors from original and upgraded mammograms using haze-reduced adaptive methods, data augmentation, and the EfficientNet-B4 pre-trained architecture. The vectors are then categorized by ML methods. The framework achieved classifications with 98.459% and 96.175% accuracy.

Authors in [13] constructed a breast cancer classification model employing a Deep Neural Network (DNN), a GA, and an Egret Swarm Optimization (ESO). The model implements ILDA for data preprocessing, DNN for outlier identification, and GA for feature selection. The ESO algorithm identifies data as benign or malignant. The model demonstrated an accuracy of 99.30% in the classification of WBC data and 99.45% in the classification of WDBC data.

Authors in [14] focused on the challenge of accurately predicting breast cancer by comparing various feature selection methods integrated with ML algorithms. The authors employed techniques such as GAs, ant colony optimization, and the Hybrid Hopfield Neural Network-E2SAT (HHNN-E2SAT) model to enhance the predictive performance of classifiers. Their methodology involved applying these feature selection methods to identify the most relevant attributes from breast cancer datasets, followed by training ML models to assess improvements in prediction accuracy. The study revealed that the incorporation of these optimization-based feature selection techniques significantly improved the classifiers' ability to predict breast cancer, thereby contributing to the development of more reliable diagnostic tools.

Authors in [15] introduced a breast cancer classification system that utilizes histopathological images and integrates deep learning with optimization techniques. The methodology incorporates Wiener filtering for image preprocessing, ResNeXt for feature extraction, and a hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) model for classification. Hyperparameter tuning is achieved through the implementation of Sunflower Optimization (SFO). The results demonstrate the system's effectiveness, achieving high accuracy rates of 96.94% and 98.69% on diverse datasets, outperforming existing methods. The findings highlight the potential of this integrated approach for enhancing breast cancer detection and classification.

Authors in [16] have developed a methodology for enhancing breast cancer prediction that utilizes GAs, Chemical Reaction Optimization (CRO), and ML. The incorporation of GA and CRO has been demonstrated to enhance the process of feature selection and the optimization of hyperparameters, resulting in a significant enhancement of classifier performance. On three datasets, the suggested fusion strategy exhibited superior performance in terms of accuracy, precision, recall, and F1-score when compared to conventional models. This implies a solid and scalable clinical decision-making solution.

Authors in [17] proposed a hybrid model for early breast cancer diagnosis using a quantum-inspired binary grey wolf optimizer with radial basis function kernel Support Vector Machines (SVMs). The model optimizes SVM parameters and selects important features to improve diagnosis accuracy. The model exceeds SVMs and other optimization methods in terms of classification accuracy, making it a viable early breast cancer screening tool.

Several recent studies have explored the integration of ML techniques and advanced feature selection methods for breast cancer prediction. In [18], a hybrid Whale Optimization Algorithm and Dragonfly Algorithm (WOADA) was employed for selecting optimal features from mammographic data, demonstrating superior performance when combined with Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS) classifiers, achieving up to 98.00% accuracy. Similarly, authors in [19] proposed a feature selection method based on Water Wave Optimization (WWO) applied to the WDBC dataset, achieving an accuracy of 97.96% and proving effective as a clinical decision support tool by reducing redundant information and boosting classifier performance. Further extending the scope, authors in [20] introduced three metaheuristic feature selection strategies: Gravitational Search Algorithm (GSA), Emperor Penguin Optimization (EPO), and a hybrid of both (hGSAEPO). Their hybrid method achieved 98.31% accuracy and exceptionally high AUC, precision, and specificity scores, thereby confirming the value of combining feature selection and ML in binary classification tasks.

This comprehensive review of the literature reveals several key gaps in existing research on breast cancer classification. While numerous studies employ deep learning or hybrid optimization models, these approaches often require high computational resources and are less interpretable, making them challenging to implement in real-world clinical environments. Moreover, limited research has explored the integration of traditional ML models with efficient feature selection techniques, such as BPSO, on tabular diagnostic datasets like Fine Needle Aspirate (FNA) data. Many prior works focus on single-model evaluation and overlook comparative performance analysis across multiple classifiers. To address these gaps, this study presents a robust, interpretable, and computationally efficient classification framework. This framework applies BPSO for feature selection and evaluates six ML models. By reducing feature dimensionality from 30 to 13 while achieving a high classification accuracy of 99.2% using Random Forest (RF), the proposed approach demonstrates strong potential for scalable, accurate, and real-time clinical breast cancer diagnosis.

II. MATERIALS AND METHODS

This section describes data preprocessing, feature extraction, and the breast cancer detection ML model. It describes ML classifier training methodologies and performance measures including accuracy, precision, recall, and ROC-AUC. These methods make the chosen model efficient and successful in distinguishing malignant from benign situations.

A. Dataset Depiction

The dataset employed in this study was obtained from [21]. It comprises 569 instances, each corresponding to a separate FNA sample of a breast mass. From the digitized images of these samples, 30 numerical features describing the characteristics of cell nuclei were extracted. These features include measurements such as radius, texture, perimeter, area, and smoothness, which help classify the tumor as malignant or benign. The dataset is widely used for training ML models

aimed at breast cancer detection and prediction. A statistical evaluation of selected dataset attributes is presented in Table I.

Figure 1 displays the correlation matrix, which presents the correlation coefficients between a set of variables in a dataset. It provides insights into how pairs of variables are related, whether positively or negatively, and to what degree. A thorough examination of this matrix can facilitate the identification of redundant features, dependencies, or potential multicollinearity issues. These findings can serve as valuable insights, informing the process of feature selection and model development in ML.

TABLE I. STATISTICAL ANALYSIS OF SELECTED UTILIZED DATASET ATTRIBUTES

	Diagnosis	**Radius_mean**	**Perimeter_mean**	**Area_mean**	**Symmetry_mean**	**Texture_se**
Count	569	569	569	569	569	569
Mean	0.372583	14.12729	91.96903	654.8891	0.181162	1.216853
Std	0.483918	3.524049	24.29898	351.9141	0.027414	0.551648
Min	0	6.981	43.79	143.5	0.106	0.3602
25%	0	11.7	75.17	420.3	0.1619	0.8339
50%	0	13.37	86.24	551.1	0.1792	1.108
75%	1	15.78	104.1	782.7	0.1957	1.474
Max	1	28.11	188.5	2501	0.304	4.885

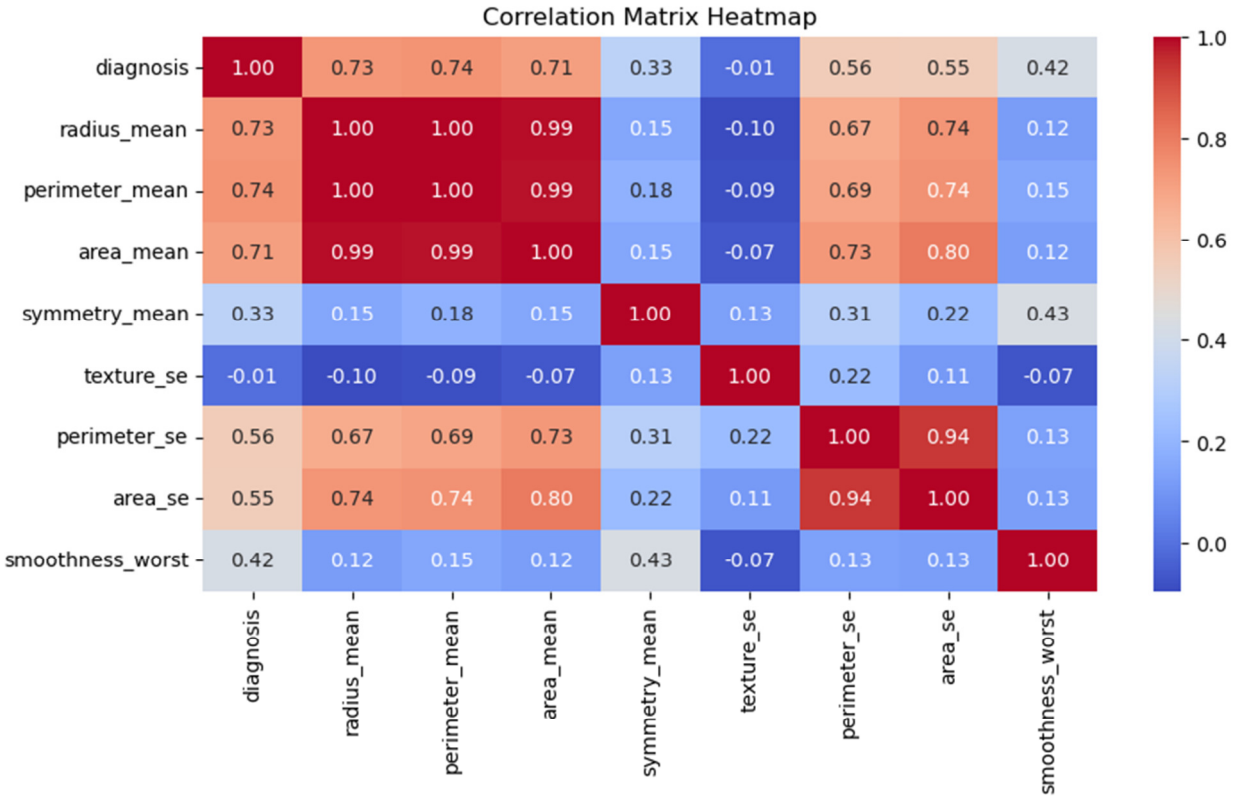


Fig. 1. Correlation matrix of the utilized dataset attributes.

B. Feature Selection

To optimize feature selection, the BPSO method [22-25] is applied as a specialized technique for reducing dimensionality. In this approach, binary particles represent whether specific features are included (1) or excluded (0). Although BPSO is a classical soft-computing algorithm, it remains widely used in recent biomedical applications due to its simplicity, interpretability, and competitive performance in feature selection. Its low computational cost makes it particularly attractive for resource-constrained environments, such as clinical diagnostic systems. Furthermore, recent studies have shown that BPSO can still outperform more complex and newer optimization algorithms in specific domains, particularly when paired with robust classifiers. By establishing a strong and interpretable baseline, this work enables future comparisons and provides a practical solution for real-world deployment.

By iteratively updating the particles' positions and velocities based on personal and global best solutions, BPSO effectively identifies a minimal subset of features [26] that enhances classification accuracy and computational efficiency. This approach achieves a balance between exploration and

exploitation, making it ideal for optimizing datasets like breast cancer diagnosis, where precise and efficient feature selection is critical.

For breast cancer data [27-28], BPSO selected 14 features, including radius mean, perimeter mean, and area worst, achieving high accuracy (0.98), with a mean fitness of 0.9886, and a low error rate (0.0412). This demonstrates its efficacy in optimizing classification performance while reducing dimensionality. Figure 2 illustrates the convergence curve of the BPSO algorithm.

C. Hyperparameter Tuning and Final Parameter Settings

To optimize the performance of the applied ML models, a grid search strategy was used to fine-tune the hyperparameters for each classifier. The tuning process was carried out using 5-fold cross-validation on the training set, ensuring that the models generalize well without overfitting. For BPSO, key parameters such as population size, inertia weight, and acceleration coefficients were finalized based on literature benchmarks and experimental evaluation for optimal feature subset selection. The final values utilized for each algorithm are presented in Table II.

D. Methodology

This study proposes a robust methodology for breast cancer classification by integrating advanced ML models with feature selection and optimization techniques. The process begins with data preprocessing, wherein missing values, outliers, and feature scaling are addressed to prepare the breast cancer dataset for analysis. The BPSO is applied to select the most relevant features, reducing the dataset from 30 attributes to 13, which enhances classification accuracy and computational efficiency. The preprocessed dataset is then split into training (80%) and testing (20%) subsets to facilitate model training and evaluation. Figure 3 presents the proposed methodology.

TABLE II. FINALIZED HYPERPARAMETERS FOR BPSO AND ML MODELS

Model	Hyperparameters	Value/Setting
Binary Particle Swarm Optimization (BPSO)	Population size	30
	Max iterations	100
	Inertia weight (w)	0.7
	Cognitive coefficient (c1)	1.5
	Social coefficient (c2)	1.5
Random Forest (RF)	Number of estimators	100
	Max depth	-
	Criterion	Gini
Logistic Regression (LR)	Regularization (C)	1.0
	Solver	Liblinear
Gradient Boosting (GB)	Number of estimators	100
	Learning rate	0.1
	Max depth	3
Support Vector Machine (SVM)	Kernel	RBF
	Regularization parameter (C)	1
	Gamma	Scale
K-Nearest Neighbors (KNN)	Number of Neighbors (k)	5
	Distance metric	Euclidean
Naive Bayes (NB)	Assumed distribution	Gaussian

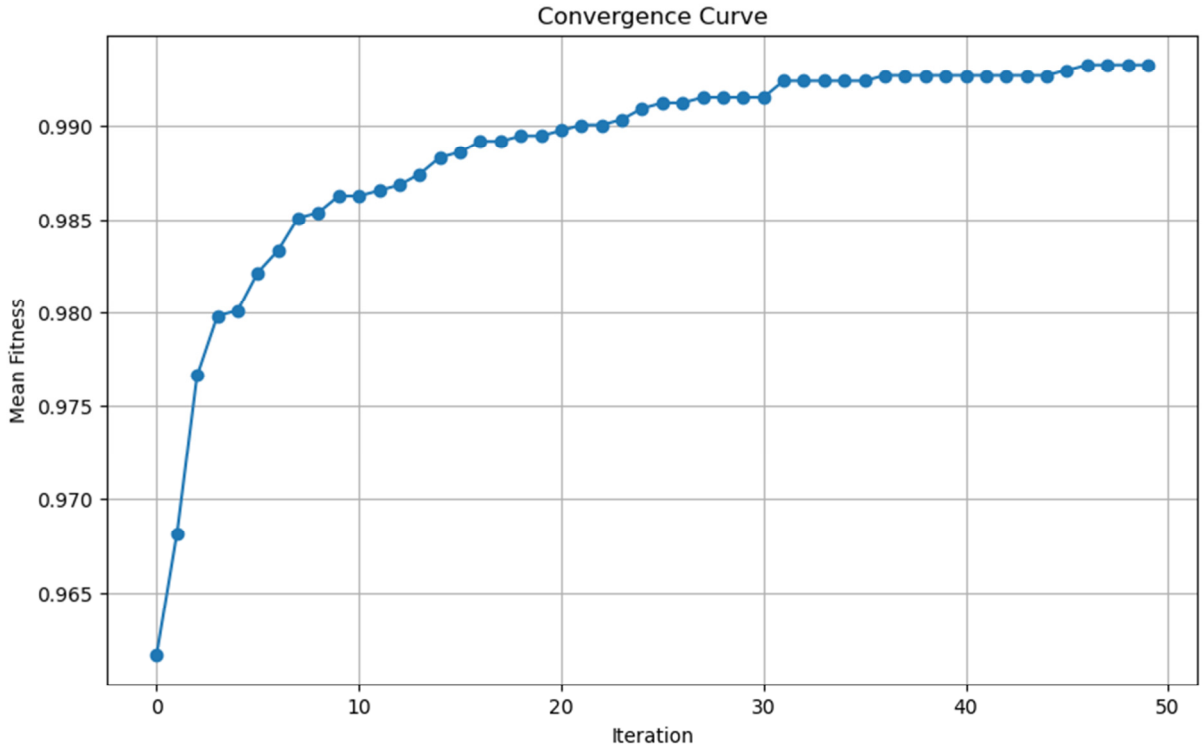


Fig. 2. Convergence curve of the BPSO algorithm.

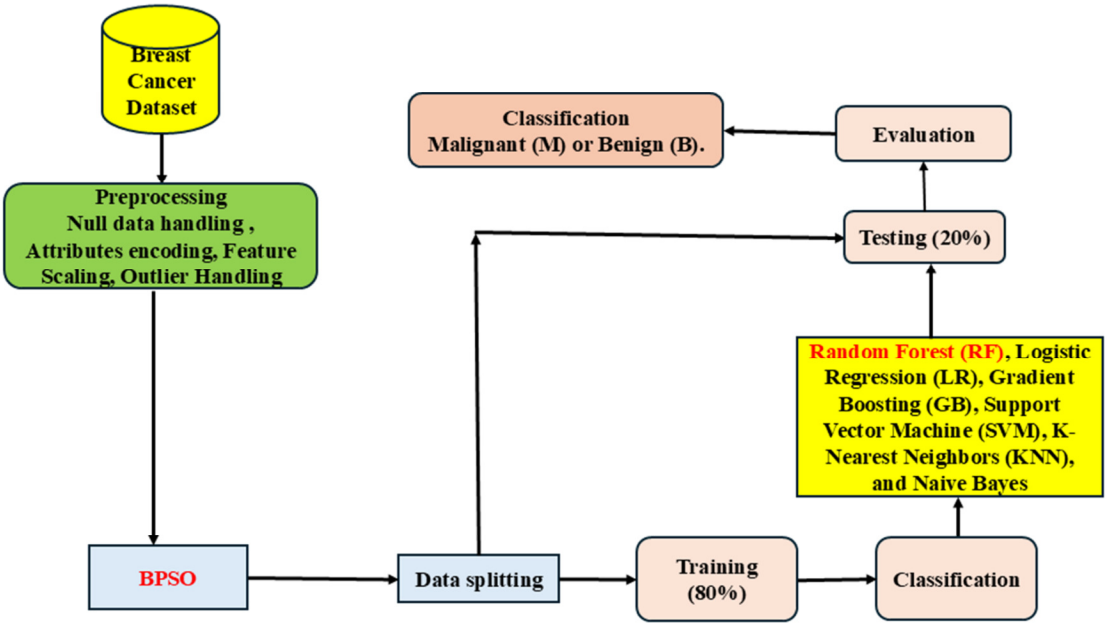


Fig. 3. The methodology of the proposed approach.

III. RESULTS AND DISCUSSION

This section presents the experimental results and evaluates the performance of the methods applied to breast cancer diagnosis. Key metrics such as accuracy, precision, recall, F1-score, and AUC are used for comparison, detailed in Table III. These insights highlight the models' potential for accurate and efficient breast cancer detection, offering a balanced view of their practical utility [29, 30]. The evaluation metric is calculated with the following equations:

$$\text{Accuracy} = \frac{\text{TPos}+\text{TNeg}}{\text{TPos}+\text{FPos}+\text{FNeg}+\text{TNeg}} \quad (1)$$

$$\text{Precision} = \frac{\text{TPos}}{\text{TPos}+\text{FPos}} \quad (2)$$

$$\text{Recall} = \frac{\text{TPos}}{\text{TPos}+\text{FNeg}} \quad (3)$$

$$\text{F1} - \text{score} = \frac{2\times\text{Recall}\times\text{Precision}}{\text{Recall}+\text{Precision}} \quad (4)$$

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR})\, d(\text{FPR}) \quad (5)$$

where:

- $\text{TPR} = \frac{\text{True Positives (TPos)}}{\text{True Positives (TPos)}+\text{False Negatives (FNeg)}}$
- $\text{FPR} = \frac{\text{False Positives (FPos)}}{\text{False Positives (FPos)}+\text{True Negatives (TNeg)}}$

TABLE III. THE PERFORMANCE EVALUATION BETWEEN ML MODELS

Model	Accuracy	Precision	Recall	F1-score	ROC-AUC
RF	0.992456	1.000000	0.982476	0.986878	1.000000
LR	0.982456	0.976744	0.976744	0.976744	0.999345
GB	0.973684	0.976190	0.953488	0.964706	0.998690
SVM	0.947368	1.000000	0.860465	0.925000	0.993449
KNN	0.947368	0.974359	0.883721	0.926829	0.994104
NB	0.956140	1.000000	0.883721	0.938272	0.989191

Table III summarizes the performance evaluation of 6 ML models used for breast cancer classification experiment after BPSO feature selection. Among all the models, RF achieved the best performance, with an accuracy of 99.25%, precision of 100%, recall of 98.25%, F1-score of 98.68%, and ROC-AUC score of 1.0, indicating exceptional classification ability without compromising sensitivity and specificity. LR followed closely, achieving an accuracy of 98.25%, precision and recall of 97.67%, and an F1-score of 97.67%, with a ROC-AUC of 0.9993. The GB model showed slightly lower performance, with an accuracy of 97.37%, precision of 97.61%, and recall of 95.35%, leading to an F1-score of 96.47% and a ROC-AUC of 0.9987. SVM and KNN both achieved an accuracy of 94.74%, but differed in recall (86.04% for SVM and 88.37% for KNN) and precision (100% for SVM and 97.43% for KNN), resulting in F1-scores of 92.50% and 92.68%, respectively. Lastly, NB demonstrated solid performance, with an accuracy of 95.61%, a perfect precision of 100%, a recall of 88.37%, an F1-score of 93.83%, and a ROC-AUC of 0.9891.

The results indicate that RF achieved superior performance across all evaluation metrics, including perfect precision and ROC-AUC, demonstrating its robustness and suitability for breast cancer classification tasks. This can be attributed to RF's ensemble structure, which mitigates overfitting and improves generalization by averaging multiple decision trees. LR and GB also performed well, showing that even simpler or additive models can achieve high predictive accuracy when combined with optimal feature subsets. The lower recall values observed for SVM and KNN suggest sensitivity to the reduced feature set or to specific class imbalances, despite the overall effectiveness of BPSO in selecting relevant features.

Figure 4 presents the ROC curve, a diagram for assessing classification model performance, which shows the true positive rate compared to the false positive rate for different threshold values. Higher AUC indicates better performance. It

is useful for unbalanced datasets and for selecting classification thresholds.

Figure 5 displays the confusion matrices for the applied models, offering a visual representation of their performance in binary [31-36]. This approach assists in identifying strengths, such as handling class imbalances, and in highlighting areas that require improvement, such as reducing specific error types. This comprehensive analysis supports the selection of the most effective model for the task, ensuring reliable and accurate classification outcomes.

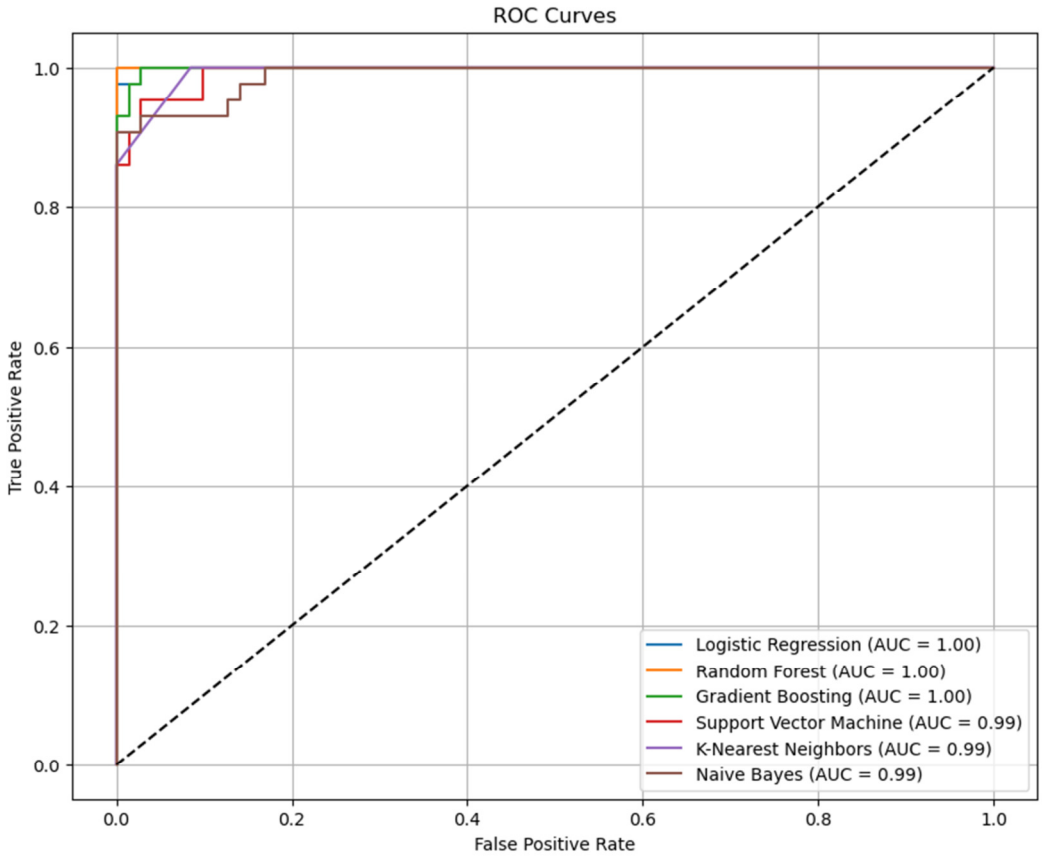
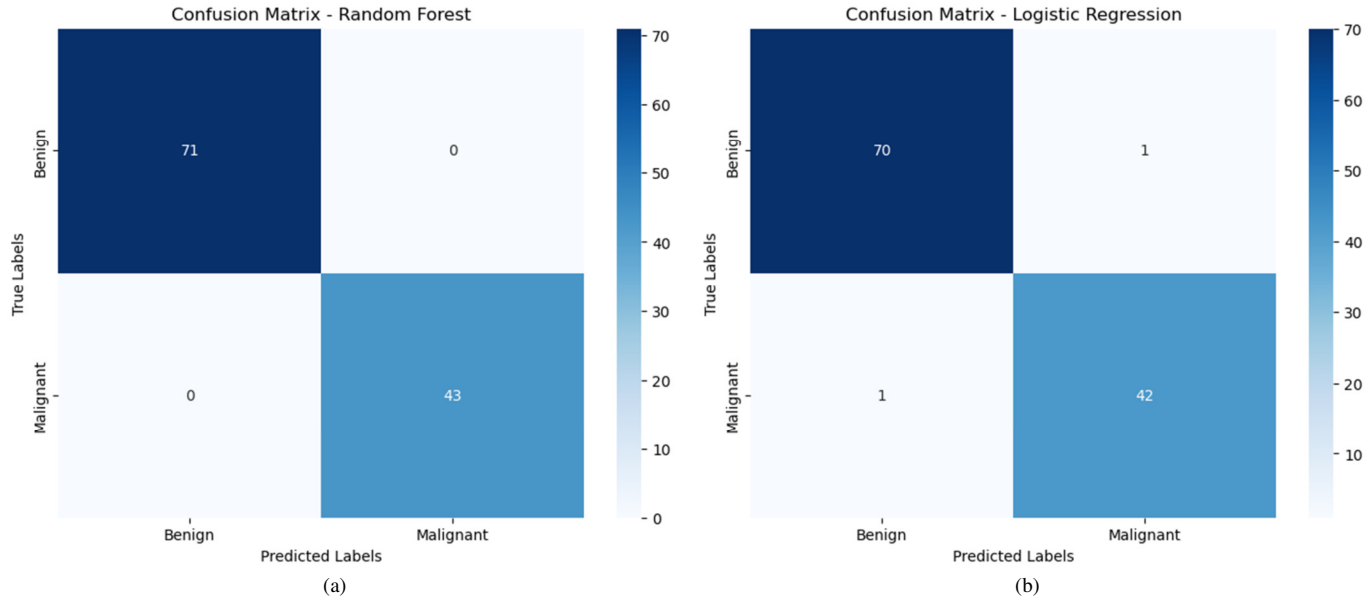
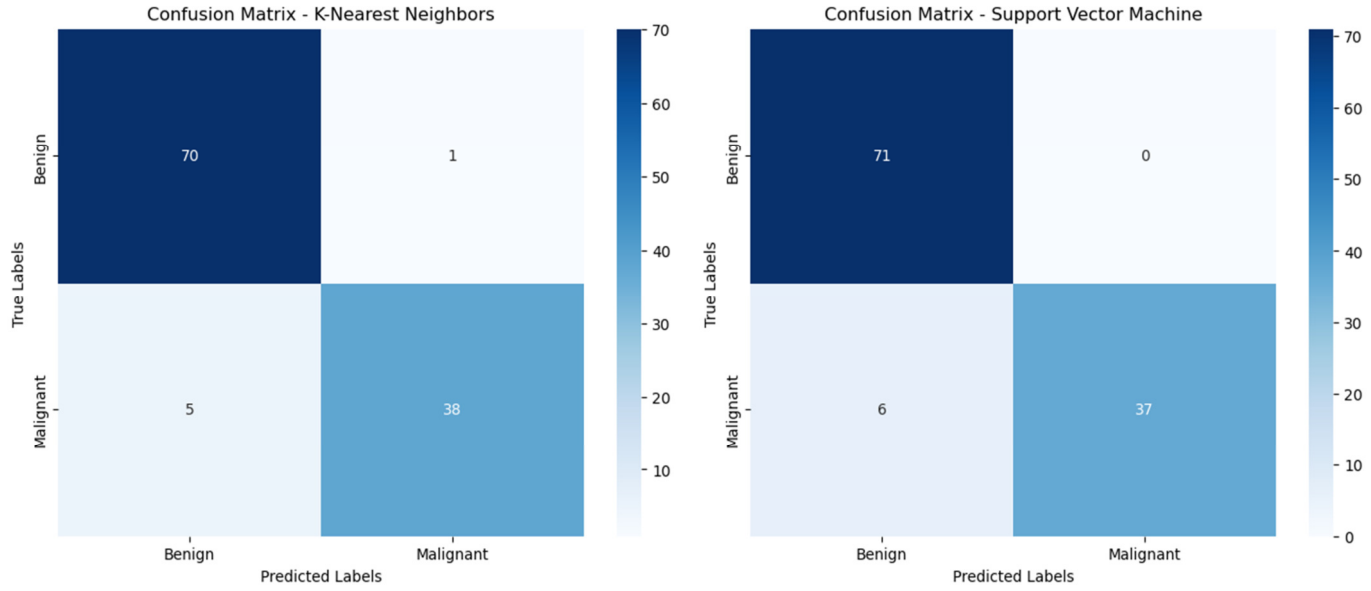


Fig. 4. ROC curve of the suggested ML models.

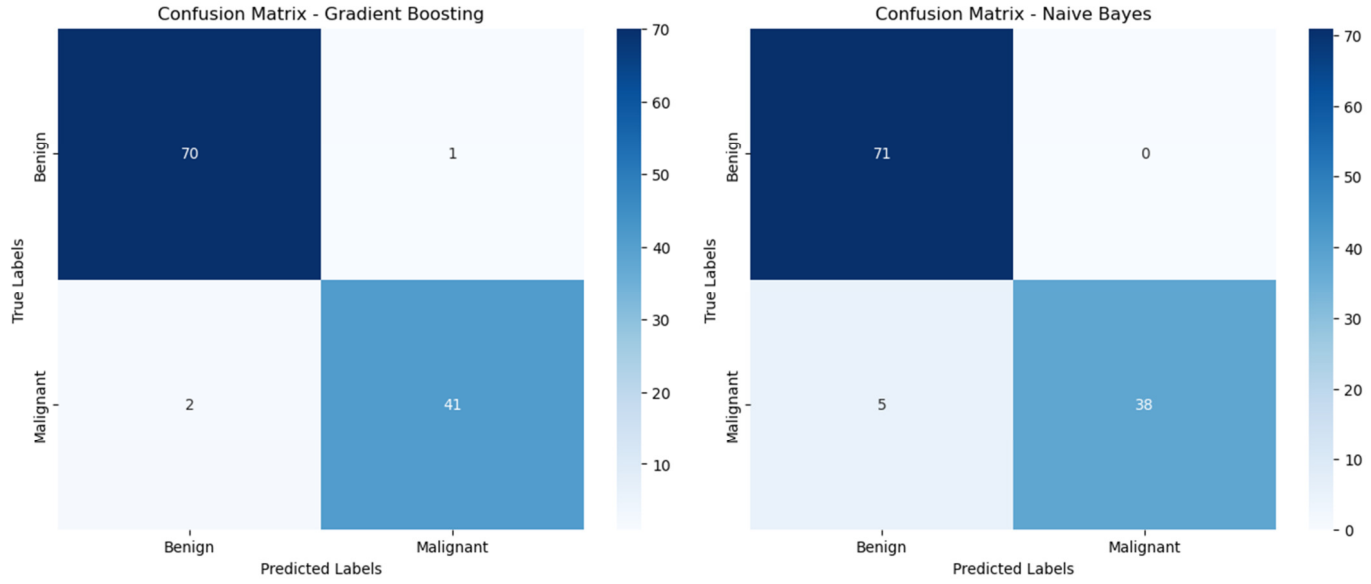


(a) (b)



(c)

(d)



(e)

(f)

Fig. 5. The confusion metrices of the suggested ML models: (a) RF, (b) LR, (c) KNN, (d) SVM, (e) GB, and (f) NB.

Table IV presents a comparative analysis with other studies. It includes four studies, with the first being "our Study," which employs RF with BPSO and achieves the highest accuracy of 99.2%. The second study [37] applies a Multilayer Perceptron (MLP) model with 5-fold cross-validation, achieving 99.12% accuracy. The third study [38] utilizes KNN and an ANN, reporting two accuracy values: 97.7% and 98.6%. Lastly, the fourth study [39] implements Principal Component Analysis (PCA) and Tomek link resampling with KNN, resulting in an accuracy of 95.29%. This comparison highlights the effectiveness of RF with BPSO in achieving superior classification performance compared to other techniques.

TABLE IV. COMPARATIVE ANALYSIS WITH OTHER STUDIES

Ref.	Methodology	Accuracy (%)
Our study	RF with BPSO	99.2
[37]	MLP with 5-fold cross-validation	99.12
[38]	KNN and ANN	97.7 and 98.6
[39]	PCA and Tomek link with KNN	95.29

IV. CONCLUSION AND FUTURE WORK

The present study demonstrates the efficacy of integrating Binary Particle Swarm Optimization (BPSO) with Machine Learning (ML) models for breast cancer classification. By reducing the dataset's dimensionality while retaining critical

diagnostic features, the proposed Random Forest (RF)-BPSO model achieves an optimal classification accuracy of 99.2%. The results indicate that feature selection has a substantial impact on enhancing model performance, primarily by improving classification precision and reducing computational overhead. Furthermore, a comparative analysis with recent methodologies confirms the superiority of the proposed approach, as it outperforms several state-of-the-art techniques [37-39] in terms of accuracy and robustness. Future work should explore deep learning architectures such as Convolutional Neural Networks (CNNs) and transformers, to further improve breast cancer classification. In addition, the practical deployment of the proposed RF-BPSO model necessitates attention to critical aspects such as scalability, interoperability, and regulatory compliance. The model's low computational overhead makes it inherently scalable for larger patient datasets and real-time diagnosis scenarios. To ensure interoperability, future extensions should focus on integrating the model with existing electronic health record systems using healthcare data exchange standards such as HL7 or FHIR. Additionally, real-world applications must consider regulatory frameworks such as the Health Insurance Portability and Accountability Act (HIPAA), the General Data Protection Regulation (GDPR), and guidelines established by regional health authorities. Incorporating privacy-preserving mechanisms and explainable AI components can further support ethical and compliant deployment in clinical environments, reinforcing the model's readiness for translation into practical healthcare solutions.

The study's limitations include its experimental evaluation based on a single publicly available dataset, the use of BPSO, the assumption that all input features are accurate and unbiased, the lack of exploration of class imbalance mitigation techniques beyond basic performance metrics, and the absence of external validation on independent datasets and real-time clinical deployment scenarios. These factors may limit the generalizability of the findings to other data types, such as medical imaging or genomic profiles.

REFERENCES

[1] B. Ajlouni, T. Mukattash, A. Al-Nabulsi, R. A. Farha, W. Ta'an, and R. Itani, "Evaluating Nutrition-related Knowledge, Attitudes, and Practices for the Prevention of Breast Cancer among Women in Jordan," *Jordan Journal of Nursing Research*, vol. 2, no. 4, pp. 289–298, Dec. 2023, https://doi.org/10.14525/JJNR.v2i4.07.

[2] Z. Lafi *et al.*, "Synergistic combination of doxorubicin with hydralazine, and disulfiram against MCF-7 breast cancer cell line," *Plos One*, vol. 18, no. 9, Sep. 2023, Art. no. e0291981, https://doi.org/10.1371/journal.pone.0291981.

[3] A. A. Tawil, A. Shaban, and L. Almazaydeh, "A comparative analysis of convolutional neural networks for breast cancer prediction," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 3, pp. 3406–3414, Jun. 2024, https://doi.org/10.11591/ijece.v14i3.pp3406-3414.

[4] A. A. Alhussan *et al.*, "Classification of Breast Cancer Using Transfer Learning and Advanced Al-Biruni Earth Radius Optimization," *Biomimetics*, vol. 8, no. 3, Jul. 2023, Art. no. 270, https://doi.org/10.3390/biomimetics8030270.

[5] S. Thirumalaisamy *et al.*, "Breast Cancer Classification Using Synthesized Deep Learning Model with Metaheuristic Optimization Algorithm," *Diagnostics*, vol. 13, no. 18, Sep. 2023, Art. no. 2925, https://doi.org/10.3390/diagnostics13182925.

[6] K. M. M. Uddin, N. Biswas, S. T. Rikta, and S. K. Dey, "Machine learning-based diagnosis of breast cancer utilizing feature optimization technique," *Computer Methods and Programs in Biomedicine Update*, vol. 3, Jan. 2023, Art. no. 100098, https://doi.org/10.1016/j.cmpbup.2023.100098.

[7] M. Kumar, S. Singhal, S. Shekhar, B. Sharma, and G. Srivastava, "Optimized Stacking Ensemble Learning Model for Breast Cancer Detection and Classification Using Machine Learning," *Sustainability*, vol. 14, no. 21, Nov. 2022, Art. no. 13998, https://doi.org/10.3390/su142113998.

[8] M. S. A. Reshan *et al.*, "Enhancing Breast Cancer Detection and Classification Using Advanced Multi-Model Features and Ensemble Machine Learning Techniques," *Life*, vol. 13, no. 10, Oct. 2023, Art. no. 2093, https://doi.org/10.3390/life13102093.

[9] M. Obayya *et al.*, "Hyperparameter Optimizer with Deep Learning-Based Decision-Support Systems for Histopathological Breast Cancer Diagnosis," *Cancers*, vol. 15, no. 3, Feb. 2023, Art. no. 885, https://doi.org/10.3390/cancers15030885.

[10] E. Michael, H. Ma, H. Li, and S. Qi, "An Optimized Framework for Breast Cancer Classification Using Machine Learning," *BioMed Research International*, vol. 2022, no. 1, Feb. 2022, Art. no. 8482022, https://doi.org/10.1155/2022/8482022.

[11] M. R. Islam *et al.*, "Enhancing breast cancer segmentation and classification: An Ensemble Deep Convolutional Neural Network and U-net approach on ultrasound images," *Machine Learning with Applications*, vol. 16, Jun. 2024, Art. no. 100555, https://doi.org/10.1016/j.mlwa.2024.100555.

[12] S. Chakravarthy, B. Nagarajan, V. V. Kumar, T. R. Mahesh, R. Sivakami, and J. R. Annand, "Breast Tumor Classification with Enhanced Transfer Learning Features and Selection Using Chaotic Map-Based Optimization," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, Feb. 2024, Art. no. 18, https://doi.org/10.1007/s44196-024-00409-8.

[13] S. M. Sylviaa and N. Sudha, "Enhancing Breast Cancer Classification: A Deep Learning Approach with Outlier Detection with Egret Swarm Optimization Algorithm and Feature Selection Integratio," *Journal of Angiotherapy*, vol. 8, no. 3, pp. 1–13, Mar. 2024, https://doi.org/10.25163/angiotherapy.839520.

[14] C. S. Rajpoot, G. Sharma, P. Gupta, P. Dadheech, U. Yahya, and N. Aneja, "Feature Selection-based Machine Learning Comparative Analysis for Predicting Breast Cancer," *Applied Artificial Intelligence*, vol. 38, no. 1, Dec. 2024, Art. no. 2340386, https://doi.org/10.1080/08839514.2024.2340386.

[15] R. Gurumoorthy and M. Kamarasan, "Breast Cancer Classification from Histopathological Images using Future Search Optimization Algorithm and Deep Learning," *Engineering, Technology & Applied Science Research*, vol. 14, no. 1, pp. 12831–12836, Feb. 2024, https://doi.org/10.48084/etasr.6720.

[16] Md. R. Islam, Md. S. Islam, and S. Majumder, "Breast Cancer Prediction: A Fusion of Genetic Algorithm, Chemical Reaction Optimization, and Machine Learning Techniques," *Applied Computational Intelligence and Soft Computing*, vol. 2024, no. 1, Aug. 2024, Art. no. 7221343, https://doi.org/10.1155/2024/7221343.

[17] A. Bilal, A. Imran, T. I. Baig, X. Liu, E. Abouel Nasr, and H. Long, "Breast cancer diagnosis using support vector machine optimized by improved quantum inspired grey wolf optimization," *Scientific Reports*, vol. 14, no. 1, May 2024, Art. no. 10714, https://doi.org/10.1038/s41598-024-61322-w.

[18] S. Thawkar, V. Katta, A. R. Parashar, L. K. Singh, and M. Khanna, "Breast cancer: A hybrid method for feature selection and classification in digital mammography," *International Journal of Imaging Systems and Technology*, vol. 33, no. 5, pp. 1696–1712, Sep. 2023, https://doi.org/10.1002/ima.22889.

[19] A. Bansal, V. K. Lohan, M. Khanna, and S. Agnihotri, "A Novel Efficient Approach for Feature Selection for Enhanced Performance in Breast Cancer Prediction," in *4th International Conference on Advanced Network Technologies and Intelligent Computing*, Varanasi, India, 2024, pp. 478–488, https://doi.org/10.1007/978-3-031-83793-7_30.

[20] L. K. Singh, M. Khanna, and R. Singh, "An enhanced soft-computing based strategy for efficient feature selection for timely breast cancer prediction: Wisconsin Diagnostic Breast Cancer dataset case," *Multimedia Tools and Applications*, vol. 83, no. 31, pp. 76607–76672, Sep. 2024, https://doi.org/10.1007/s11042-024-18473-9.

[21] "Breast Cancer Dataset." Kaggle. [Online]. Available: https://www.kaggle.com/datasets/adhamelkomy/breast-cancer.

[22] M. Kaddes, Y. M. Ayid, A. M. Elshewey, and Y. Fouad, "Breast cancer classification based on hybrid CNN with LSTM model," *Scientific Reports*, vol. 15, no. 1, Feb. 2025, Art. no. 4409, https://doi.org/10.1038/s41598-025-88459-6.

[23] M. Y. Shams, Z. Tarek, and A. M. Elshewey, "A novel RFE-GRU model for diabetes classification using PIMA Indian dataset," *Scientific Reports*, vol. 15, no. 1, Jan. 2025, Art. no. 982, https://doi.org/10.1038/s41598-024-82420-9.

[24] Z. Tarek, A. A. Alhussan, D. S. Khafaga, E.-S. M. El-Kenawy, and A. M. Elshewey, "A snake optimization algorithm-based feature selection framework for rapid detection of cardiovascular disease in its early stages," *Biomedical Signal Processing and Control*, vol. 102, Apr. 2025, Art. no. 107417, https://doi.org/10.1016/j.bspc.2024.107417.

[25] E.-S. M. Elkenawy, A. A. Alhussan, D. S. Khafaga, Z. Tarek, and A. M. Elshewey, "Greylag goose optimization and multilayer perceptron for enhancing lung cancer classification," *Scientific Reports*, vol. 14, no. 1, Oct. 2024, Art. no. 23784, https://doi.org/10.1038/s41598-024-72013-x.

[26] A. M. Elshewey, R. Y. Youssef, H. M. El-Bakry, and A. M. Osman, "Water potability classification based on hybrid stacked model and feature selection," *Environmental Science and Pollution Research*, vol. 32, no. 13, pp. 7933–7949, Mar. 2025, https://doi.org/10.1007/s11356-025-36120-0.

[27] O. I. Ramadan *et al.*, "Co-administration of either curcumin or resveratrol with cisplatin treatment decreases hepatotoxicity in rats via anti-inflammatory and oxidative stress-apoptotic pathways," *PeerJ*, vol. 12, Jul. 2024, Art. no. e17687, https://doi.org/10.7717/peerj.17687.

[28] L. S. Ali, G. Mohamed, M. Akeel, K. S. ElBayoumi, and A. E.-F. B. M. El-Beltagy, "Zingiber officinale ethanolic extract improved organs function in lipopolysaccharides-induced organ toxicity by modulating inflammation and oxidative stress in male rats," *Egyptian Journal of Basic and Applied Sciences*, vol. 11, no. 1, pp. 213–231, Dec. 2024, https://doi.org/10.1080/2314808X.2024.2332849.

[29] A. Jafari, "Machine-learning methods in detecting breast cancer and related therapeutic issues: a review," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 12, no. 1, Dec. 2024, Art. no. 2299093, https://doi.org/10.1080/21681163.2023.2299093.

[30] Y. Guo *et al.*, "Machine learning and new insights for breast cancer diagnosis," *Journal of International Medical Research*, vol. 52, no. 4, Apr. 2024, Art. no. 03000605241237867, https://doi.org/10.1177/03000605241237867.

[31] S. A. Z. Hassan, "An AI healthcare ecosystem framework for Covid-19 detection and forecasting using CronaSona," *Medical & Biological Engineering & Computing*, vol. 62, no. 7, pp. 1959–1979, Jul. 2024, https://doi.org/10.1007/s11517-024-03058-3.

[32] S. A. Z. Hassan, "MemorySona: Illuminating Cognitive Health with Deep Learning - A Mobile Medical App for Alzheimer's Patients, Emphasizing Detection through Brain MRI Images," in *2024 6th International Conference on Computing and Informatics*, New Cairo - Cairo, Egypt, 2024, pp. 83–90, https://doi.org/10.1109/ICCI61671.2024.10485057.

[33] L. B. Ammar, "Enhanced Diagnosis of Lung Cancer through an Ensemble Learning Model leveraging an Adaptive Optimization Algorithm," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 18518–18524, Dec. 2024, https://doi.org/10.48084/etasr.9096.

[34] E.-S. M. El-Kenawy, N. Khodadadi, A. Ibrahim, M. M. Eid, A. M. Osman, and A. M. Elshewey, "An optimized model for Liver disease classification based on BPSO Using Machine learning models," *Mesopotamian Journal of Computer Science*, vol. 2024, pp. 214–223, Dec. 2024, https://doi.org/10.58496/MJCSC/2024/017.

[35] A. M. Elshewey and A. M. Osman, "Orthopedic disease classification based on breadth-first search algorithm," *Scientific Reports*, vol. 14, no. 1, Oct. 2024, Art. no. 23368, https://doi.org/10.1038/s41598-024-73559-6.

[36] Y. Fouad, A. M. Osman, S. A. Z. Hassan, H. M. El-Bakry, and A. M. Elshewey, "Adaptive Visual Sentiment Prediction Model Based on Event Concepts and Object Detection Techniques in Social Media," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 7, pp. 252-256, 2023, https://doi.org/10.14569/IJACSA.2023.0140728.

[37] S. Aamir *et al.*, "Predicting Breast Cancer Leveraging Supervised Machine Learning Techniques," *Computational and Mathematical Methods in Medicine*, vol. 2022, no. 1, Aug. 2022, Art. no. 5869529, https://doi.org/10.1155/2022/5869529.

[38] T. Khater *et al.*, "An Explainable Artificial Intelligence Model for the Classification of Breast Cancer," *IEEE Access*, vol. 13, pp. 5618–5633, 2025, https://doi.org/10.1109/ACCESS.2023.3308446.

[39] S. B. Manir and P. Deshpande, "Critical Risk Assessment, Diagnosis, and Survival Analysis of Breast Cancer," *Diagnostics*, vol. 14, no. 10, May 2024, Art. no. 984, https://doi.org/10.3390/diagnostics14100984.