Indonesian Journal of Electrical Engineering and Computer Science
Vol. 37, No. 2, February 2025, pp. 849~858
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v37.i2.pp849-858

Breast cancer prediction using genetic algorithm and sand cat swarm optimization algorithm

Velu Sangeetha[1], Maniyambadi Vaneeta[2], Arjuna Mamatha[1], Muniya Shoba[3], Sugatur Ramamurthy Deepa[4], Velusamy Sujatha[5], Shanmugam Sujatha[6]
[1]Department of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore, India
[2]Department of Artificial Intelligence and Data Science, Ramaiah Institute of Technology, Bangalore, India
[3]Information Science and Engineering, BMS Institute of Technology and Management, Bangalore, India
[4]Department of Computer Science and Design, K. S. Institute of Technology, Bangalore, India
[5]Department of Electronics and Communication Engineering, Shree Sathyam College of Engineering and Technology, Sangagiri, India
[6]Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India

Article Info

Article history:

Received Jul 4, 2024
Revised Sep 6, 2024
Accepted Sep 30, 2024

Keywords:

Breast cancer
Classification
Feature selection
Optimization
Women

ABSTRACT

Breast cancer is the second leading type of cancer, which is mainly found in women and which increases the death rate among women. Early detection and diagnosis of breast cancer can reduce its occurrence and the death rate. Unfortunately, even if cancer treatment is initiated quickly after diagnosis, cancer may relapse because cancer cells may continue to exist in the body, which is also a major problem faced by women who fear facing the same treatment twice. So, detecting cancer at its early stage and predicting the recurrence of it is a major issue in the medical field that needs to be solved. Machine learning (ML) algorithms such as support vector machine (SVM), random forest (RF), k-nearest neighbor (KNN), and voting classifier (VC) are used for breast cancer prediction. Due to high-dimension data, the predicted results using Machine learning algorithms will increase the errors and decrease the accuracy. So, bioinspired algorithms such as the genetic algorithm (GA) and sand cat swarm optimization (SCSO) are used to reduce the data dimension. Convolutional neural network (CNN) is used for feature extraction from the image dataset. CNN algorithms are used for feature selection, which selects the important features for classification and prediction by applying 10 cross-validation methods. The proposed model using bioinspired optimization algorithms outcomes will yield high accuracy and the best solution.

Corresponding Author:

Velu Sangeetha
Department of Computer Science and Engineering, Ramaiah Institute of Technology
Bangalore, Karnataka, India
Email: drsangeethav@msrit.edu

1. INTRODUCTION

Ninety precent of cancer patients dead from metastasis that is an unrestrained cell development which can rapidly move to other human body parts. Cancer develops when genetic alterations known as mutations, which control cell growth, manifest and grow uncontrollably [1]. One of the greatest health hazards people face today is cancer [2]. The second major cause of death worldwide is cancer. A right cancer diagnosing is necessitated for a appropriate and effectual course of therapy. A dangerous development that occurs in the breast cells is called breast cancer [3]. For the diagnosis of cancer in tissues, a breast biopsy, thermography, fine-needle cytology, ultrasonography, mammography, and other techniques are utilized.

According to current medical theory, mammography is the most trustworthy approach [4]. Early identification and precise classification of breast cancer tumors are an essential to good treatment planning as well as patient prognosis. Prediction of cancer recurrence plays a vital role in determining appropriate follow-up care for patients [5]. The main objetive of the proposed system is to growth a robust and accurate model for breast cancer classification and recurrence diagnosis, utilizing bio-inspired algorithms. Deep learning techniques such as the convolutional neural network (CNN) model are used for feature extraction [6]. SCSO optimization techniques are used for feature optimization. Support vector machine (SVM), k-nearest neighbor (KNN), random forest (RF), and voting classifier (VC) machine learning algorithms are applied for breast cancer classification. Genetic algorithms are used to optimize the classification model [7].

The contributions of proposed method is specified: develop a breast cancer classification model capable of accurately distinguishing between malignant and benign breast tumors based on diverse clinical and diagnostic features. Utilized a genetic algorithm to optimize the classification model's performance, ensuring high accuracy, precision, and recall. Apply bio-inspired algorithms such as sand cat swarm optimization (SCSO) to identify significant recurrence predictors by analyzing a wide range of patient-specific factors, including genetic markers, treatment history, and post-treatment follow-up data.

The paper is structured as follows: section 2 discusses existing and prior research efforts, providing an overview of related work. Section 3 outlines the dataset used in this study and the methods and methodology of the proposed study. Section 4 is dedicated to presenting the results and conducting an evaluation analysis of the performance. Finally, in section 5, conclusions based on the findings of this study are made.

Many technologies are emerging in computer science and have been utilized in the medical field in recent years. So, it is essential to understand which technology must be applied to different diseases found in the medical field [8]. Machine learning (ML) and artificial intelligence (AI) are emerging technologies used in the medical field for all kinds of diseases and cancers throughout recent years [9]. ML algorithms are used as both classifiers and regressors to predict the presence of cancer by giving some inputs to the developed model [10]. However, as the data generated in the medical field is quite large and of high dimension, the predicted accuracy is becoming low, in addition to many errors in the developed model. This leads to the use of Bio-inspired algorithms to lessen the higher data dimension. The paper [11] proposes an ML model for the classification of breast cancer. Throughout the paper [12], an automated decision-making pipeline for BC detection has been proposed, incorporating ML algorithms like gaussian Naive Bayes (GNB), RF, XGBoost (XGB), AdaBoost (AdB), and pre-processing such as outlier rejection (OR) and attribute selection (AS) to increase the survival rate of the patient suffering from BC [13].

The paper [14] proposed a computer-aided ensemble method for diagnosing BC using ReNet18 and SVM. A pre-trained ReNet18 model is used to extract the features from the X-ray image, and SVM is utilized in diagnosing cancer. The paper's author used an MPSO with an SVM on a wisconsin breast cancer UCI dataset to detect breast cancer. The proposed method can also pinpoint useful breast cancer characteristics [15]. The presented MPSO utilizes an adaptive strategy to enhance performance while solving difficult issues and preventing the premature convergence characteristic of PSO. The research gap identified is that most of the researchers use Bio-inspired algorithms such as grey wolf optimization, SCSO, artificial rabbit optimization, and brainstorm optimization algorithms [16]. The study discusses which feature selection algorithm performs better for each classification method and provides high accuracy [17].

2. BREAST CANCER PREDICTION USING GENETIC ALGORITHM AND SCSO ALGORITHM

2.1. Breast cancer classification

Figure 1 explains the research methodology for the classification system. It contains 5 phases. The first phase of this methodology was the data-collecting phase, during which the necessary information for the investigation was gathered [18]. The data preprocessing stage was the second phase when the gathered data was combined, cleaned, and processed to make the datasets acceptable for classification prediction, as shown in Figure 1.

After that, while still in the second phase, a genetic algorithm [19] was used to extract features. Once the preprocessing phase [20] (Phase 2) was complete, the data were passed over to Phase 3, where they were used to build models both with and without feature extraction [21]. Phase 4 involved the evaluation of the model or result analysis. In the final phase of the research, a comparison of the feature-selection-based models versus models that did not employ feature selection was done [22].

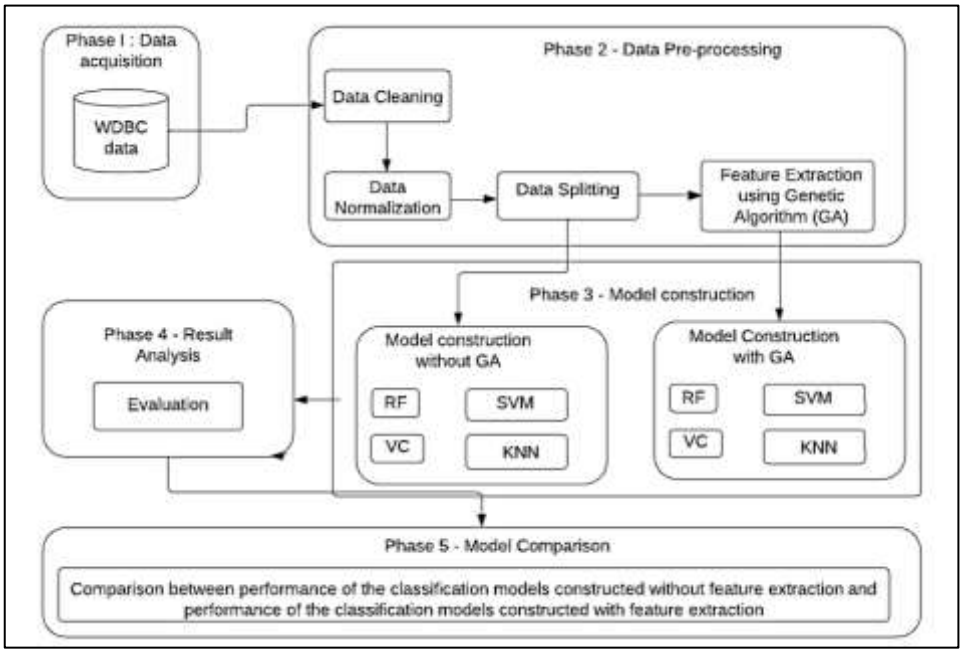

Figure 1. Research methodology for classification system

2.2. Dataset description

The wisconsin diagnostic breast cancer (WDBC) dataset is a widely used dataset in machine learning and breast cancer research [23]. Dr. William H. Wolberg created dataset contains features computed from digitized images of breast masses' fine needle aspirates (FNA). The goal of this work to predict whether a given mass is benign (non-cancerous) or malignant (cancerous). Mammography images are used to distinguish the cells as recurrence or non-recurrence. The total number of images used is 500, of which 250 images that tell the cells show recurrence signs and 250 images that show non-recurrence signs [24].

2.3. Algorithms

2.3.1. Genetic algorithm (GA)

GA is a search and optimization technique [25] inspired by the evolution and natural selection processes. It is a type of evolutionary algorithm that uses the ideas of genetics and natural selection to look for optimum or almost optimal solutions to challenging issues. A genetic algorithm effectively imitates the process of natural selection by developing a population of candidate solutions over numerous generations. A chromosome, which is analogous to the DNA in biological things and commonly encoded as a string of binary digits (0s and 1s) in Algorithm 1, is used to represent each solution in the population [26].

Algorithm 1. Genetic_algorithm

```
Genetic_Algorithm()
{
        Initialize random population
        Evaluate the population
        Generation=0
While the termination criterion is not satisfied
{
        Generation=Generation + 1
        Select good chromosomes by reproduction procedure
        Perform crossover with probability crossover (Pc)
        Perform mutation with probability mutation (Pm)
Evaluate the population
        }
}
```

2.3.2. Sand cat swarm optimization (SCSO)

In the SCSO algorithm [27], the search for prey is empowered by each sand cat's ability in the low-frequency noise emission. The range (R) for sensitivity of every search agent is preset among 2 to 0. In the SCSO algorithm in Algorithm 2, each search agent moves circularly; intimal, to motion in several

directions to seek for feasible global solutions, the arbitrary angle α among 0 to 360 is represented as a cos(α) function. The structural formulas of the SCSO is specified (1) to (8).

$$\overrightarrow{r_G} = s_M - \left(\frac{s_M \times iter_c}{iter_{max}}\right) \quad (1)$$

$$R = 2 \times \overrightarrow{r_G} \times rand(0,1) - \overrightarrow{r_G} \quad (2)$$

$$\vec{r} = \overrightarrow{r_G} \times rand(0,1) \quad (3)$$

$$\vec{X}(t+1) = 2 \times \overrightarrow{r_G} \times rand(0,1) - \overrightarrow{r_G} \quad (4)$$

$$\vec{X}(t+1) = \vec{r} \cdot \left(\overrightarrow{X_b}(t) - rand(0,1) \cdot -\overrightarrow{X_c}(t)\right) \quad (5)$$

$$\overrightarrow{X_{rnd}} = \left|rand(0,1) \cdot \overrightarrow{X_b}(t) - \overrightarrow{X_c}(t)\right| \quad (6)$$

$$\vec{X}(t+1) = \overrightarrow{X_b}(t) - \vec{r} \cdot \overrightarrow{X_{rnd}} \cos(\alpha) \quad (7)$$

$$\vec{X}(t+1) = \begin{cases} \vec{r} \cdot \left(\overrightarrow{X_b}(t) - rand(0,1) \cdot -\overrightarrow{X_c}(t)\right) & |R| > 1 \\ \overrightarrow{X_b}(t) - \vec{r} \cdot \overrightarrow{X_{rnd}} \cos(\alpha) & |R| \leq 1 \end{cases} \quad (8)$$

Algorithm 2. SCSO

```
Step1: Initialize:
           Initialize a population of sand cats with random positions
           Evaluate the fitness of every sand cat
Step2:  Repeat until a termination condition is met:
              For each sand cat:
                  Update the sand cat's position using the following equation:
       newPosition=current position + random() * (best position-current position)
Step 3: Apply position constraints to ensure the sand cat remains within the search space
Step 4: Measure the fitness of the new position
                  If the new position is better than the previous best position:
              Update the sand cat's best position
                  Update global best position if applicable
Step 5: Return global best position
Step 6: End Procedure
```

3. RESULTS AND DISCUSSION

A proposed (GA+VC) predictive model in Algorithm 3 is developed to categorize the cancerous cells into malignant or benign and to tell whether the cancer recurrence has happened for a particular cancer-diagnosed person. GA is applied to the pre-processed dataset to extract the important features from a high-dimensional dataset using different base SVM, KNN, RF, and VC classifiers [25]. Classification is applied to the extracted features for further processing. For recurrence, CNN is used for feature extraction, SCSO is used for feature optimization, and SVM [28] is used to classify the binary values as 0 or 1. The model is trained with the WDBC and WPBC datasets for classification and recurrence, respectively, which provides accurate and efficient results. The proposed model performs better than others in accuracy, precision, recall, and f1-score [29]. The evaluation metrics are compared for all the classification algorithms before the GA is applied to the data shown in Table 1.

Algorithm 3. Proposed model (GA+VC)

```
Input: WDBC and mammography dataset
Output: Classify as Malignant or benign
          Classify as recurrence or non-recurrence
Begin
        If the WDBC dataset performs:
        Begin:
               Perform data collection and pre-processing, then normalize the data.
               Apply GA for feature extraction.
             Build the model using SVM, KNN, RF, VC
            Train the model
            Feed testing data to models and get the results
```

```
            Construct the graphs for the obtained results.
        End
        elif mammography images:
       Perform data collection and pre-processing, then normalize the data.
              Apply CNN for feature extraction.
             SCSO is applied for feature Optimization.
            Build the model using SVM.
           Train the model
             Feed testing data to models and get the results
       Construct the graphs for the obtained results.
         End
End
```

Table 1. Evaluation metrics before applying GA

Models	Accuracy	Precision	Recall	F1-score
SVM	92.26	96.67	82.07	88.77
KNN	95.60	96.98	96.98	96.98
RF	98.90	98.09	97.08	98.04
VC	99.96	98.90	98.87	98.68

Table 1 explains the accuracy, precision, recall, and F1-score percentage of the SVM, KNN, RF, and VC models. Compared to all models, VC has a 99.26% accuracy, 98.80% precision, 98.87% recall, and 98.68% F1-score ratio. All models are not capable of getting 100% values. The evaluation metrics are compared for all the classification algorithms after applying the GA to the data shown in Table 2.

Table 2. Evaluation metrics of models after applying GA

Models	Accuracy	Precision	Recall	F1-score
SVM	94.72	95.5	90.09	92.71
KNN	95.62	95.60	95.60	95.60
RF	100	100	100	100
VC	100	100	100	100

From Table 2, the SVM, KNN, RF, and VC models obtain accuracy, precision, recall, and F1-score percentage values after applying the GA algorithm. Compared to all models, the VC reaches 100% accuracy, precision, recall, and F1-score. In addition, the DR model also reaches 100% for all evaluation parameters. Figure 2 explains the accuracy comparisons of before and after utilizing the GA algorithm.

Figure 2 explains the Accuracy of models before and after utilizing the GA algorithm, such as SVM, KNN, RF, and VC models. Compared to all models, the after-applying GA algorithm provides a high percentage accuracy. However, the RF and VC models provide a higher accuracy than other models. Figure 3 shows a precision comparison of all models before and after applying GA.

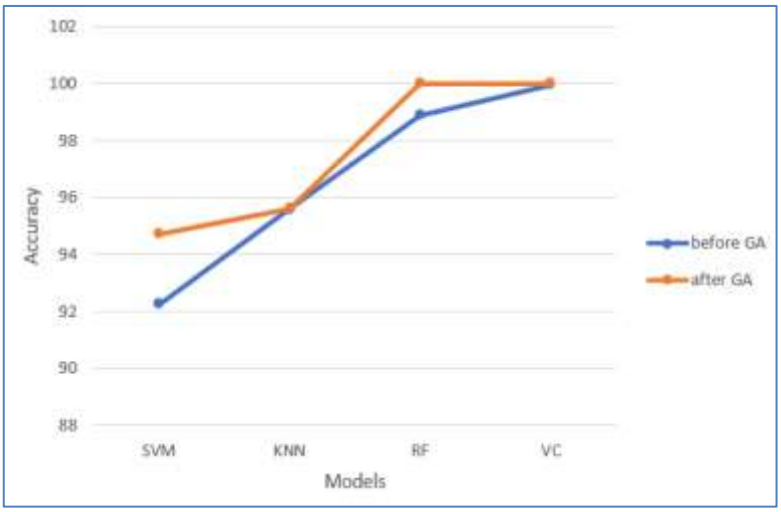

Figure 2. Accuracy comparisons of all models before and after utilizing the GA algorithm

Figure 3 explains the before-and-after use of the GA algorithm for comparing several models, such as SVM, KNN, RF, and VC. The outcomes show that the GA algorithm offers better accuracy than GA models. The SVM and KNN models provide lesser precision than the RF and VC models. Figure 4 demonstrates the Recall comparison of all models before and after applying the GA algorithm.

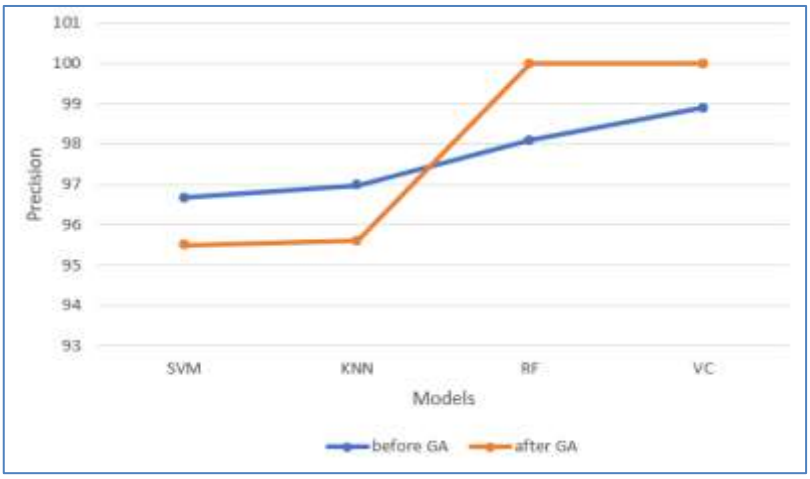

Figure 3. Precision comparison of all models before and after applying GA

Figure 4 explains the recall comparison of SVM, KNN, RF, and VC models before and after applying the GA algorithm. Compared to before and after utilizing the GA algorithm, utilizing the GA algorithm provides a higher recall percentage. However, the SVM model does not apply a GA algorithm, which provides lesser performance. Figure 5 explains the F1-score comparison of all models before and after applying GA models.

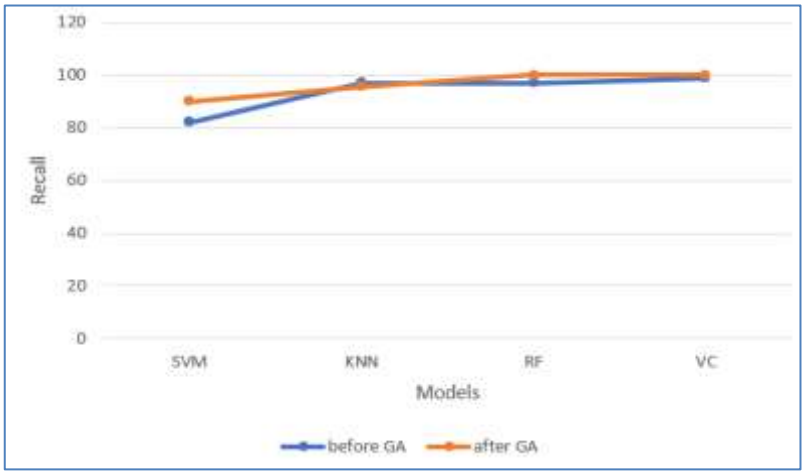

Figure 4. Recall the comparison of all models

Figure 5 explains the F1-score comparison among SVM, KNN, RF, and VC models before and after applying the GA algorithm. Compared to all models, the GA algorithm provides the highest F1 score. Particularly, the RF and VC models have a higher F1 Score than the SVM and KNN models. Figure 6 explains the comparison of proposed model (GA+VC) to the existing LDA-SVM models for classification diagnosis.

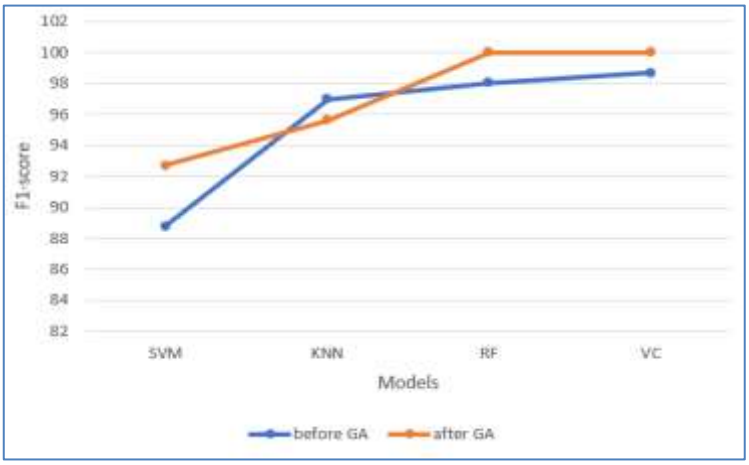

Figure 5. F1-score comparison of all models before and after applying GA

Figure 6 compares the proposed model (GA+VC) to the existing LDA-SVM: Egwom [16]. The GA+VC model provides a greater accuracy, precision, as well as recall. The LDA-SVM model's precision rate is lower than that of the accuracy and recall model. The GA+VC for breast cancer classification gives higher accuracy, precision, and recall when compared with the authors' Egwon work, as shown in Table 3.

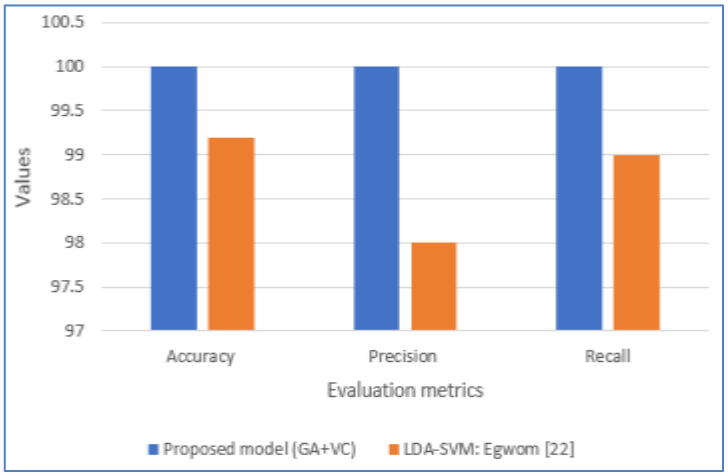

Figure 6. Comparison of existing and proposed models for classification diagnosis

Table 3. The comparison between the existing and proposed models for classification diagnosis

Classifier	Accuracy	Precision	Recall
Proposed model (GA+VC)	100.00	100.00	100.00
LDA-SVM: Egwom	99.2	98.0	99.0

4. CONCLUSION

The application of genetic algorithms for feature selection in breast cancer classification offers significant benefits and potential improvements to the overall accuracy and efficiency of the categorization process. By training the CNN on a large dataset of breast cancer samples, it learned to identify relevant patterns and characteristics indicative of cancer recurrence. This process helped reduce the input data's dimensionality and highlight the most discriminative features. After feature selection, the SCSO algorithm was applied to refine and optimize the selected features further. This paper compared before and after utilizing GA algorithms like SVM, KNN, RF, and VC models; from the experimental results, after utilizing GA algorithm that enhance the accuracy, precision, recall, and F1-score. Future research can focus on developing more advanced machine learning and deep learning algorithms to improve the accuracy and

efficiency of breast cancer classification. This could involve exploring ensemble methods, neural networks, and hybrid models that combine multiple techniques.

REFERENCES

[1] K. Karthika, S. Dhanalakshmi, S. M. Murthy, N. Mishra, S. Sasikala, and S. Murugan, "Raspberry Pi-enabled wearable sensors for personal health tracking and analysis," in *2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, IEEE, Oct. 2023, pp. 1254–1259, doi: 10.1109/ICSSAS57918.2023.10331909.

[2] A. Deepa, M. N K, R. Latha, J. Preetha, T. S. Kumar, and S. Murugan, "IoT-based wearable devices for personal safety and accident prevention systems," in *2023 Second International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, IEEE, Aug. 2023, pp. 1510–1514, doi: 10.1109/SmartTechCon57526.2023.10391691.

[3] S. Selvarasu, K. Bashkaran, K. Radhika, S. Valarmathy, and S. Murugan, "IoT-enabled medication safety: real-time temperature and storage monitoring for enhanced medication quality in hospitals," in *2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS)*, IEEE, Dec. 2023, pp. 256–261, doi: 10.1109/ICACRS58579.2023.10405212.

[4] M. S. Kumar, H. Azath, A. K. Velmurugan, K. Padmanaban, and M. Subbiah, "Prediction of Alzheimer's disease using hybrid machine learning technique," in *AIP Conference Proceedings*, 2023, p. 020091, doi: 10.1063/5.0110283.

[5] R. Raman, K. Dhivya, P. Sapra, S. Gurpur, S. P. Maniraj, and S. Murugan, "IoT-driven smart packaging for pharmaceuticals: ensuring product integrity and patient safety," in *2023 International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI)*, IEEE, Dec. 2023, pp. 1–6, doi: 10.1109/ICAIIHI57871.2023.10489420.

[6] S. Simsek, U. Kursuncu, E. Kibis, M. AnisAbdellatif, and A. Dag, "A hybrid data mining approach for identifying the temporal effects of variables associated with breast cancer survival," *Expert Systems with Applications*, vol. 139, p. 112863, Jan. 2020, doi: 10.1016/j.eswa.2019.112863.

[7] P. E. Jebarani, "Grey wolf optimization based breast cancer detection using 1D convolution LSTM classifier," *Przegląd Elektrotechniczny*, vol. 1, no. 1, pp. 3–11, Jan. 2023, doi: 10.15199/48.2023.01.01.

[8] A. Sahu, P. K. Das, and S. Meher, "High accuracy hybrid CNN classifiers for breast cancer detection using mammogram and ultrasound datasets," *Biomedical Signal Processing and Control*, vol. 80, Feb. 2023, doi: 10.1016/j.bspc.2022.104292.

[9] N. Chen, B. Han, Z. Li, and H. Wang, "Breast cancer prediction based on the CNN models," *Highlights in Science, Engineering and Technology*, vol. 34, pp. 103–109, Feb. 2023, doi: 10.54097/hset.v34i.5388.

[10] B. M. Azman, S. I. Hussain, N. A. Azmi, M. Z. A. Abd Ghani, and N. I. D. Norlen, "Prediction of distant recurrence in breast cancer using a deep neural network," *Revista Internacional de Metodos Numericos para Calculo y Diseno en Ingenieria*, vol. 38, no. 1, 2022, doi: 10.23967/j.rimni.2022.03.006.

[11] A. Begum, V. D. Kumar, J. Asghar, D. Hemalatha, and G. Arulkumaran, "A combined deep CNN: LSTM with a random forest approach for breast cancer diagnosis," *Complexity*, vol. 2022, no. 1, Jan. 2022, doi: 10.1155/2022/9299621.

[12] T. N. Mahbub, M. A. Yousuf, and M. N. Uddin, "A modified CNN and fuzzy AHP based breast cancer stage detection system," in *2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, IEEE, Feb. 2022, pp. 1–6, doi: 10.1109/ICAEEE54957.2022.9836541.

[13] C. B. Goncalves, J. R. Souza, and H. Fernandes, "CNN optimization using surrogate evolutionary algorithm for breast cancer detection using infrared images," in *2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)*, IEEE, Jul. 2022, pp. 84–89, doi: 10.1109/CBMS55023.2022.00022.

[14] D. A. Zebari, H. Haron, D. M. Sulaiman, Y. Yusoff, and M. N. M. Othman, "CNN-based deep transfer learning approach for detecting breast cancer in mammogram images," in *2022 IEEE 10th Conference on Systems, Process and Control (ICSPC)*, IEEE, Dec. 2022, pp. 256–261, doi: 10.1109/ICSPC55597.2022.10001781.

[15] S. Ruban, M. Moosa Jabeer, and R. Shenoy Basti, "Improvising breast cancer detection using CNN, VGG and SSD algorithms," in *2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT)*, IEEE, Aug. 2022, pp. 1772–1776, doi: 10.1109/ICICICT54557.2022.9918009.

[16] E. A. Mohamed, T. Gaber, O. Karam, and E. A. Rashed, "A novel CNN pooling layer for breast cancer segmentation and classification from thermograms," *PLOS ONE*, vol. 17, no. 10, Oct. 2022, doi: 10.1371/journal.pone.0276523.

[17] S. R. Sannasi Chakravarthy, N. Bharanidharan, and H. Rajaguru, "Multi-deep CNN based experimentations for early diagnosis of breast cancer," *IETE Journal of Research*, vol. 69, no. 10, pp. 7326–7341, Oct. 2023, doi: 10.1080/03772063.2022.2028584.

[18] M. A. Mezher, A. Altamimi, and R. Altamimi, "An enhanced genetic folding algorithm for prostate and breast cancer detection," *PeerJ Computer Science*, vol. 8, Jun. 2022, doi: 10.7717/peerj-cs.1015.

[19] E. Badr, S. Almotairi, M. A. Salam, and H. Ahmed, "New sequential and parallel support vector machine with grey wolf optimizer for breast cancer diagnosis," *Alexandria Engineering Journal*, vol. 61, no. 3, pp. 2520–2534, Mar. 2022, doi: 10.1016/j.aej.2021.07.024.

[20] P.-T. Yang, W.-S. Wu, C.-C. Wu, Y.-N. Shih, C.-H. Hsieh, and J.-L. Hsu, "Breast cancer recurrence prediction with ensemble methods and cost-sensitive learning," *Open Medicine*, vol. 16, no. 1, pp. 754–768, May 2021, doi: 10.1515/med-2021-0282.

[21] S. S. Ittannavar and R. H. Havaldar, "Detection of breast cancer using the infinite feature selection with genetic algorithm and deep neural network," *Distributed and Parallel Databases*, vol. 40, no. 4, pp. 675–697, Dec. 2022, doi: 10.1007/s10619-021-07355-w.

[22] O. J. Egwom, M. Hassan, J. J. Tanimu, M. Hamada, and O. M. Ogar, "An LDA–SVM machine learning model for breast cancer classification," *BioMedInformatics*, vol. 2, no. 3, pp. 345–358, Jun. 2022, doi: 10.3390/biomedinformatics2030022.

[23] H. Wang, Y. Li, S. A. Khan, and Y. Luo, "Prediction of breast cancer distant recurrence using natural language processing and knowledge-guided convolutional neural network," *Artificial Intelligence in Medicine*, vol. 110, Nov. 2020, doi: 10.1016/j.artmed.2020.101977.

[24] T. Loughman *et al.*, "Analytical validation of a novel 6-Gene signature for prediction of distant recurrence in estrogen receptor-positive, HER2-negative, early-stage breast cancer," *Clinical Chemistry*, vol. 68, no. 6, pp. 837–847, Jun. 2022, doi: 10.1093/clinchem/hvac028.

[25] K. Davoudi, and P. Thulasiraman, "Evolving convolutional neural network parameters through the genetic algorithm for the breast cancer classification problem," *Simulation*, vol. 97, no. 8, pp. 511-527, https://doi.org/10.1177/0037549721996.

[26] Z. M. Zain *et al.*, "Predicting breast cancer recurrence using principal component analysis as feature extraction: an unbiased comparative analysis," *International Journal of Advances in Intelligent Informatics*, vol. 6, no. 3, Nov. 2020, doi: 10.26555/ijain.v6i3.462.

[27] S. P. H. Shah, A. Jaffar, S. Nawaz, W. Arshad, M. Izhar, and A. Iqbal, "Early detection and classification of breast cancer using machine learning techniques," *Tianjin University Science and Technology*, vol. 55, no. 7, pp. 102–120, 2022.
[28] M. Kamrul Hasan and M. Tasnim Jawad, "Breast cancer classification using ensemble of machine learning boosting algorithms," in *2022 International Conference on Inventive Computation Technologies (ICICT)*, IEEE, Jul. 2022, pp. 444–451, doi: 10.1109/ICICT54344.2022.9850750.
[29] Y. Sahu, A. Tripathi, R. K. Gupta, P. Gautam, R. K. Pateriya, and A. Gupta, "A CNN-SVM based computer aided diagnosis of breast Cancer using histogram K-means segmentation technique," *Multimedia Tools and Applications*, vol. 82, no. 9, pp. 14055–14075, Apr. 2023, doi: 10.1007/s11042-022-13807-x.

BIOGRAPHIES OF AUTHORS

Velu Sangeetha is an Associate Professor, Department of Computer Science and Engineering, Ramaiah Institute of Technology, Bangalore is been serving in academia for 18 years. She pursued her B.E. in Computer Science and Engineering under Visveswaraya Technological University (VTU), M.Tech. in Computer Science and Engineering, R V College of Engineering, VTU. And received Ph.D. from VTU, Belagavi. Research specialization includes networks and security, artificial intelligence and machine learning, blockchain. She has published over 45 research papers in reputed National and International Journals and Conferences. Some of papers published by her enjoy good citation in Google Scholar, Scopus, and Web of Science. She has also been the member of various professional societies including IEEE (IEEE Senior Member, IEEE SIGHT Bangalore Section 2024-Vice Chair), IEI, ISTE (Life Member), IAENG (Member). She can be contacted at email: drsangeethav@msrit.edu.

Maniyambadi Vaneeta is an Associate Professor, Ramaiah Institute of Technology is highly skilled, and qualified Associate Professor serving academia from 21 years at various institutes. She received her B.E (CSE) in 1999, M.E (CSE) in 2009, and Ph.D. (Data Security) in 2021 from Visveswariah Technological University (VTU), Belagavi. Research specialization includes networks and security, artificial intelligence and machine learning, blockchain. Published over 40 research papers in reputed National and International Conferences and Journals. Some of papers published by her enjoy good citation in Google Scholar, Scopus, and Web of Science. She has received fund from KSCST in the year 2021. She can be contacted at email: vaneetam@msrit.edu.

Arjuna Mamatha is working as Assistant Professor in Department of Computer Science and Engineering, Ramaiah Institute of Technology. She received her M. Tech degree in CSE from Visvesvaraya Technological University, Karnataka and B.E degree in ISE from Visvesvaraya Technological University, Karnataka. Currently pursuing Ph.D. in Visvesvaraya Technological University, Karnataka. Having a total of 1.5 decade of teaching experience. She has authored and coauthored 7 research papers in various international conferences. She has published 2 book chapters and 2 patents. She is a member of IEEE, ISTE and IAENG. Her areas of interest include image processing, network security, artificial intelligence and machine learning. She can be contacted at email: mamatha.a@msrit.edu.

Muniya Shoba Associate Professor, Department of ISE, BMS Institute of Technology and Management College is extremely passionate towards Teaching Learning Process and R&D activities. With a total teaching experience of 18 years and 6 years in research. She pursued her B.Edegree in Information Science and Engineering from JSSATE, Visveswaraiah Technological University, Belagavi, M.Tech degree in Computer Science and Engineering from RVCE, Visveswaraiah Technological University, Belagavi and Doctoral of Philosophy in School of Computing and IT from REVA University. She is a recognized Research Supervisor under VTU and 1 Research Scholar is awarded with Ph.D. under her research supervision.She has authored more than 15+ research papers in various renowned Scopus-indexed/WoSJournalsand International conferences. She has published 1 patent. She has received a National Award for Faculty in Higher Education (Karnataka, South Zone) 2020 and Research Excellence Award in 2020. She has authored a textbook on "Big data analytics using python". She has delivered many technical sessions as resource person, session chair in International Conferences, Panel member and Jury member for Project exhibitions in various engineering colleges across Karnataka. She is an active member of IEEE-ITS Chapter, ISTE, IEI and InSc professional bodies. She can be contacted at email: dr.smc@bmsit.in.

Sugatur Ramamurthy Deepa is currently working as Professor and Head of the Computer Science and Design Department at K. S. Institute of Technology, Bengaluru. Received Ph.D. from VIT, Chennai from the School of Computer Science and Engineering in the year 2020, M. Tech Degree in Computer Science and Engineering from VTU in 2009, and BE Degree in Electronics and Communication from VTU in 2002. Research interests are wireless sensor networks, the internet of things, soft computing, and computational intelligence. Published 18 papers in International Journals, Presented 5 papers at international conferences, and 8 papers at National Conferences. Obtained 3 Indian and 1 Australian patent. Reviewerfor the Peer-reviewed Scopus-indexed Journal-Wireless Networks, Springer. Obtained a grant of Rs. 1,00,000 from AICTE under the Scheme for promoting interests, creativity, and ethics among students (SPICES). Received Young Achiever Award by IEAE in the year 2019 and the Longest Continuous SBC award by CSI in the year 2018. Resource person for various technical talks and workshops. Organized several workshops, technical talks, guest lectures, workshops, faculty development programs, and CSI Karnataka state student conventions. She can be contacted at email: deepaprashanth.g@gmail.com.

Velusamy Sujatha received her doctoral degree and post-graduation from Anna University, Chennai, and her undergraduate degree from Madras University. Her research focuses on designing charge pump Phase-Locked Loops, with a keen interest in Analog VLSI. She began her career as a lecturer in 1993 and has served as an academician in various reputed institutions, holding several positions and guiding numerous UG and PG projects. Currently, she is a Professor in the Department of Electronics and Communication Engineering (ECE) at Shree Sathyam College of Engineering and Technology, Sankari, Tamil Nadu. She can be contacted at email: info@shreesathyam.edu.in.

Shanmugam Sujatha is an adjunct professor, Saveetha School of Engineering,Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India. She published her research articles in many international and national conferences and journals. Her research areas include network security and machine learning. She can be contacted at email: sujathasmvr@gmail.com.