Original Article

Scale Invariant Deep Neural Multiple Feature Learning Based Boosted Support Vector Entropy Classification for Breast Cancer Diagnosis using Mammograms

K. Sai Krishna[1], P. Grace Kanmani Prince[2]

[1,2]*Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, Chennai.*

[1]*Corresponding Author : saiphd810@gmail.com*

Abstract - An earlier diagnosis of cancer rises the living days of patients. The evolution of medical systems supports to identify the presence of cancer. In the state-of-the-art works, few methods were implemented for diagnosing breast cancer with different Deep Neural Network (DNN) support. However, the traditional DNN techniques were computationally expensive as it does provide better feature extraction performance with minimal time usage. Besides that, the misclassification error observed during the diagnosis process was higher, impacting the accuracy of early breast cancer prediction. Therefore, a novel Scale Invariant Deep Feature Learning Based Boosted Support Vector Entropy Classification (SIDFL-BSVEC) method is introduced. The SIDFL-BSVEC method is proposed to get better accuracy for the early identification of breast cancer with lesser time. The SIDFL-BSVEC method initially designs Scale Invariant Deep Neural Multiple Feature Extraction (SIDNMFE) algorithms to discover critical features in input mammograms with minimal time by working as robust to illumination variations, noise, partial occlusion, and minor viewpoint changes in mammograms images. These traits are significant for early cancer disease prediction when cells in the mammogram image have dissimilar sizes and orientations. After that, Support Vector Entropy Boosted Cancer Classifier (SVEBCC) algorithm is designed to minimize the error rate determined during the early analysis of breast cancer. The performance of the proposed CWDCMFE-MBRC technique is verified by taking the parameters such as disease diagnosis accuracy, disease diagnosis time, sensitivity, specificity and error rate along with various numbers of input images.

Keywords - Scale invariant features transform, Deep learning, Boosting, Support vector entropy, Morphological features, Texture features, Density features.

1. Introduction

Breast cancer is a cell with superfluous mass development in women's breasts. This breast tissue formulates the tumour, labelled as benign or malignant. The researchers studied diverse automated systems to find breast cancer. Early discovery and analysis is the preeminent and effectual approach to managing the tumour progress. Mammography is suggested imaging for early prediction and identification of breast tumours. Categorizations of masses in mammograms are still a considered problem to be resolved.

The two-staged classifier was employed in [1] to boost the efficiency of breast abnormality identification. However, the abnormality classification performance was insufficient when taking mammograms with illumination variations, noise, and partial occlusion as input. DL-enabled Capsule Network was constructed in [2] to obtain higher breast cancer finding performance. However, the accuracy of cancer classification was insufficient for early prediction.

The Modified Entropy Whale Optimization Algorithm (MEWOA) was implemented in [3] to perform the deep feature selection and classification task more accurately. However, the error rate was not minimized.

In-Depth Feature Learning Algorithm was presented in [4] to accomplish breast abnormality prediction and robust classification. However, the complexity measured during the breast abnormality classification was much higher. Deep Breast Cancer Net DL representation was developed in [5] to increase breast cancer discovery and categorization performance. However, the recall determined during the classification process was not better.

An automatic diverse feature depending on the breast cancer recognition system was intended in [6] to label an input mammogram as normal or abnormal. However, the cancer classification performance was ineffective, considering the massive number of input images.

Deep Learning-Based, Real-Time Discriminate Correlation examination was performed in [7] to decrease breast cancer recognition error rate. However, the time taken for accurate disease diagnosis was more. Various machine learning algorithms were constructed for breast cancer image categorization was discussed in [8].

The Deep Learning (DL) concept was presented in [9] to find the features by analyzing the high dimensional and associated information. However, the time taken for robust feature extraction was not reduced. Novel CNN classifiers were employed in [10] to predict breast cancer via labelling mammogram images as benign or cancer. However, the classification performance was poor in attaining early prediction of cancer.

2. Literature Survey

A Convolution Neural Network (CNN) based classification method was planned in [11] for attaining accurate diagnosis. However, a misclassification error was observed while considering the large number of images as input. Intuitionistic Fuzzy Evidential Reasoning (IFER) concept was intended in [12] to give better accuracy for early diagnosis. However, the sensitivity and specificity of cancer prediction were insufficient.

Nature-inspired metaheuristic optimized CNN was utilized in [13] to classify the diseases in breast cancer images using less time. However, robust feature extraction from input images was not attained. Deep CNN was implemented in [14] to attain better breast segmentation and mass recognition performance. However, the time required for practical cancer analysis was not reduced.Snapshot Ensemble Deep Learning algorithm was designed in [15] to categorize breast cancer. However, the accuracy was not higher. Gated Graph Attention Network (GGAT) was presented in [16] for precise cancer prediction. However, the mammogram images were not considered. The Back Propagation Boosting Recurrent Wienmed model was implemented in [18] to identify breast cancer in an earlier phase. Though, the computational complexity was not lower.

A study of different CAD systems intended for mass recognition and categorization [19] was presented in [22]. A CAD system using normalized CNN features was designed in [23] to predict two categories of mammogram cancers. Adaptive Convolutional Feature Descriptor Selection (AFDS) was performed in [21] to increase mammograms' classification performance. However, the time for the significant feature selection process was not reduced.This work aims to address the above talked drawbacks in state-of-the-art algorithms, so the SIDFL-BSVEC method is introduced.

The significant functions of the proposed SIDFL-BSVEC method are discussed below,

- To increase feature extraction's robustness while taking mammogram images as input with minimal complexity, the SIDNMFE algorithm is intended in the proposed SIDFL-BSVEC method, which finds only significant features in mammograms through deep analysis using SIFT. Contrary to traditional works, SIFT is utilized in the SIDNMFE algorithm because it is very proficient and assists in decreasing the dimensions of the feature space by eliminating redundant features.

- To minimize the misclassification error involved during the early breast cancer diagnosis [20] with lesser computational time, BSVEC is introduced in the proposed SIDFL-BSVEC method, which boosts breast cancer classification accuracy by combining several weak SVE classifiers into stronger learners. This supports for proposed SIDFL-BSVEC method to achieve higher diagnosis accuracy.

3. Proposed SIDFL-BSVEC Method

The Proposed SIDFL-BSVEC Method is planned in this paper by integrating the SIDNMFE and BSVEC concepts to reduce the error rate of cancer analysis with better time complexity. Unlike the conventional feature extraction algorithm, SIDFL-BSVEC Method implements a novel SIDNMFE to accomplish a robust feature extraction process by combining the SIFT in existing DL concepts. The state-of-the-art DL concepts were computationally expensive, and their feature extraction performance was impacted when considering many input mammogram images with illumination variations, noise, partial occlusion, and minor viewpoint changes. For that reason, SIDNMFE is introduced in this research work.

Also, SIDFL-BSVEC Method implements a novel BSVEC to boost the classification performance of mammograms by applying a weighted gradient boosting concept in the base Support Vector Entropy (SVE) classifier. Thus, BSVEC work as a robust classifier, classifying all input mammogram images as benign or malignant with less time utilization. The architecture design of the proposed SIDFL-BSVEC Method is depicted in Figure 1 illustrates the overall processing diagram of the SIDFL-BSVEC Method. At first, SIDFL-BSVEC Method considers several mammograms as input, and subsequently, SIDNMFE is applied where robust feature extraction is done and thereby predicts multiple vital features (i.e. morphological, texture and density features) with lesser time complexity. For extracted key features, SVEBCC is applied to boost the mammogram classification performance and achieve more cancer diagnosis accuracy.
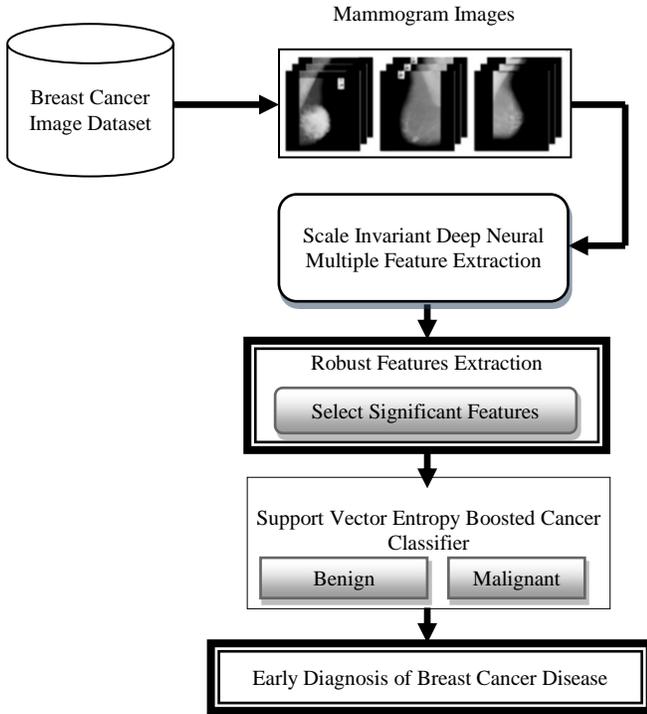
Mammogram Images

Breast Cancer Image Dataset

Scale Invariant Deep Neural Multiple Feature Extraction

Robust Features Extraction

Select Significant Features

Support Vector Entropy Boosted Cancer Classifier

Benign

Malignant

Early Diagnosis of Breast Cancer Disease

Fig. 1 Architecture of proposed SIDFL-BSVEC method for early diagnosis

3.1. Scale Invariant Deep Neural Multiple Feature Extraction

The Scale Invariant Deep Neural Multiple Feature Extraction (SIDNMFE) algorithm is constructed to raise the feature extraction performance of cancer identification when getting a massive number of input mammogram images with illumination differences, noise, partial occlusion, and little viewpoint variations. The SIDNMFE algorithm employs several layers to extract significant features from input mammograms.

The SIDNMFE is a deep feed forward neural network that takes the number of mammogram images. $\mu_1, \mu_2, .. \mu_z$ as input at the input layer. After that, hidden layers in SIDNMFE perform a robust feature extraction process, finding the most significant features presented in mammograms with minimal time consumption. The output layer in SIDNMFE presents extracted key cancerous features for correctly identifying the existence of cancer diseases. The process of SIDNMFE is depicted in Figure 2. Figure 2 explains the processing diagram of SIDNMFE to reduce the computational time of feature extraction and thereby attain better breast cancer analysis performance. As demonstrated in the above processing design, SIDNMFE, at the start, defines a deep network with arbitrary weights.

The input layer in SIDNMFE obtains the number of mammogram images as input, and then it is forwarded to hidden layers. Let us consider the number of mammogram

images in the input database is defined as '$\mu_1, \mu_2, .. \mu_z$' whereas z depicts the total number of mammograms considered for the simulation process.

Scale-Space Peak Selection

Keypoint Localization

Orientation Assignment

Keypoint Descriptor

Input Mammograms

Robust Feature Extraction
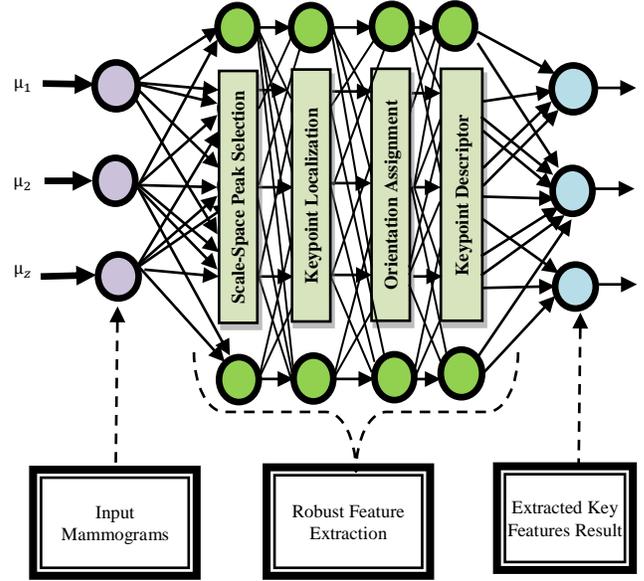
Extracted Key Features Result

Fig. 2 Processing diagram of SIDNMFE

Consequently, the input layer '(E)' process is mathematically done using the below expression,

$$E(t) = \sum \mu_i \, £_{EF} + bi_j \qquad (1)$$

In the above mathematical description (1), '$E(t)$' defines the tasks of the input layer. Here, 'μ_i' describes an input mammogram image taken for diagnosing the breast cancer and. '$£_{EF}$' depicts the weight of input and hidden layer and. 'bi_j' expresses a bias. After taking the number of mammograms input, SIDNMFE applies SIFT concepts at the hidden layers to enhance the feature extraction performance of cancer detection.

Because SIFT helps to precisely predict the key cancerous features in input mammograms by operating well to illumination deviations, noise, partial occlusion, and minor viewpoint variations. SIDNMFE performs a scale-space peak selection process at the first hidden layer. Followed by the first hidden layer task at the time 't' is performed using the below equation.

$$F_1(t) = \sum E(t) \, £_{F_1}[SPS] \qquad (2)$$

From the above mathematical determination (2), '$F_1(t)$' expressed the first hidden layer outcomes. Here, '$E(t)$' depicts the mammograms. '$\mu_1, \mu_2, .. \mu_z$' get from the input layer and. '$£_{F_1}$' describes the weight of the primary hidden layer, whereas 'SPS' symbolizes scale-space peak selection. During this process, an input mammogram image is

described in multiple scales to find exciting points (i.e. significant key features) across different scales with the support of Gaussian kernel using the below formulation,

$$G(a, b, \beta) = \frac{1}{2\pi\beta^2} \frac{-(a^2+b^2)}{2\beta^2} \qquad (3)$$

In the above mathematical presentation (3), where a, b depicts the coordinates of each pixel, and β explains the parameter associated with the scale. To define mammogram image. 'μ_i' in multiple scales, the convolution of the image with the kernel at each scale is mathematically obtained using below,

$$L(a, b, \beta) = G(a, b, \beta) * I(a, b) \qquad (4)$$

Then, SIDNMFE accomplishes key-point localization at the second hidden layer with the aid of the below formulation,

$$F_2(t) = \sum F_1(t)\, \pounds_{F_2}[KPL] \qquad (5)$$

In the above mathematical calculation (5), '$F_2(t)$' illustrates second hidden layer output and. '$F_1(t)$' point outs the primary hidden layer result and '$\pounds_{F_2}$' defines the weight of the second hidden layer. Here, 'KPL' defines key-point localization where interesting points are predicted with the support of the difference of Gaussians using the following mathematical equation,

$$DoG(u, v) = \mu_i \left(\frac{1}{2\pi\beta^2} e^{-\frac{u^2+v^2}{2\beta^2}} - \frac{1}{2\pi K^2\beta^2} e^{-\frac{u^2+v^2}{2K^2\beta^2}} \right) \qquad (6)$$

After that, SIDNMFE carried out the orientation assignment task at the third hidden layer using the below,

$$F_3(t) = \sum F_2(t)\pounds_{F_3}[OTA] \qquad (7)$$

In the above mathematical depiction (5), '$F_3(t)$' expresses third hidden layer outcomes and. '$F_2(t)$' describes the result of subsequent hidden layers. Here '$\pounds_{F_3}$' is the weight of third hidden layers, and 'OTA' defines processes of orientation assignment where several exciting points are predicted at specific scales through estimating the gradient magnitude $m(u, v)$ and orientation $O(u, v)$ using below,

$$m(u, v) = Lu + 1, v - Lu - 1, v2 + \left(L(u, v+1) - L(u, v-1) \right)^2 \qquad (8)$$

$$O(u, v) = \tan^{-1} \left(\frac{L(u, v+1) - L(u, v-1)}{L(u+1, v) - L(u-1, v)} \right) \qquad (9)$$

From the above mathematical equations (8) and (9), SIDNMFE predicts several exciting mammogram points.

Followed by the fourth hidden layer. '$F_4(t)$' act as a keypoint descriptor where useful discriminative cancerous features are obtained using the below,

$$F_4(t) = \sum F_3(t)\pounds_{F_4}[KPD] \qquad (10)$$

In the above mathematical expression (10), '$F_3(t)$' defines the input acquired from the third hidden layer and. '$\pounds_{F_4}$' illustrates the fourth hidden layer weight, and 'KPD' describes the critical point descriptor. The discovered key cancerous features are after that forwarded to the output layer '$G(t)$' which procedure is mathematically done using,

$$G(t) = X\pounds_{FG}F_4(t) \qquad (11)$$

In the above mathematical estimation (11), '$\pounds_{FG}$' describes hidden and output layer weight, whereas 'X' defines an activation function. At last, the output layer in SIDNMFE presents discovered morphological, texture, and density features as key cancerous features. With the support of the above procedure of SIDNMFE, the proposed method achieved better feature extraction performance with minimal computational time for early recognition of breast cancer compared to existing works.

3.2. Support Vector Entropy Boosted Cancer Classifier

Support Vector Entropy Boosted Cancer Classifier (SVEBCC) is constructed in SIDFL-BSVEC Method to reduce the error of diagnosing breast cancer at an early phase. The SVEBCC integrates Support Vector Entropy Classifier (SVEC) and weighted gradient boosting concepts. In SVEBCC, base SVEC is a classifier that returns a relatively poor classification for correctly recognizing the existence of breast cancer in input mammograms.

With aiming at raising the classification performance of base SVEC, weighted gradient boosting concepts is employed in this research works in which diagnosis result of several base SVEC are united to lessen error. At the start, each input mammogram image. 'μ_i' and their extracted key features (ω_i) are trained with the support of base SVEC, which is defined as '$\{(\omega_1, c_1), (\omega_2, c_2), \dots (\omega_n, c_n)\}$'. Here, ω_i describes the set of training mammogram images and C_i depicts the output class (disease diagnosis result). The base SVEC separates benign and malignant classes via a marginal hyperplane. The SVEC in SVEBCC predicts optimal marginal hyperplane to categorize the all mammogram with the aid of the entropy technique, where it utilizes a split approach and is estimated concerning their class label.

The base SVEC classification task is done by computing the divide with the better information gain. For mammograms, i.e. input database '$DB = \mu_1, \mu_2, \mu_3, \dots \mu_z$', if '$DB$' is partitioned into two intervals. 'N_1' and 'N_2' using boundary 'α', the entropy is obtained as,

$$I(DB, \alpha) = \frac{|N_1|}{DB} Ent(N_1) + \frac{|N_2|}{DB} Ent(N_2) \qquad (12)$$

In SVEBCC, the chance of class 'i' in the interval. 'N_i' is estimated using,

$$Ent(N_1) = -\sum_{i=1}^{z} \mu_i \log_2(\mu_i) \qquad (13)$$

Using the mathematical formulation (12) and (13), boundary that lessens the entropy function is predicted as a binary discretization to categorize all mammograms into benign or malignant classes.

For boosting the classification accuracy of cancer diagnosis, 'k' number of base classifier output for every mammogram is obtained, consequently arbitrary weight. 's_i' defined for each base SVEC. '$w_i(d_i)$' using below,

$$s_i \to \sum_{i=1}^{k} BC_i(\mu_i) \qquad (14)$$

Here, 's_i' describes the initialized weight of base SVEC. '$BC_i(d_i)$' and 'μ_i' defines input mammograms. Next, negative gradient 'ε' is obtained using the below,

$$\varepsilon = \left(AR_i - BC(\mu_i)\right)^2 \qquad (15)$$

Where, 'AR_i' shows the actual classification outcome, and '$BC(\mu_i)$' describes predicted output with the support of a base classifier. At that moment, SVEBCC fits a base classifier '$BC(\mu_i)$' to negative gradient 'ε' using below,

$$\rho_i(\mu_i) = (P, (AC_i - BC(\mu_i)) \qquad (16)$$

Following, SVEBCC modifies the weights of weak classifiers by considering their negative gradient using,

$$\Delta s_i \to \sum_{i=1}^{n} BC_i(\mu_i) \qquad (17)$$

Here, '$\Delta \pi_i$' expresses the modified weight of. '$BC_i(\mu_i)$'. Thus, $BC_i(\mu_i)$ with the most significant weight determined as a robust classifier using,

$$Y_i' = arg \max_{k} s (BC_i(\mu_i)) \qquad (18)$$

Here, 'Y_i'' shows the final strong classification output for an input mammogram image. Here '$arg \max_{k} s$' is described to identify $BC_i(\mu_i)$ with better weight as a solid classifier, it correctly categorizes all mammograms into benign or malignant classes with less time consumption.

The exhaustive algorithmic procedure of the SIDFL-BSVEC Method is described as follows,

Algorithm 1 Scale invariant deep feature learning based boosted support vector entropy classification

// Scale Invariant Deep Feature Learning Based Boosted Support Vector Entropy Classification Algorithm
Input : Number of mammogram images in Database '$DB = \mu_1, \mu_2, .. \mu_z$';
Output : Enhanced diagnosis accuracy
Step 1 : Begin
Step 2 : Consider DNN with arbitrary weights
Step 3 : For '$\mu_i \in DB$'
Step 4 : While (the error is lesser) do
Step 5 : Obtains input mammograms using (1) // Input Layer
Step 6 : Give taken mammograms 'μ_i' to the hidden layer
Step 7 : Apply SIFT // Hidden Layers
Step 8 : Perform Scale-space peak selection using (2), (3) and (4)
Step 9 : Locate Keypoints using (5) and (6)
Step 10 : Orientation Assignment using (7), (8) and (9)
Step 11 : Keypoint descriptor finds key features using (10)
Step 12 : Returns extracted key features using (11) // Output Layer
Step 13 : For each extracted vital feature of the mammogram 'μ_i'
Step 14 : Design 'k' number of base classifier output using (12) and (13)
Step 15 : For weak classifier '$BC_i(\mu_i)$'
Step 16 : Consider weight 's_i' using (14)
Step 17 : Estimate 'ε' using (15)
Step 18 : Fit a $BC_i(\mu_i)$ to 'ε' using (16)
Step 19 : Update weights 'Δs_i' using (17)
Step 20 : Find out robust classifier using (18)
Step 21 : Strong classifier presents exact categorization output 'y_i''
Step 22 : Correctly find the occurrence of breast cancer in mammograms
Step 23 : End For
Step 24 : End For
Step 25 : End While
Step 26 : End For

Algorithm 1 describes the pseudo code representation of the SIDFL-BSVEC Method to attain higher breast cancer disease diagnosis performance. With the application of the above algorithmic procedures, SIDFL-BSVEC Method provides better accuracy, time and error rate to precisely discover the presence of breast cancer disease when compared to conventional works.

4. Simulation Setup

SIDFL-BSVEC Method and conventional Two-staged classifier [1] and DL-enabled Capsule Network (DLCN) [2] are implemented with the support of MATLAB simulator. For analyzing their simulation performance, CBIS-DDSM: Breast Cancer Image Dataset (CBIS-DDSM: Breast Cancer Image Dataset | Kaggle) is taken as input, comprising 10239 mammograms with distinct categories, i.e. normal, benign, and malignant. The testing mammogram images are taken in the range of 100 to 1000 in order to test the breast cancer diagnosis performance. The performance outcome of the proposed SIDFL-BSVEC method is compared with the traditional Two-staged classifier [1] and DLCN [2]. The testing results of the CWDCMFE-MBRC technique are analyzed by considering the following metrics,

❖ Disease Diagnosis Accuracy
❖ Disease Diagnosis Time
❖ Error Rate
❖ Sensitivity
❖ Specificity

4.1. Scenario 1: Performance Measure of Accuracy for Breast Cancer Diagnosis

Disease Diagnosis Accuracy 'DDA' is the ratio of input mammograms that are rightly classified into corresponding classes (i.e. benign or malignant). The DDA is mathematically obtained as,

$$DDA = \frac{\mu_c}{z} * 100 \qquad (19)$$

From the above mathematical representation (19), 'μ_c' defines the number of input mammogram images correctly categorized, and 'z' expresses the total number of mammogram images acquired as input to carry out simulation evaluation. The DDA of breast cancer is obtained in percentage (%).

Table 1. Accuracy of early breast cancer analysis

Number of Input Mammogram Images (z)	Disease Diagnosis Accuracy (%)		
	SIDFL-BSVEC	Two-Staged Classifier	DLCN
100	96	89	87
200	94	86	82
300	96	83	80
400	93	85	83
500	94	83	81
600	95	84	80
700	98	86	81
800	95	84	82
900	97	88	84
1000	96	90	85

Table 1 and Figure 3 present the testing results of accuracy observed during the simulation process for achieving better early breast cancer diagnosis performance versus the varied number of input mammograms for the proposed SIDFL-BSVEC method and two-staged classifier [1] and DLCN [2].
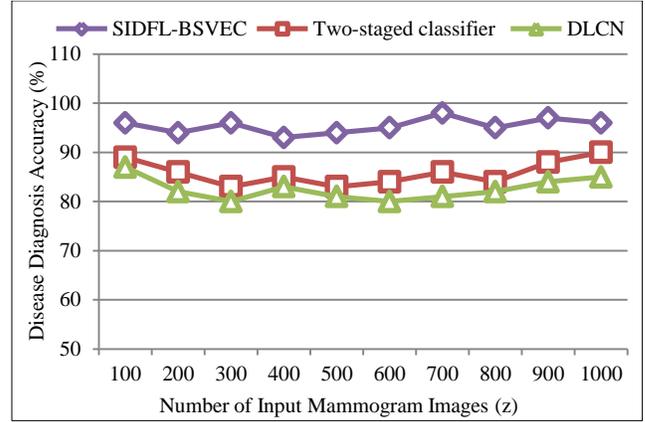

Fig. 3 Comparative graphical presentation of disease diagnosis accuracy

From the above comparative performance validation in tabulation and graphical diagram, it is considerable that the proposed SIDFL-BSVEC method gets more diagnosis accuracy for finding the presence of breast cancer in mammogram images as compared to the existing Two-staged classifier [1] and DLCN [2]. The cancer diagnosis accuracy is enhanced in the proposed SIDFL-BSVEC method because of implementing a novel SIDNMFE algorithm for the key features selection process and the SVEBCC model for predicting cancer occurrence contrary to traditional works. As a result, the proposed SIDFL-BSVEC method raises the number of input mammograms that are rightly categorized into benign or malignant classes compared to existing classification systems. Thus, the proposed CWDCMFE-MBRC technique increased disease diagnosis accuracy by 11 % and 16 % compared to traditional Two-staged classifiers [1] and DLCN [2].

4.2. Scenario 2: Performance Measure of Disease Diagnosis Time

Disease Diagnosis Time 'DDT' is determined as the amount of time taken to classify an input mammogram as benign or malignant precisely. Thus, DDT of breast cancer is mathematically estimated as,

$$DDT = z * time(RCSM) \qquad (20)$$

In the above mathematical demonstration (20), '$time(RCSM)$' depicts an amount of time utilized to rightly classify the one mammogram image as benign or malignant and 'z' portrays the number of mammogram images get as input during the implementation. The DDT is estimated in milliseconds (ms). Table 2 and Figure 4 depict the disease

diagnosis time testing performance versus the diverse number of input mammogram images using the SIDFL-BSVEC method and Two-staged classifier [1] and DLCN [2].

Table 2. Diagnosis time of breast cancer disease

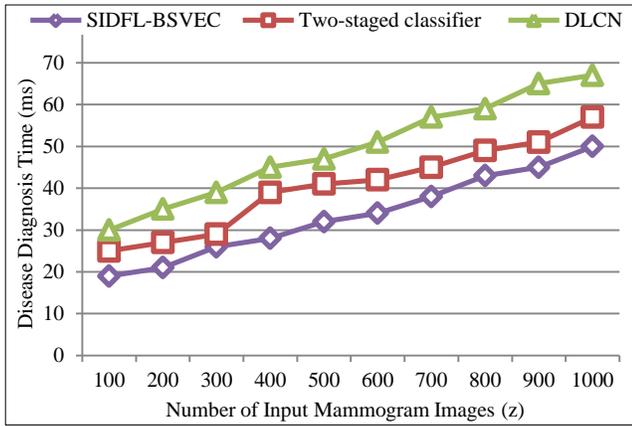| Number of Input Mammogram Images (*z*)	Disease Diagnosis Time (ms)		
	SIDFL-BSVEC	Two-Staged Classifier	DLCN
100	19	25	30
200	21	27	35
300	26	29	39
400	28	39	45
500	32	41	47
600	34	42	51
700	38	45	57
800	43	49	59
900	45	51	65
1000	50	57	67

Fig. 4 Comparative graphical presentation of disease diagnosis time

From the above table value and graphical assessment, it is expressive that the disease diagnosis time using the proposed SIDFL-BSVEC method is better when compared to the existing Two-staged classifier [1] and DLCN [2].

The minimal value of diagnosis time is determined in the proposed SIDFL-BSVEC method because it planned a new SIDNMFE for detecting key features with less time consumption and SVEBCC for effectively diagnosing the occurrence of cancer with reduced time complexity. Therefore, the proposed SIDFL-BSVEC method lessens the time required to rightly classify an input mammogram as benign or malignant compared to the Two-staged classifier [1] and DLCN [2]. Hence, the proposed CWDCMFE-MBRC technique minimized disease diagnosis time by 18 % and 33 % compared to traditional Two-staged classifiers [1] and DLCN [2].

4.3. Scenario 3: Performance Measure of Error Rate

Error Rate '$Error_R$' is determined by considering the number of input mammograms erroneously classified as benign or malignant. The misclassification error of breast cancer diagnosis is mathematically obtained as,

$$Error_R = \frac{\mu_{iacc}}{z} * 100 \qquad (21)$$

From the above mathematical expression (21), 'μ_{iacc}' defines the number of input mammogram images inaccurately classified, and 'z' states the total number of input mammograms assumed as input during the simulation process. The false rate of breast cancer analysis is determined in percentage (%).

Table 3. Error rate for early diagnosis of breast cancer

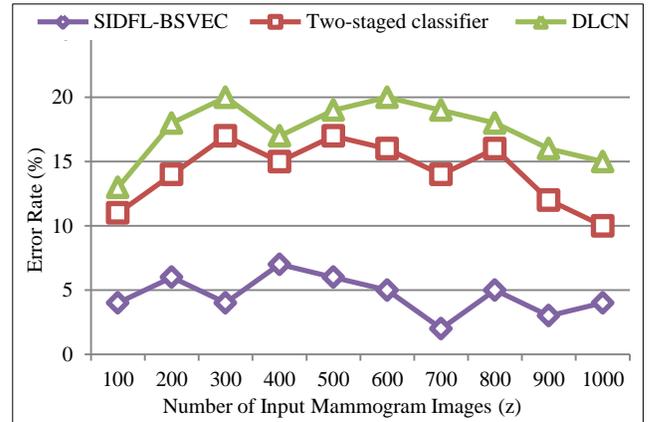| Number of Input Mammogram Images (*z*)	Error Rate (%)		
	SIDFL-BSVEC	Two-Staged Classifier	DLCN
100	7	12	14
200	9	16	19
300	8	18	22
400	11	17	21
500	10	18	22
600	9	19	23
700	6	15	20
800	9	18	21
900	7	13	17
1000	8	12	18

Fig. 5 Comparative graphical presentation of error rate

Table 3 and Figure 5 display the comparative testing evaluation of the error rate for breast cancer disease diagnosis versus the varied number of input mammograms using the SIDFL-BSVEC method and Two-staged classifier [1] and DLCN [2]. In the above tabulation and comparative graphical judgment, it is expressive that the error rate using the proposed CWDCMFE-MBRC technique is relatively better for recognizing breast cancer earlier compared to the

existing Two-staged classifier [1] and DLCN [2]. The lesser error rate value is obtained in proposed SIDFL-BSVEC method since it develops SVEBCC to decrease the misclassification performance by combining several base classifier results of cancer diagnosis. As a consequence, the proposed SIDFL-BSVEC method correctly notices the incidence of breast cancer in mammograms. Therefore, the SIDFL-BSVEC method lessens the number of input mammograms inaccurately categorized. Therefore, the proposed SIDFL-BSVEC method reduced error rates by 67 % and 73 % compared to existing Two-staged classifiers [1] and DLCN [2], respectively.

4.4. Scenario 4: Performance Measure of Sensitivity

Sensitivity ($Sens_{ty}$) is determined as the ratio of the number of mammograms which contains malignant disease that is rightly categorized as malignant to the total quantity of mammograms considered. Thus, sensitivity is obtained as,

$$Sens_{ty} = \frac{TP}{TP+FN} * 100 \tag{22}$$

In the above mathematical depiction (22), 'TP' expresses the true positive (i.e. the number of mammograms with malignant disease is strictly classified as malignant), and FN states that mistakenly classified mammograms. The sensitivity of early cancer diagnosis is estimated in percentage (%).

Table 4. Sensitivity for early classification of breast cancer

Number of Input Mammogram Images (z)	Sensitivity (%)		
	SIDFL-BSVEC	Two-Staged Classifier	DLCN
100	92	85	82
200	89	81	77
300	90	78	75
400	88	80	76
500	89	79	75
600	86	78	74
700	89	82	76
800	86	79	75
900	91	84	80
1000	90	83	79

Table 4 and Figure 6 demonstrate simulation outcomes of sensitivity for analyzing the breast cancer disease versus different numbers of input mammograms using the SIDFL-BSVEC method and Two-staged classifier [1] and DLCN [2]. As demonstrated in both the tabulation and graphical results investigation, it is significant that the sensitivity using the proposed SIDFL-BSVEC method is better when compared to the existing Two-staged classifier [1] and DLCN [2]. The sensitivity is enhanced in the proposed SIDFL-BSVEC method because the SVEBCC process

constructs a robust classifier to increase breast cancer diagnosis performance with lesser error considerably.
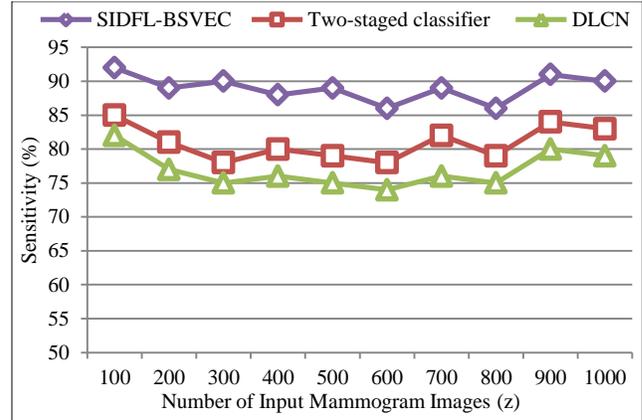

Fig. 6 Comparative graphical presentation of sensitivity for breast cancer diagnosis

Table 4 and Figure 6 demonstrate simulation outcomes of sensitivity for analyzing the breast cancer disease versus different numbers of input mammograms using the SIDFL-BSVEC method and Two-staged classifier [1] and DLCN [2]. As demonstrated in both the tabulation and graphical results investigation, it is significant that the sensitivity using the proposed SIDFL-BSVEC method is better when compared to the existing Two-staged classifier [1] and DLCN [2].

The sensitivity is enhanced in the proposed SIDFL-BSVEC method because the SVEBCC process constructs a robust classifier to increase breast cancer diagnosis performance with lesser error considerably. Therefore, the proposed SIDFL-BSVEC method boosts the ratio of mammograms, including a malignant disease that is rightly classified as malignant. Hence, the proposed SIDFL-BSVEC method enlarged the sensitivity of cancer diagnosis by 10 % and 13 % compared to existing Two-staged classifier [1] and DLCN [2], respectively.

4.5. Scenario 5: Performance Measure of Specificity

Specificity ($Spec_{ty}$) is determined as the ratios of the number of mammograms, including benign, are precisely categorized as benign to the total quantity of mammograms assumed. Therefore, specificity is obtained as,

$$Spec_{ty} = \frac{TN}{TN+FP} * 100 \tag{23}$$

In the above mathematical presentation (23), TN describes the True Negative (i.e. rightly classified mammogram images which contain benign as normal patients) and False Positive (FP) defines the inexactly classified as mammogram as benign or malignant. The specificity of breast cancer disease prediction [17] is acquired in percentage (%).

Table 5. Specificity for early recognition of breast cancer

Number of Input Mammogram Images (*z*)	Specificity (%)		
	SIDFL-BSVEC	Two-Staged Classifier	DLCN
100	90	83	80
200	86	79	74
300	88	75	72
400	86	78	75
500	87	77	73
600	84	76	74
700	87	80	75
800	84	77	76
900	89	82	79
1000	88	81	78
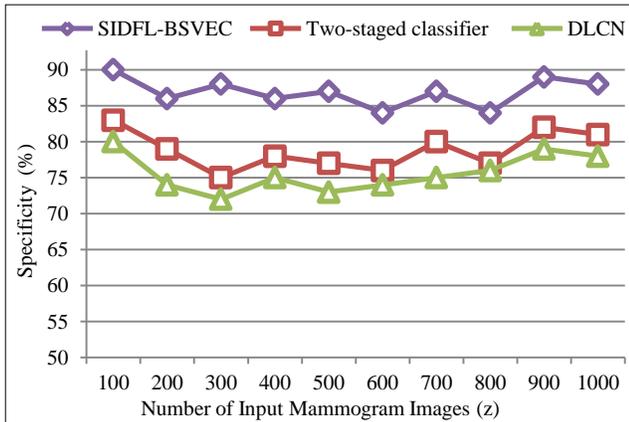
Fig. 7 Comparative graphical presentation of specificity

Table 5 and Figure 7 describe the graphical performance of specificity for classifying breast cancer versus dissimilar number of input mammograms using the SIDFL-BSVEC method and Two-staged classifier [1] and DLCN [2]. As the above comparative graphical analysis exposed, the CWDCMFE-MBRC specificity was better when compared to the existing Two-staged classifier [1] and DLCN [2]. The specificity is improved in the proposed SIDFL-BSVEC method, where SVEBCC identifies a robust classifier.

Accordingly, the proposed SIDFL-BSVEC method boosts the ratios of mammograms, which are rightly classified as benign. Hence, the proposed SIDFL-BSVEC method enhanced the specificity of cancer diagnosis by 10 % and 15 % compared to the existing Two-staged classifier [1] and DLCN [2], respectively.

5. Conclusion

This article introduces the SIDFL-BSVEC method to get better mammogram classification accuracy for analyzing breast cancer with reduced time utilization. The aim of the proposed SIDFL-BSVEC method is achieved by implementing SIDNMFE and SVEBCC for efficiently diagnosing breast cancer disease earlier using mammograms. The proposed SIDFL-BSVEC method decreases the misclassification performance of mammogram images and thereby significantly recognizes the presence of cancer disease with the support of SVEBCC.

The proposed SIDFL-BSVEC method also reduced the time to effectively extract the critical features in mammograms by employing the SIDNMFE process. Also, the proposed SIDFL-BSVEC method lessens the error rate of breast cancer diagnosis by designing a solid classifier with the best gradient descent step size. The testing performance results prove that the proposed SIDFL-BSVEC method enhanced disease diagnosis performance of breast cancer, an improvement of accuracy averagely by 14% and minimization of time by 25 % when compared to conventional diagnosis systems.

References

[1] Farnoosh Azour, and Azzedine Boukerche, "An Efficient Transfer and Ensemble Learning Based Computer Aided Breast Abnormality Diagnosis System," *IEEE Access*, vol. 11, pp. 21199-21209, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] T. Kavitha et al., "Deep Learning Based Capsule Neural Network Model for Breast Cancer Diagnosis using Mammogram Images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 14, pp. 113–129, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Saliha Zahoor, Umar Shoaib, and Ikram Ullah Lali, "Breast Cancer Mammograms Classification using Deep Neural Network and Entropy-Controlled Whale Optimization Algorithm," *Diagnostics*, vol. 12, no. 2, p. 557, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] Tariq Mahmood et al., "An Automated In-Depth Feature Learning Algorithm for Breast Abnormality Prognosis and Robust Characterization from Mammography Images using Deep Transfer Learning," *Biology*, vol. 10, no. 9, p. 859, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Asaf Raza et al., "Deep Breast Cancer Net: A Novel Deep Learning Model for Breast Cancer Detection using Ultrasound Images," *Applied Sciences*, vol. 13, no. 4, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Naveed Chouhan et al., "Deep Convolutional Neural Network and Emotional Learning Based Breast Cancer Detection using Digital Mammography," *Computers in Biology and Medicine*, vol. 132, p. 104318, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Manisha Bhende et al., "Deep Learning-Based Real-Time Discriminate Correlation Analysis for Breast Cancer Detection," *BioMed Research International*, vol. 2022, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Abdullah-Al Nahid, and Yinan Kong, "Involvement of Machine Learning for Breast Cancer Image Classification: A Survey," *Computational and Mathematical Methods in Medicine*, vol. 2017, pp. 1-30, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[9] Nusrat Mohi ud din et al., "Breast Cancer Detection using Deep Learning: Datasets, Methods, and Challenges Ahead," *Computers in Biology and Medicine*, vol. 149, p. 106073, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Subasish Mohapatra et al., "Evaluation of Deep Learning Models for Detecting Breast Cancer using Histopathological Mammograms Images," *Sustainable Operations and Computers*, vol. 3, pp. 296-302, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] K. Shwetha et al., "Breast Cancer Detection using Deep Learning Technique," *International Journal of Engineering Research & Technology (IJERT)*, vol. 6, no. 13, pp. 1-4, 2018. [Google Scholar] [Publisher Link]

[12] Susmita Mishra, and M. Prakash, "Digital Mammogram Inferencing System using Intuitionistic Fuzzy Theory," *Computer Systems Science and Engineering*, vol. 41, no. 3, pp. 1099-1115, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Olaide N. Oyelade, and Absalom E. Ezugwu, "Characterization of Abnormalities in Breast Cancer Images using Nature-Inspired Metaheuristic Optimized Convolutional Neural Networks Model," *Concurrency and Computation Practice and Experience*, vol. 34, no. 4, p. e6629, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Han Jiao et al., "Deep Convolutional Neural Networks-Based Automatic Breast Segmentation and Mass Detection in DCE-MRI," *Computational and Mathematical Methods in Medicine*, vol. 2020, pp. 1-12, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Nonita Sharma et al., "Breast Cancer Classification using Snapshot Ensemble Deep Learning Model and T-Distributed Stochastic Neighbor Embedding," *Multimedia Tools and Applications*, vol. 82, pp. 4011-4029, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Linling Qiu et al., "Gated Graph Attention Network for Cancer Prediction," *Sensors*, vol. 21, no. 6, pp. 1-21, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] G. Rajasekaran, and C. Sunitha Ram, "Effective Breast Cancer Prediction Based on Feature Extraction, Fusion and Selection using Hybrid Methodologies," *SSRG International Journal of Electrical and Electronics Engineering*, vol. 10, no. 5, pp. 131-142, 2023. [CrossRef] [Publisher Link]

[18] Nada M. Hassan, Safwat Hamad, and Khaled Mahar, "Mammogram Breast Cancer CAD Systems for Mass Detection and Classification: A Review," *Multimedia Tools and Applications*, vol. 81, pp. 20043–20075, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Mariam Shadan et al., "Histological Categorization of Stromal Desmoplasia in Breast Cancer and Its Diagnostic and Prognostic Utility," *SSRG International Journal of Medical Science*, vol. 4, no. 6, pp. 8-11, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[20] A. Zulaiha Parveen, and T. Senthil Kumar, "The Deep Learning Methodology for Improved Breast Cancer Diagnosis in MRI," *International Journal of Computer and Organization Trends*, vol. 11, no. 3, pp. 11-14, 2021. [CrossRef] [Publisher Link]

[21] T. Rajendran et al., "A Deep Learning Based Methodological Analysis for Breast Cancer Classification," *SSRG International Journal of Electronics and Communication Engineering*, vol. 10, no. 6, pp. 52-68, 2023. [CrossRef] [Publisher Link]

[22] Farnoosh Azour, and Azzedine Boukerche, "Design Guidelines for Mammogram-Based Computer-Aided Systems using Deep Learning Techniques," *IEEE Access*, vol. 10, pp. 21701-21726, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[23] Dong Li et al., "Convolutional Feature Descriptor Selection for Mammogram Classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 3, pp. 1467-1476, 2023. [CrossRef] [Google Scholar] [Publisher Link]