International Journal of Operations Research and Information Systems
Volume 13 • Issue 2

Artificial Intelligence-Based Breast Cancer Detection Using WPSO

Murali Krishna Doma, Shri Vishnu Engineering College for Women, India*
Kayal Padmanandam, BVRIT Hyderabad College of Engineering for Women, India
Sunil Tambvekar, Nowrosjee Wadia Maternity Hospital, India
Keshav Kumar K., Jawaharlal Nehru Technological University, India
https://orcid.org/0000-0002-9211-2960
Bilal Abdualgalil, Mahatma Gandhi University, India
R. N. Thakur, LBEF Campus, Nepal
https://orcid.org/0000-0003-3911-4358

ABSTRACT

To detect breast cancer in the early stages, microcalcifications are considered a key symptom. Several scientific investigations were performed to fight against this disease for which machine learning techniques can be extensively used. Particle swarm optimization (PSO) is recognized as one among several efficient and promising approach for diagnosing breast cancer by assisting medical experts for timely and apt treatment. This paper uses weighted particle swarm optimization (WPSO) approach for extracting textural features from the segmented mammogram image for classifying microcalcifications as normal, benign, or malignant, thereby improving the accuracy. In the breast region, tumor part is extracted using optimization methods. Here, artificial intelligence (AI) is proposed for detecting breast cancer, which reduces the manual overheads. AI framework is constructed for extracting features efficiently. This designed model detects the cancer regions in mammogram (MG) images and rapidly classifies those regions as normal or abnormal. This model uses MG images obtained from hospitals.

KEYWORDS

Artificial Intelligence Tools, Breast Cancer, Convolutional Neural Networks, Data Analytics for Disease Prevention, Data Investigation, Mammogram, Microcalcifications

INTRODUCTION

Breast cancer is the most commonly found in women which causes deaths who are aged from 20 to 59. According to the Ministry of Health and Medical Education, it has become the most common disease in recent years in Iran (Ganggayah-Taib, et al., 2019). Today, 88% of women diagnosed with breast cancer have a life expectancy of 10 years. In the United States, it has been reported that about 12% of women were identified during their lifetime, and were referred to as the second cause of women's

DOI: 10.4018/IJORIS.306195 *Corresponding Author

death (Houssein-Emam, et al., 2021). Diagnosing the disease at the earlier stages is important because in the early stages, cancer masses are restricted to the breast and the chance of surgical treatment in a less invasive manner is increased. The mortality rate is also decreased in the early stage (Beura-Majhi, et al., 2015; Arulkumar, Lakshmi, and Rao, 2021). Also, the use of classifiers such as artificial neural networks in various fields of engineering sciences is increasing to analyze the time series and various issues of classification. Due to the invention of techniques in the recent era for early diagnosis of breast cancer, the survival rate of the patients is improved. Nowadays, X-ray mammography and MRI (Magnetic Resonant Imaging) techniques are widely utilized with few implications and limitations. X-ray is very harm due to the ionizing radiation and thus its contact with patients has to be only for very short duration. Conversely, MRI technique is expensive while mammography is of less cost, but hard to provide consistency and accuracy in analysing breast cancer. Moreover, errors occur while analysis (Hamian-Darvishan, et al., 2018).

To increase the rate of accuracy and reduce the occurrence of errors, supervised machine learning approaches like KNN, SVM, LSSVM are developed. These models efficiently classify the features as normal or abnormal classes. These methods are complex and even tedious with low CR. Therefore, to provide a solution for all the drawbacks of breast cancer, an optimal classification model is required for which machine learning approaches based on image processing are developed to classify cancer and non-cancer images which involved mammogram images. As the features are essential to discriminate breast cancer as benign or malignant, feature extraction process is of most important. Once the features are extracted, properties of the image like depth, coarseness, smoothness, and regularity are obtained with the help of segmentation process (Leng-Li, et al., 2018). Scientifically, with breast cancer, division of tumor cells is uncontrolled and abnormal tumor cells need more nutrients for growing continuously and to reproduce (Gupta, 2021). Image analysis is the process of producing images on a computer with the goal of determining what things are visible in the image. The process of segmenting a picture into its constituent parts is known as image segmentation. It is one of the most important tasks in autonomous image analysis since the outcomes of segmentation will influence all subsequent tasks, such as feature extraction and object classification. Because of its importance, the segmentation procedure and approach have received a lot of attention in recent decades. This has resulted in a large number of (thousands) of distinct algorithms, and the number is continually growing.

The cancer cells penetrate into the surrounding for gaining nutrients. There is a heterogeneous variation in the circulation of blood with various tumors and hence lesion morphology characteristics and ambiguity with edges in diagnosing images are significant indicators for evaluation. The paramagnetic contrast agent spreads in blood which enters into the blood vessel and passes in the intercellular space as well as cells easily via penetrable capillary wall, hence the sputum concentration is high in the tumor rich region (Obulesu-Kallam, et al., 2021). This abnormality can be found using TIC when DCE-MRI is utilized for several imaging of the same tissue in various stages. Thus, edge, shape etc which are static characteristics and initial increase and change in signal which are dynamic characteristics of the lesion plays a major role in identifying the tumor as benign or malignant. MRI images are usually clear and complete with multi-angle, multi-faceted imaging. With the breast, surface coil has been used for clinical purpose, and MRI technology are improved to be much clear. However, the true positive rate and the true negative rate obtained while diagnosing breast cancer are also improved simultaneously (Pi-Chen, et al., 2020). The remaining part of this work is presented as follows. An outline of relating works is discussed in Section 2; Section 3 elaborates the proposed methodology while Section 4describes the experiments and the discusses the obtained results. Finally, Section 5concludes the work with future improvements (Libson and Lippman, 2014).

The most frequent malignancy among women is breast cancer. Over half a million individuals die each year as a result of this disease. While early detection is critical for successful treatment, breast cancer diagnosis is challenging due to the dense breast tissues. The doctors sought a technique to increase the accuracy of the diagnosis because it was prone to human error. According to a new study, AI algorithms trained on a dataset of a large number of breast pictures can aid in the detection of evidence of cancer with greater diagnostic accuracy. Breast cancer claims the lives of 40,000 women in the United States each year. Cancers can often be healed if discovered early. Although mammograms are the most accurate diagnostic available, they are nonetheless flawed and frequently produce false positive results, leading to unneeded biopsies and procedures. So-called "high-risk" lesions that appear worrisome on mammograms and exhibit abnormal cells when analysed by needle biopsy are a common source of false positives. In this instance, the patient usually has surgery to remove the lesion; nevertheless, 90% of the time, the lesions turn out to be benign during surgery. This implies that thousands of women are subjected to painful, costly, and scar-inducing operations that aren't even necessary. So, how can unneeded procedures be minimised while keeping mammography's critical function in cancer detection? Artificial intelligence, according to researchers at MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL), Massachusetts General Hospital, and Harvard Medical School, is the answer (AI).

Breast cancer is the most frequent disease in women, and despite significant advances in treatment, it remains a leading cause of cancer-related death, accounting for around 500 000 deaths each year globally. Mammography-based population-based breast cancer screening programmes are thought to be beneficial in lowering breast cancer-related mortality. Current screening programmes, on the other hand, are labor-intensive due to the enormous number of women examined each diagnosed cancer and the use of double reading, notably in European screening programmes, which adds to the expense. Furthermore, despite this technique, up to twenty five percent of mammographically evident tumours go undetected during screening. Alternative techniques to allow the continuation of present screening programmes are required, given the rising scarcity of radiologists in several countries, notably breast screening radiologists. Furthermore, it is critical to avoid apparent lesions in digital mammography (DM) from being ignored or misconstrued.

Artificial intelligence (AI) systems that evaluate digital mammography (DM) at radiologist-like levels might increase breast cancer screening accuracy and efficiency. We wanted to see how well an AI system could diagnose breast cancer in DM without the help of radiologists. Breast cancer mortality has been reduced as a result of population-wide breast cancer screening programmes, and screening adherence is high. However, two issues remain unresolved: the vast number of radiologist hours spent analysing mostly healthy women, and the relatively high fraction of women whose cancer is not found during screening despite regular involvement. In biennial screening programmes, about five women will have their disease identified, and two women will have a normal screening evaluation and be clinically diagnosed with breast cancer in the interval before the next scheduled screening (interval cancer). To add to the early detection potential, approximately twenty percent of screen-caught tumours are big (>2 cm), suggesting that many of these cancers could have been detected on previous screening if MRI had been utilised. MRI, on the other hand, is expensive, time-consuming, and impractical to utilise as a first-line screening method for all women. A triaging model is needed to identify the mammograms for which radiologist examination is unnecessary, as well as the women who are most at risk of leaving screening facilities with cancer that has not been discovered. For mammography, several artificial intelligence (AI) cancer detection software algorithms have been created. Even while validation in a fully representative screening cohort is still missing, several software algorithms are currently performing at a level comparable to radiologists when assessing mammograms. There are a number of potential functions for AI in the screening process that has yet to be completely explored. While using an AI cancer detector as a contemporaneous assistance to a radiologist may uncover more malignancies, it will not achieve the goal of lowering radiologist resources.

Computer-aided detection methods have been used to detect and classify breast lesions in mammography since the 1990s. The widespread use of DM for breast cancer imaging has accelerated the development of automated breast cancer detection systems. Unfortunately, due to a low specificity, no studies have proven that classical computer-aided detection systems directly improve screening performance or cost-effectiveness. As a result, they can't be used as a stand-alone reader for mammography screening. However, because to the development of novel algorithms based on deep learning convolutional neural networks, the field of artificial intelligence (AI) is rapidly changing. Self-driving automobiles and sophisticated voice recognition are two examples of these methods' success in automating cognitively challenging jobs. Deep learning-based AI is also rapidly reducing the gap between humans and computers in medical imaging. Such algorithms, it has been proposed, may have the ability to increase the benefit-to-harm ratio of breast cancer screening programmes. Several deep learning-based algorithms for automated mammography analysis have been developed in recent years, with some of them already showing encouraging results when compared to radiologists in constrained and homogeneous circumstances.

The teams worked to develop an AI system that employs machine learning to predict whether a high-risk lesion discovered on needle biopsy after a mammogram will upgrade to cancer during surgery as a first attempt to apply AI to better detection and diagnosis. When compared to previous methodologies, the model correctly recognised 97 percent of high-risk lesions as malignant and reduced the frequency of benign procedures by more than 30 percent when evaluated on 335 high-risk lesions. "Because diagnostic technologies are so inexact, doctors have an understandable inclination to over-screen for breast cancer," explains Regina Barzilay, a breast cancer survivor and Delta Electronics Professor of Electrical Engineering and Computer Science at MIT. The programme, which was trained on data from over 600 existing high-risk lesions, looks for trends in a variety of data factors such as demographics, family history, previous biopsies, and pathology reports. "To our knowledge, this is the first study to apply machine learning to the task of distinguishing high-risk lesions that require surgery from those that do not," says collaborator Constance Lehman, professor at Harvard Medical School and chief of MGH's Department of Radiology's Breast Imaging Division. "We believe that by doing so, we will be able to help women make more informed treatment decisions and deliver more targeted health care in general."

RELATED WORKS

This section discusses few related works carried out for diagnosing breast cancer which involved various optimization techniques. It is well known that breast cancer is one serious and dangerous cancers among women and hence diagnosing at the earlier stages is more effective to provide treatment and protect the lives of patients. Till now, several approaches are coined for detecting breast cancer which addresses different sorts of challenges and few of them are reviewed here. Asri et al., in 2016 (Asri-Mousannif, 2016), employed machine learning methods for the prediction and classification of WBC actual dataset. Various classifiers were used which includes SVM, Naïve Bayes, KNN and decision tree C4. SVM produced accuracy with Weka tool. In (Acharjya, 2016; Varatharajan and Varatharajan, 2020; Sampathkumar-Rastogi, et al., 2020), Chowdhary et al. utilized mammography images for the detection of breast cancer using intuitive fuzzy histogram magnification approach thereby data was processed and image quality was improved (Kelly-Dean, et al., 2010):

Then, probabilistic Fuzzy Clustering approach was employed for segmenting and separating the cancer tissues. Hence, this model was suitable for processing larger cancer datasets with the objective to offer better accuracy. Next, with the methods like grey area coefficient and linear binary pattern, textural properties were extracted. The accuracy obtained was 94% but hard while dealing with larger datasets and extends the processing time. Aalaei et al. (Aalaei-Shahraki, et al., 2016; Juneja, 2021) employed genetic meta-specificity reduction for classifying breast cancer. Three datasets namely WBC, WDBC and WPBC were used for evaluating which used Artificial Neural Network (ANN) cluster. (Rudra and Kautish, 2021)

The accuracy estimated for the method used with WBC, WDBC and WPBC datasets were 96, 96.1 and 76.3, respectively. Even though feature set was reduced, accuracy could be improved. Nilashi et al., in 2017 (Nilashi-Ibrahim, et al., 2017), designed a knowledge-based system which involved fuzzy logic. The process was carried out in three steps: initially Wisconsin Breast Cancer data was processed. Then, data with similar groups was clustered by the use of Expectation Maximization (EM) clustering technique. Finally, once the features were reduced by PCA, fuzzy rule set was categorized as data by means of regression tree (Reyana and Kautish, 2021):

The accuracy obtained was 93.2%. Sometimes when learning rules are applied on datasets, the classification task is complicated. In (Arulkumar and Vivekanandan, 2015; Jeyasingh and Veluchamy, 2017), the use of Bat algorithm selected optimal features for diagnosing breast cancer. 286 samples were selected from WDBC dataset for which simple random sampling approach was involved for feature selection. After selecting the features, according to the classification similarity which involved Random Forest (RF), overall ranking was performed and obtained accuracy. As samples are selected at random, selection of features was sometimes difficult. In (Doreswamy, 2016), Dore swamy et al. improved Bat algorithm to classify breast cancer images. 569 samples of UCI data were involved in experimenting this method. The accuracy for training set was 92.61 while that of the testing was 89.95. In (Muslim-Rukmana, 2018), an approach using PSO was utilized for reducing the specificity in diagnosing breast cancer. (CHU-LEE, et al., 2021)

The objective was to estimate the level of breast cancer. 699 pre-processed samples of UCI data after reducing the specificity were used by PSO algorithm along with decision tree C4.5 to classify the samples into two classes namely malignant and benign. The accuracy achieved was 95.61%. Sahu et al. (Sahu, Mohanty and Rout, 2018) used a hybrid approach for classifying and diagnosing breast cancer. With PCA feature reduction and various clusters, it was found that ANN classification produced 97% of performance than other clusters. 699 samples with 9 features were used in the experiment to label them as benign and malignant. Even though results achieved are better, every method has few weaknesses and limitations. In (Gao-Wu, et al., 2018), Gao et al. integrated shallow AI with deep AI and formulated SD-AI. Shallow AI was used with the intention to extract "virtual" recombination of images which has lower energy, while deep AI extracted the novel features related to LE (Reyana-Krishnaprasath, et al., 2020).

Additionally, knowledge of nonlinear mapping was gathered from LE for recombining images; shallow AI and deep-AI was created with 49 CEDM each. The performance was enhanced in terms of AUC and was accurate than the methods existing. In (Ting, Tan and Sim, 2019), Ting et al. developed AI-BCC which helped medical experts in diagnosing breast cancer at the earlier stage. This model enhanced the classification using AI and the breast cancer images were classified as various types like benign, malignant, and healthy. From the numerous experiments conducted, this model enhances accuracy, sensitivity as well as AUC. Diagnosing and classifying approaches involved for breast cancer were not tested and evaluated with three various datasets of breast cancer. The specialties of the present investigation are reduction is detecting costs, using better classifier with no adverse effects of aggressive approaches, higher accuracy of detection than the paper cited, choosing titles appropriate with the data available and comprehensive comparison with the researches made so far (Bhattad and Jain, 2020).

We looked at the impact of using artificial intelligence (AI) cancer detection software to triage certain screening examinations into a no-radiologist work stream, and then triaging certain screening examinations into an enhanced assessment work stream after regular radiologist assessment of the rest. The goal of improved assessment was to imitate women being selected for more sensitive screening, allowing for early detection of malignancies that would otherwise be identified as interval cancers or next-round screen-detected tumours. The study's goal was to see if AI could minimise radiologist effort while while increasing cancer diagnosis.

PROPOSED METHODOLOGY

The workflow of the methodology developed is illustrated in Figure 1 below. The phases like pre-processing, segmentation and feature extraction are discussed below. AI classifier is involved to obtain the accuracy of classification.

Preprocessing Steps

Step 1: Looking for an input Breast Image.
Step 2: The raw image provided as input raw undergoes resizing to 256 x 256.
Step 3: When 3-dimensional (3D) images are provided as input, they are converted to 2D, since mostly image processing is carried out only with 2D images i.e, RGB image is converted into gray scale image.
Step 4: Two filtering techniques are applied for de-noising as described below:
Step 4.1: Out_1 = Laplacian filter is applied on the gray scale image.
Step 4.2: Out_2 = Then mean filter is applied on the gray scale image.
Step 4.3: Out_3 = Out_1 – Out_3.
Step 4.4: The final output of the pre-processing stage is the pre-processed breast imageOut_3.
Step 5: Out_3

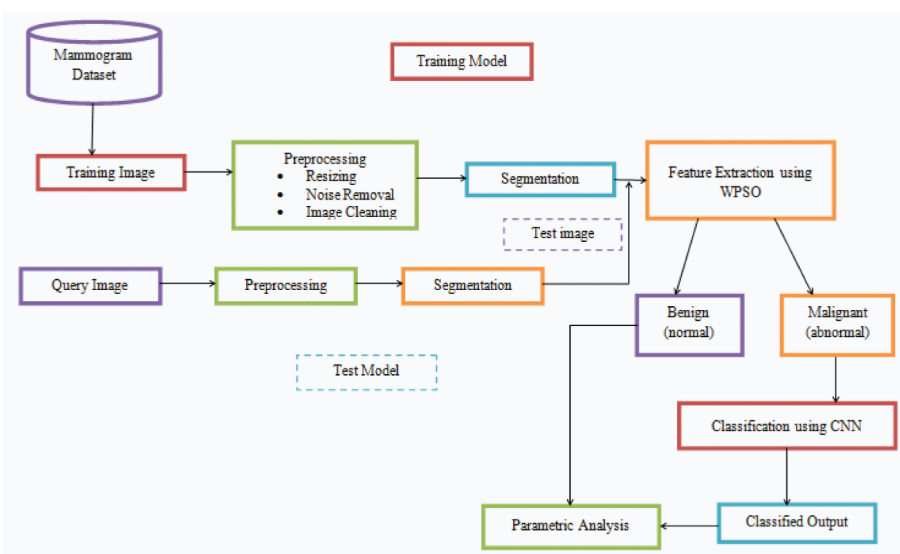
Segmentation Steps

Input: Out_3 Pre-processed Image
Step 1: Gradient is obtained along X and Y axis in variables Out X and Out Y.
Step 2: Gradient values are combined to obtain gradient vector G val which is given by:

G val=[1/(1+(Out X + Out Y))]

Step 3: G val obtained in radians is converted to degrees so that orientation information of image pixels can be attained.

Figure 1. Architecture of the proposed technique



Step 4: Out_3 image is partitioned to grids GRi.
Step 5: Threshold values are defined for intensity Ti and orientation To.
Step 6: For every grid GRi do:
- **Step 6.1:** Histogram Hi for every pixel Pjis computed over grid GRi.
- **Step 6.2:** Most frequent histogram of grid GRiis found which is represented by FreqH.
- **Step 6.3:** Any arbitrary pixel Pjis selected which is related to FreqH which is then assigned to pixel information seed point (SP) with Intensity Ip and Orientation Op.
- **Step 6.4:** Intensity along with orientation constraints for adjacent pixel is verified.
- **Step 6.5:** When both constraints are fulfilled, then decided that region is grown, or else next grid is considered for further process.

Step 7: Output: Segmented Image

Weighted PSO Based Feature Extraction

A heuristic global optimization technique named Weighted Particle Swarm Optimization (WPSO) algorithm simulates the social behaviour of flocking bird towards a position for attaining the exact objective in a multidimensional space. This approach involves a population of particles (also called swarm) in the search space. For every particle, the status is categorized based on its location $\vec{x}_i = \left\{x_{i1}, x_{i2}, \ldots . x_{id}\right\}$ and the velocity of particle i is given by $\vec{v}_i = \left\{v_{i1}, v_{i2}, \ldots . v_{id}\right\}$. To find the optimal solution, every particle deviates from its actual searching direction to a new direction based on two concepts namely the best location of the given particle (pbest) and the one obtained so far by swarm (gbest). WPSO identifies the optimal solution after velocity and position of every particle are updated in relation with the equations:

$$v_{id}^{(t+1)} = wv_{id}^{t} + c_1 r_1 \left(p_{id} - x_{id}^{t}\right) + c_2 r_2 \left(p_{gd} - x_{id}^{t}\right) \tag{1}$$

$$x_{id}^{(t+1)} = x_{id}^{t} + v_{id}^{(t+1)} \tag{2}$$

where t symbolize the iteration in the evolutionary space and d denote the dimension in the search space respectively. w represent the weight of inertia. c1 andc2 represent the personal and social learning factors. r1 and r2 are uniformly distributed random values ranging between 0 to 1. pid and pgd denotes pbest and gbest in the dimension d.

The basic steps performed in W PSO algorithm are as described below:

1. **Initialization:** Random positions and velocities are used to initialize the particles.
2. **Evaluation:** For every particle, value of the objective function is estimated.
3. **Finding pbest:** When the value obtained with the objective function is better than the pbest for particle i, then the current value is assigned as the new pbest.
4. **Finding gbest:** When pbest is better than gbest, then gbest is assigned to the current value.
5. **Updating the position and velocity:** For every particle, velocity is updated using Equation 1, and the particle is moved to the next position based on Equation 2.
6. **Terminating Criteria:** When the required number of iterations are reached, the process ends or else repeated from step 2.

For the search space, exploration and exploitation are controlled by weight as velocity is adjusted dynamically. Moreover, weight controls the impact of the previous velocities on the current one. Thus, the exploration capabilities are compromised between global and local swarm. Larger weight simplifies global search for new areas while the smaller weight simplifies local search. When the weight is chosen properly, global and local exploration of swarm is balanced providing better solution. Hence, weight can basically set to a larger value for better global exploration of the search space and then decrease it gradually to obtain refined solution. When the weight decreases linearly, exploration from global to local change linearly. Search algorithms are required to have non-linear searching ability. With few statistical features obtained, PSO search is easily understood and the suitable weight is calculated for the next iteration. Here, when there is an increase in total generation, there is a linear decrease in weight w while optimization in relation to:

$$W = w_{max} - \left(\frac{w_{max} - w_{min}}{iter_{max}} \right) * iter \tag{3}$$

where w_{max} and w_{min} denote the maximum and minimum inertia weight respectively, $iter$ and $iter_{max}$ are the current iteration and maximum number of iterations respectively. For particle i, the best position is position that the particle visited (past value of Xi), which provides highest fitness value. For minimization, a position with small function value is considered to have fitness. f(X) denotes the minimized objective function for which the updated equation is:

$$P_{bestid}^{(t+1)} = \{x_{id} t \, if \, f(x_{id}\left(t+1\right) \geq P_{bestid}^{t} \tag{4a}$$

or:

$$P_{bestid}^{(t+1)} = \{x_{id}\left(t+1\right) if \, f(x_{id}\left(t+1\right) \geq P_{bestid}^{t} \tag{4b}$$

A faster rate convergence is provided by the gbest with the expense of robustness and only a single best solution is maintained termed as global best particle. The role of this particle is to act as an attractor and hence pulls every particle towards it. Ultimately, every particle converges at this position and thus has to be regularly updated if not swarm converges prematurely. For every particle in the swarm, fitness value is computed using the objective function. Then, Pid and Pgd values are evaluated and updated with the global best position or better particle best position if obtained.

Steps for WPSO:

1. Initialize the function
2. Create objective function
3. Objective function is based on intensity of pixel
4. Set iteration count = 1000
5. Calculate pixel intensity of images
6. Optimize the cancer image pixel intensity
7. Calculate optimal value for input image pixel
8. Extract tumor part with maximum pixel intensity
9. Calculate accuracy

Classification Using Convolution Neural Networks

AI takes the breast cancer image dataset an input for classification. Then, deep convolutional kernels are trained using the introduced AI architecture. RELU nonlinearity is used in convolution layers and are defined as:

$$f\left(x\right)=\begin{cases} x, if\, x>0 \\ ax, otherwise \end{cases} \tag{5}$$

Generally, convolution layer is stated as:

$$y^j = fb^j + \sum_i k^{ij} + x^i \tag{6}$$

Here, xi represents the i^{th}input map and yi denotes the j^{th}output map. bjis the bias parameter of j^{th}map, convolution process between two functions is given by *, and convolutional kernel involved between i and j maps is bij. Max-pooling layer was the next layer following the convolutional layer. In max-pooling layer, every neuron provides yi pools in the output map yi pools against s * s non-overlapping areas of xi. In general, max-pooling layer is defined as:

$$y_j^i = max_{0\le m\le s}\left\{x_{j.s+m}^i\right\} \tag{7}$$

Convolutional as well as max-pooling layers are fully connected which is followed by Softmax classifier containing output classes which equals the number of outputs. In the architecture introduced, tanh is used as a non-linear protocol in connecting one layer with another. The function of Softmax function equals squashing, and dataset with k-dimension is re-normalized producing real values ranging from 1 to 2. This is represented mathematically as:

$$\sigma\left(z\right)_j = \frac{e^{zj}}{\sum_{k=1}^{k} e^{zk}}, for\, j = 1,\ldots k \tag{8}$$

Error obtained while developing ML approaches are training and generalization errors. The former is observed while training the neural network, where the latter is produced while testing the proposed classifier. In the process of deep learning, training is frequently affected with the process of overfit and under-fit. To surpass these issues, after every layer, batch normalization is applied in the proposed architecture for BCC. Dropout layer was added next to the first fully connected layer. The entire architecture developed for breast cancer classification is illustrated in Figure 2.

Training of AI

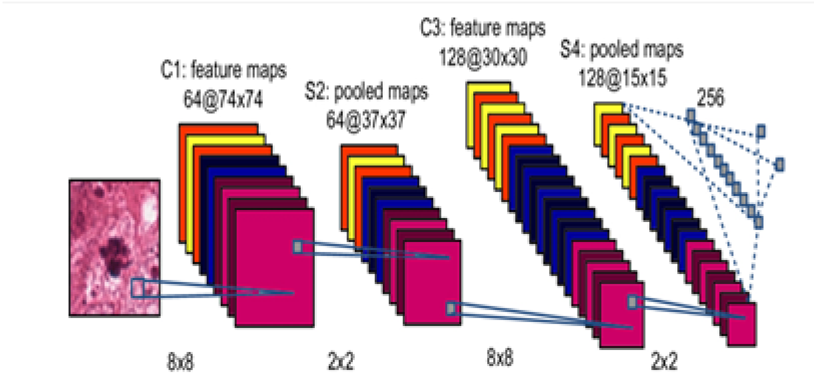
The proposed AI architecture has two classes namely benign and malignant. Weighted loss function was employed for training the proposed AI classifier:

$$\xi\left(w, x_n, y_n\right) = -\frac{1}{N}\sum_{n=1}^{N}\alpha_n \sum_{k=1}^{k} t_{kn} ln y_{kn} \tag{9}$$

Here xn represents input vector, yn the prediction obtained from classifier for n^{th}clinical input, and tn its actual response. K and N are the number of classes total clinical samples.

For recognizing, patch results of the entire image are combined. As the model is trained with image patches, strategy is necessary for partitioning the actual testing images into patches, then executing and combining the results obtained to get optimal result but is computationally too complex. Rather,

Figure 2. Proposed AI Architecture for Breast Cancer Classification



grid patches are obtained from the images which provides the set of non-overlapping patches, and this was reasonable and balanced the performance of classification as well as computational cost. By implementing this model, every patch produced the probability of every possible class for the given patch of the image. For combining the results produced by the patches for the test image, three various fusion rules were involved and found that Sum rule produced better results.

PERFORMANCE ANALYSIS

Detection of breast cancer in the earlier stages is critical for treatment and managing the its condition. This study presented a detailed derivation methods and processes along with way it is applied in detecting tumor. According to the tissue segmenting method, the effects of the obtained number of glandular tissues is analyzed. It is observed that presence of numerous glandular tissues worsen the imaging effect later. Simultaneously, a progressive approach to detect multiple tumors is also introduced. Imaging is done in three steps: preliminary examination, refocusing, and image optimization by which every tumor is detected successfully. Here, WPSO-AI is used for extraction of features and classification of tumor and this has obtained enhanced accuracy. Features were obtained and classified the image of histopathology. The below Figure 3 shows extracted feature of histopathology image using WPSO-AI. The Figure 4 shows classified image with malignancy.

Figure 3. Feature extracted tumor image of histopathological image

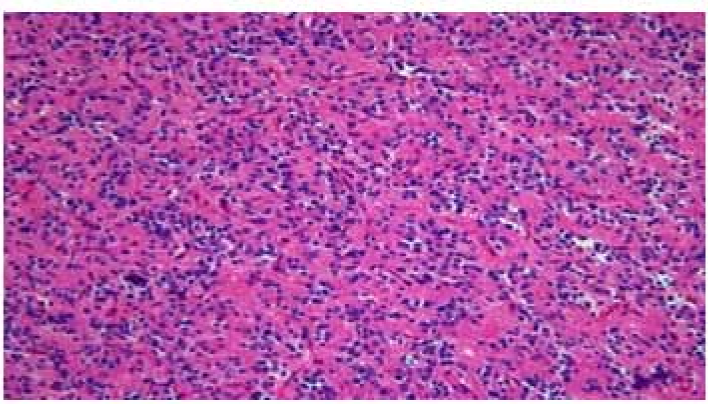


Figure 4. Classified image of histopathology detecting malignancy

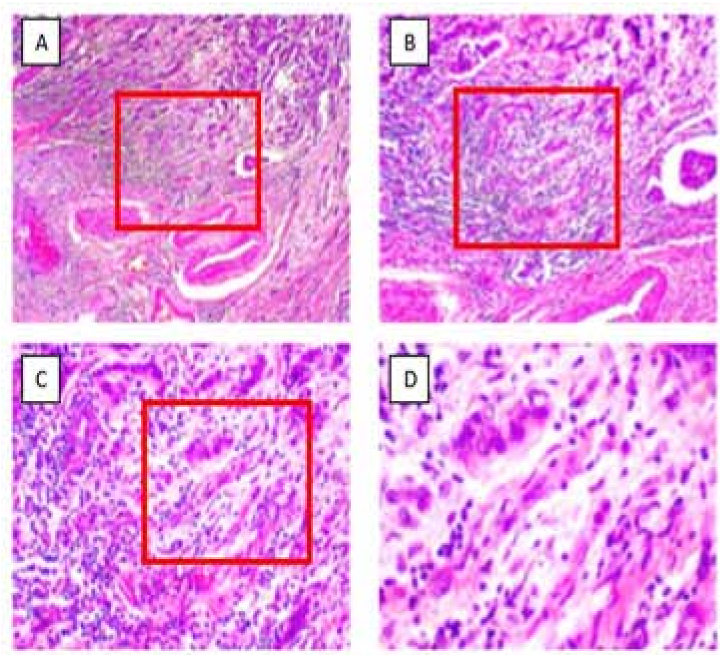


Table 1. Comparison of Accuracy

Number of Epochs	AI-BCC	WBPC-ANN	WPSO-AI
25	45	49	53
85	52	56	59
125	64	66	66
165	72	76	76
200	81	85	89

The accuracy, precision, recall and F-1 score graph has been shown below in Figure 5,6,7,8. The below graphs show comparison of parameters between existing and proposed techniques.

CONCLUSION

The objective to carry out this research is to improve accuracy of detection using CAD technique for detecting breast cancer. With this objective, a framework was contributed along with its flow and parameters used for simulation. Publicly available dataset is involved for analyzing the effectiveness of the method for classifying normal and abnormal breast images of several individuals. Here, weighted particle swarm optimization (WPSO) with AI is employed named as WPSO-AI. The objective of the method is to diagnose the breast cancer using kernel density estimation based classifier by extracting

Figure 5. Comparison of Accuracy

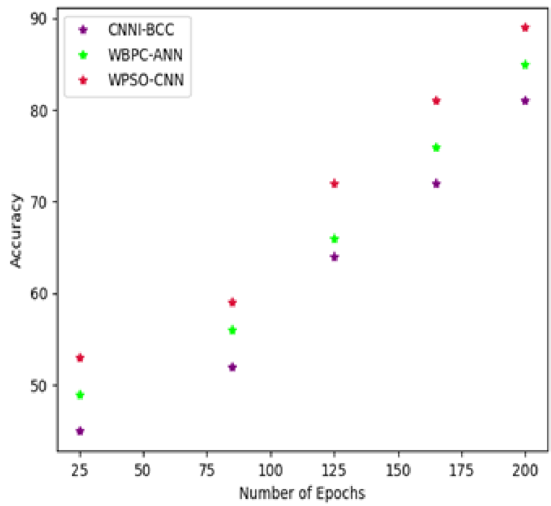


Table 2. Comparison of Precision

Number of Epochs	AI-BCC	WBPC-ANN	WPSO-AI
25	62	69	71
85	65.7	71.34	73.65
125	68.5	74.63	81.26
165	70.83	78.49	88
200	73.6	81.1	86.3

Figure 6. Comparison of Precision

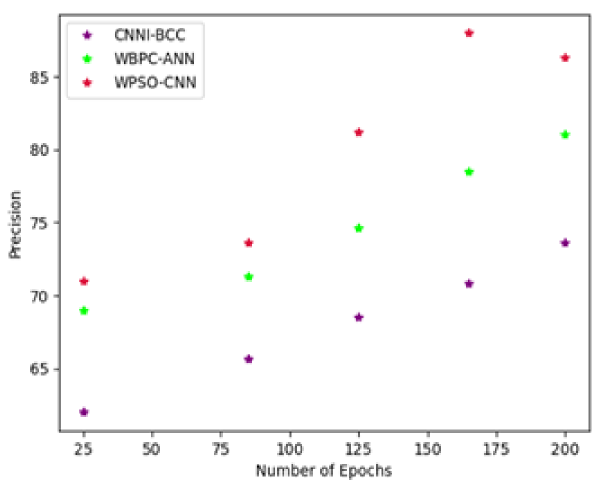


Table 3. Comparison of Recall

Number of Epochs	AI-BCC	WBPC-ANN	WPSO-AI
25	74	74	79
85	77.26	75	80
125	82	82	84.6
165	83	88	89
200	85	89	90.6

Figure 7. Comparison of Recall

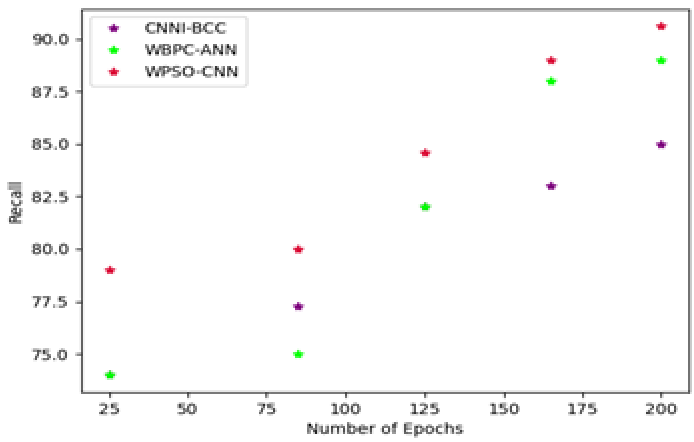
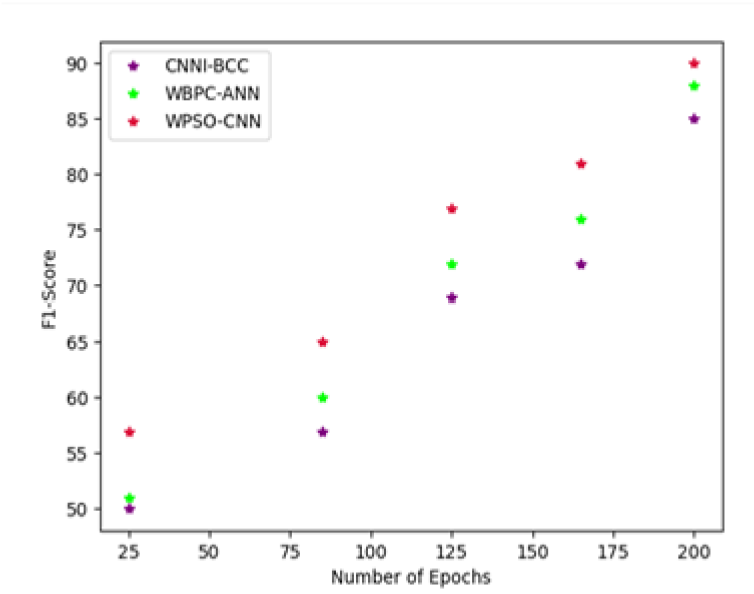


Table 4. Comparison of F-1 Score

Number of Epochs	AI-BCC	WBPC-ANN	WPSO-AI
25	50	51	57
85	57	60	65
125	69	72	77
165	72	76	81
200	85	88	90

the features and estimate the error between the estimated and true density. From the results it is observed that the performance of WPSO-AI is remarkable than existing approaches. The future work is to possibly develop an online breast cancer detection system since the detecting systems used currently are offline.

Figure 8. Comparison of F-1 Score



REFERENCES

A, R., & Kautish, S. (2021). Corona virus-related Disease Pandemic: A Review on Machine Learning Approaches and Treatment Trials on Diagnosed Population for Future Clinical Decision Support. *Current Medical Imaging: Current Medical Imaging Reviews, 17*. 10.2174/1573405617666210414101941

Aalaei, S., Shahraki, H., Rowhanimanesh, A., & Eslami, S. (2016). Feature selection using genetic algorithm for breast cancer diagnosis: Experiment on three different datasets. *Iranian Journal of Basic Medical Sciences.*, *19*(5), 476–482. PMID:27403253

Acharjya, D. P. (2016, January 1). *Breast Cancer Detection using Intuitionistic Fuzzy Histogram Hyperbolization and Possibilitic Fuzzy c-mean Clustering algorithms with texture feature based Classification on Mammography Images*. VIT University Research Platform. https://research.vit.ac.in/publication/breast-cancer-detection-using-intuitionistic

Arulkumar, C. V., & Vivekanandan, P. (2015). Multi-feature based automatic face identification on kernel eigen spaces (KES) under unstable lighting conditions. *2015 International Conference on Advanced Computing and Communication Systems*, 1-5. doi:10.1109/ICACCS.2015.7324142

Arulkumar, V. M. A., Lakshmi, M. A., & Rao, B. H. (2021). Super Resolution and Demosaicing based self learning Adaptive Dictionary Image Denoising framework. *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS).* doi:10.1109/ICICCS51141.2021.9432182

Asri, H., Mousannif, H., Moatassime, H. A., & Noel, T. (2016). Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis. *Procedia Computer Science*, *83*, 1064–1069. doi:10.1016/j.procs.2016.04.224

Beura, S., Majhi, B., Dash, R., & Roy, S. (2015b). Classification of mammogram using two-dimensional discrete orthonormal S-transform for breast cancer detection. *Healthcare Technology Letters*, *2*(2), 46–51. doi:10.1049/htl.2014.0108 PMID:26609404

Bhattad, P. B., & Jain, V. (2020). *Artificial Intelligence in Modern Medicine – The Evolving Necessity of the Present and Role in Transforming the Future of Medical Care. Cureus*. doi:10.7759/cureus.8041

Chu, W., Lee, K.-T., Luo, W., Bhambri, P., & Kautish, S. (2021). Predicting the security threats of internet rumors and spread of false information based on sociological principle. *Computer Standards & Interfaces*, *73*, 103454. doi:10.1016/j.csi.2020.103454

Doreswamy, & M, U. S. (2016). A Binary Bat Inspired Algorithm for the Classification of Breast Cancer Data. *International Journal on Soft Computing, Artificial Intelligence and Applications, 5*(2/3), 01–21. 10.5121/ijscai.2016.5301

Ganggayah, M. D., Taib, N. A., Har, Y. C., Lio, P., & Dhillon, S. K. (2019). Predicting factors for survival of breast cancer patients using machine learning techniques. *BMC Medical Informatics and Decision Making*, *19*(1), 48. Advance online publication. doi:10.1186/s12911-019-0801-4 PMID:30902088

Gao, F., Wu, T., Li, J., Zheng, B., Ruan, L., Shang, D., & Patel, B. (2018). SD-CNN: A shallow-deep CNN for improved breast cancer diagnosis. *Computerized Medical Imaging and Graphics*, *70*, 53–62. doi:10.1016/j.compmedimag.2018.09.004 PMID:30292910

Gupta, N. (2021). Transforming OMFS through Artificial Intelligence and Machine Learning. *Acta Scientific Otolaryngology*, *9–10*. Advance online publication. doi:10.31080/ASOL.2021.03.0352

Hamian, M., Darvishan, A., Hosseinzadeh, M., Lariche, M. J., Ghadimi, N., & Nouri, A. (2018). A framework to expedite joint energy-reserve payment cost minimization using a custom-designed method based on Mixed Integer Genetic Algorithm. *Engineering Applications of Artificial Intelligence*, *72*, 203–212. doi:10.1016/j.engappai.2018.03.022

Houssein, E. H., Emam, M. M., Ali, A. A., & Suganthan, P. N. (2021). Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review. *Expert Systems with Applications*, *167*, 114161. doi:10.1016/j.eswa.2020.114161

Jeyasingh, S., & Veluchamy, M. (2017). Modified Bat Algorithm for Feature Selection with the Wisconsin Diagnosis Breast Cancer (WDBC) Dataset. *Asian Pacific Journal of Cancer Prevention*, *18*(5), 1257–1264. doi:10.22034/APJCP.2017.18.5.1257 PMID:28610411

Juneja, S., Juneja, A., Dhiman, G., Behl, S., & Kautish, S. (2021b). An Approach for Thoracic Syndrome Classification with Convolutional Neural Networks. *Computational and Mathematical Methods in Medicine*, *2021*, 1–10. doi:10.1155/2021/3900254 PMID:34594396

Kelly, K. M., Dean, J., Comulada, W. S., & Lee, S. J. (2010). Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. *European Radiology*, *20*(3), 734–742. doi:10.1007/s00330-009-1588-y PMID:19727744

Leng, H., Li, X., Zhu, J., Tang, H., Zhang, Z., & Ghadimi, N. (2018). A new wind power prediction method based on ridgelet transforms, hybrid feature selection and closed-loop forecasting. *Advanced Engineering Informatics*, *36*, 20–30. doi:10.1016/j.aei.2018.02.006

Libson, S., & Lippman, M. (2014). A review of clinical aspects of breast cancer. *International Review of Psychiatry (Abingdon, England)*, *26*(1), 4–15. doi:10.3109/09540261.2013.852971 PMID:24716497

Muslim, M. A., Rukmana, S. H., Sugiharti, E., Prasetiyo, B., & Alimah, S. (2018). Optimization of C4.5 algorithm-based particle swarm optimization for breast cancer diagnosis. *Journal of Physics: Conference Series*, *983*, 012063. doi:10.1088/1742-6596/983/1/012063

Nilashi, M., Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017). A knowledge-based system for breast cancer classification using fuzzy logic method. *Telematics and Informatics*, *34*(4), 133–144. doi:10.1016/j.tele.2017.01.007

Obulesu, O., Kallam, S., Dhiman, G., Patan, R., Kadiyala, R., Raparthi, Y., & Kautish, S. (2021). Adaptive Diagnosis of Lung Cancer by Deep Learning Classification Using Wilcoxon Gain and Generator. *Journal of Healthcare Engineering*, *2021*, 1–13. doi:10.1155/2021/5912051 PMID:34691378

Pi, Y., Chen, Y., Deng, D., Qi, X., Li, J., Lv, Q., & Yi, Z. (2020). Automated diagnosis of multi-plane breast ultrasonography images using deep neural networks. *Neurocomputing*, *403*, 371–382. doi:10.1016/j.neucom.2020.04.123

Reyana, A., Krishnaprasath, V., Kautish, S., Panigrahi, R., & Shaik, M. (2020). Decision-making on the existence of soft exudates in diabetic retinopathy. *International Journal of Computer Applications in Technology*, *64*(4), 375. doi:10.1504/IJCAT.2020.112684

Rudra, T., & Kautish, S. (2021). Impact of Covid-19 Infodemic on the Global Picture. *Computational Intelligence Techniques for Combating COVID-19*, 333–353. doi:10.1007/978-3-030-68936-0_16

Sahu, B., Mohanty, S., & Rout, S. (2018). A Hybrid Approach for Breast Cancer Classification and Diagnosis. *ICST Transactions on Scalable Information Systems*, *0*(0), 156086. doi:10.4108/eai.19-12-2018.156086

Sampathkumar, A., Rastogi, R., Arukonda, S., Shankar, A., Kautish, S., & Sivaram, M. (2020). An efficient hybrid methodology for detection of cancer-causing gene using CSC for micro array data. *Journal of Ambient Intelligence and Humanized Computing*, *11*(11), 4743–4751. doi:10.1007/s12652-020-01731-7

Ting, F. F., Tan, Y. J., & Sim, K. S. (2019). Convolutional neural network improvement for breast cancer classification. *Expert Systems with Applications*, *120*, 103–115. doi:10.1016/j.eswa.2018.11.008

Varatharajan, A. C. S., & Varatharajan, V. (2020). Big Data Analytics in the Healthcare Industry: An Analysis of Healthcare Applications in Machine Learning With Big Data Analytics. In A. Haldorai & A. Ramu (Eds.), *Big Data Analytics for Sustainable Computing* (pp. 160–178). IGI Global. doi:10.4018/978-1-5225-9750-6.ch010

Kayal Padmanandam is an Associate Professor who has over a decade of credentials in the domain of Computer Science with wide exposure through teaching, research, and industry. She is appointed as a ACM Distinguished Speaker for computer community by ACM. She is passionate about research and specialized in Data Science and Machine Learning Algorithms in which she pursued her doctoral research. She has several publications especially related to machine-learning applications. Her research highly concentrates on supervised learning algorithms, measure optimization techniques, parameter selection & calibration for input-based model building, analytics on large predictor variables for any Intelligent Predictor System.

R. N. Thakur is currently working as a Program Leader/Assistant Professor in the Department of Information Technology at LBEF Campus (Asia Pacific University of Technology & Innovation), Kathmandu, Nepal. Mr. Thakur has completed his MSc in Information Technology from AAIDU, Allahabad, India in 2004 and is a PhD Scholar in Mewar University, India. His research interests include Software Development, Data Science, Machine Learning and Software Engineering, etc. He has published papers in various international journals.