E-ISSN 2581 –7957
P-ISSN 2277 – 3916

CVR Journal of Science and Technology, Volume 27, December 2024
DOI: 10.32377/cvrjst2712

Stacked Autoencoder-based ELM with Unsupervised Feature Extraction for Efficient Classification of Tumors

Dr. A. Christy Jeba Malar[1], Dr. M. Deva Priya[2]
[1]Assoc. Professor, Sri Krishna College of Technology/IT Department, Coimbatore, Tamil Nadu, India
Email: a.christyjebamalar@skct.edu.in
[2]Professor, Karpagam College of Engineering/CSE Department, Coimbatore, Tamil Nadu, India
Email: meetdevapriya@gmail.com

Abstract: **Tumors may exist in the brain, lungs, esophagus, leukemia, breast, ovary and bladder. Each tumor has its pathogenesis and categorizing them is an uphill task. Appropriate features are to be extracted, and the data of tumor should be classified either as benign or malignant. This is a critical issue as treatment of patients is solely based on this. In this paper, tumor data is classified using Stacked Extreme Learning Machine (S-ELM) and features are extracted using Statistically Controlled Activation Weight Initialization (SCAWI). It involves unsupervised clustering and is based on Neural Network (NN) framework. The proposed method offers a good decision support system for classifying tumor data. The proposed scheme is applied to data obtained from hospitals of repute and publicly available domains and tested for its outperformance. Results are compared with basic ELM and Multilayer-ELM (ML-ELM). It is seen that S-ELM with SCAWI offers better performance in contrast to diverse methods of classification.**

Index Terms: **Breast Tumor, ELM, Multilayer-ELM, Tumor Classification.**

I. INTRODUCTION

Breast cancer is seen to be common among women. Hence, timely diagnosis of the disease is of the greatest demand. As the count of patients with tumors keep increasing, it is essential to design an automated system that identifies tumors in early stages, thus saving lives. Timely detection of brain tumors plays a dominant role in improving treatment outcomes. Traditionally, breast tumors are diagnosed through biopsies, which require invasive surgery. However, the techniques of computational intelligence can aid in the identification and classification of breast tumors, offering a less invasive option.

At present, RX mammography is an extensively used screening method for breast cancer [1]. Nevertheless, it has some shortcomings. As it employs ionizing radiation, it is not suitable for young women with breasts of high-density. Identifying breast lesions is also challenging due to lacking of functional information. Ultrasound (US) is capable of detecting cancers in female patients with dense breasts as well as negative mammography. It is useful in characterizing abnormalities that are identified mammographically, evaluating tumor size as well as nodal status, and managing needle biopsy [2]. Nevertheless, its capability to detect contra or ipsi-lateral malignant lesions is limited.

The quality of images will be increased by pre-processing them using filters. Noise can be removed from edges without disturbing their edges, whereas unwanted regions can be removed by segmenting images followed by feature extraction and selection. An ideal feature set is the one that includes selective features extracted using fuzzy clustering algorithms. Tumors can be classified as benign or malignant using Artificial Neural Network (ANN) as it best suits datasets with more number of attributes and delivers a unique solution. It offers improved accuracy. It helps physicians in prompt detection and treatment of tumors, thus reducing mortality rate. Extreme Learning Machine (ELM) is a Single-layer Feed Forward Network (SLFN) with arbitrary weights. As it offers good training accuracy and circumvents local minima, it is used in disease diagnosis. It is capable of handling non-linear complex data.

In this paper, a statistically controlled weight initialized Stacked ELM (S-ELM) is proposed for multi-platform image and clinical data. It handles inter and intra-modal cancer subtypes. It divides a large ELM network into several stacked small ELMs that are linked serially. It is efficient in approximating a huge ELM network by involving less amount of memory [3].

The ELM with one hidden layer poses memory restrictions in some problems. Multilayer ELM (ML-ELM) is an extended form of ELM which supports unsupervised learning using ELM autoencoders. It eradicates the necessity of parameter tuning, facilitating improved representation learning as it includes numerous layers [4].

Cancer data represented by hidden variables are learnt using unsupervised subtractive clustering algorithm. Stacked auto-encoder framework based generative model is designed to express hidden features defined by every input. The results of hidden variables are combined to extract common features.

Clinical Features

Clinical features are based on medical records. They offer instructions or information necessary for classification. Tumors may be classified using a pattern recognition framework that is based on features and classifiers [5]. Several ML algorithms are applied to classify breast lesions based on different feature categories including dynamic, morphological and textural features [6].

Breast tumor dataset includes mammogram images. Generally, around 32 parameters are involved in classifying

Received on 23-09-2024, Revised on 23-10-2024, Accepted on 22-11-2024.

cancer [7, 8]. In the case of their values are comparatively high, it is a sign of malignancy.

- **ID:** It is a number that represents identification.
- **Membrane Diagnosis:** It represents tissue diagnosis which may take values, either malignant or benign. For varying types of cancer, it is essential to find the right tissue diagnosis if both membranes involve varying treatments.
- **Mean, Standard Error (SE) and Radius:** They represent a range between the centre and point on perimeter.
- **Radius:** Radius SE is another parameter. Radius worst represents the maximum value of centre for determined range and it is necessary to compute the distance amid centre and point as surgery is based on size. Surgery is not advisable for large tumours.
- **Texture:** The mean value denotes Standard Deviation (SD) of Gray-Scale (GS) values. SE of texture signifies the SE of determined SD for GS values. The worst value denotes the maximum mean of SD for GS values. The GS is used for determining tumor location, while SD is indispensable for finding data variation and detail of spread.
- **Perimeter:** The mean value represents the mean of core tumor, while SE of mean signifies core tumor. The worst value denotes the maximum value of core tumor inscribed on perimeter.
- **Area:** The mean, SE and worst point at related values are associated with mean areas of cancer cells.
- **Smoothness:** The mean value denotes the mean for regional disparities in range of radius, while SE of smoothness denotes SE of mean of local differences in length of radius and worst denotes the largest mean value.
- **Compactness:** The mean value represents the mean value of perimeter as well as area, SE of compactness represents SE of mean, and worst represents the highest mean of calculation.
- **Concavity:** It denotes severity of concave parts of shape, whereas mean of concave points denotes quantity of concave parts of contour. SE of concavity signifies SE of concave parts.
- **Concave points:** It denotes SE of concave parts of shape. The worst values of Concavity and concave points denote maximum value of mean.
- **Fractal Dimension (FD):** The mean value denotes the mean for approximation of coastline, FD SE is the SE of coastline approximation, whereas FD worst denotes the maximum mean value.

II. RELATED WORK

In this section, the diverse methods available for diagnosing breast tumor are detailed below.

Assiri et. al. (2020) [9] have compared hard and soft voting methods for ensemble learning. It is seen that hard voting which relies on a majority-based approach outperforms existing algorithms in the detection of breast cancer leading to enhanced accuracy in classification. Yadav & Jadhav (2022) [10] have discussed how Machine Learning (ML) serves as a powerful statistical tool for software applications, enabling them to learn from data without manual coding. By analyzing thermal scans, ML is capable of identifying areas demanding further review by healthcare professionals. Thermal imaging offers a simpler, more efficient diagnostic alternative compared to complex equipment, making it beneficial for clinics and hospitals in enhancing diagnostic accuracy.

Dewangan et. al. (2022) [11] have introduced a detection model named Back Propagation Boosting Recurrent Wienmed (BPBRW) which uses a Hybrid Krill Herd African Buffalo Optimization (HKH-ABO) mechanism to improve early-stage breast cancer diagnosis through MRI images. The model trains on breast MRI images, applies an innovative Wienmed filter for noise reduction and classifies tumors as benign or malignant. The implementation carried out in Python shows that this model achieves a higher accuracy rate with a reduced error rate, demonstrating its efficiency in early breast cancer detection. Mohamed et. al. (2022) [12] have contributed significantly by proposing a Deep Learning (DL) approach for identifying breast cancer from biopsy images using Convolutional Neural Network (CNN). The study examines the effects of different data pre-processing techniques on model performance and has introduced an ensemble learning method to integrate the best models leading to improved classification accuracy.

Dar et. al. (2022) [13] have conducted a comprehensive review of studies focusing on detection and classification of breast cancer across various imaging technologies. This work analyses existing research employing ML, DL and Deep Reinforcement Learning (DRL) for detection and classification of breast cancer. Publicly available datasets relevant to these imaging methods are surveyed to support future research and have concluded with a discussion on current challenges and opportunities in this evolving field. Abunasser et. al. (2022) [14] have developed a DL model aimed at detecting and classifying breast cancer types including benign and malignant forms. The performance of models is evaluated utilizing a dataset from Kaggle through metrics like precision, F1-score, recall and accuracy. The results indicate that DL models are highly effective for accurate detection and classification of breast cancer.

Abunasser et. al. (2023) [15] have introduced BCCNN model for breast cancer detection and classification. This model processes MRI images and incorporates effective pre-trained DL models, initially trained on ImageNet and adapted for the classification of breast cancer. The dataset sourced from Kaggle is expanded using Generative Adversarial Network (GAN) to enhance its size and diversity. Each model is evaluated using different datasets for 30 experiments to assess the effectiveness of the models. Labcharoenwongs et. al. (2023) [16] have developed an automated system for detecting, classifying and estimating the volume of breast tumors using DL techniques, specifically analyzing breast ultrasound images. They have employed various data augmentation methods like blurring and flipping for enhancing training and testing datasets. YOLOv7 architecture is utilized for tumor detection and classification yielding strong performance with increased accuracy and confidence scores, highlighting the potential of the systems to assist radiologists in improving the diagnosis of breast cancer.

Ranjbarzadeh et. al. (2023) [17] have examined segmentation methods in image processing within Computer Aided Diagnosis (CAD) systems. The study categorizes segmentation techniques into supervised, unsupervised and DL-based methods aiming at providing an overview of strengths and weaknesses of every approach to aid researchers in selecting the most appropriate technique for specific applications. Ali & Agrawal (2024) [18] have proposed a hybrid model for automatic image segmentation that combines VGG19 architecture with ResNet101 through CNN. The classifier is optimized using Bayesian methods and cross-validation is employed to identify the best parameters for optimized classifiers. Diverse Transfer Learning (TL) models are used for determining the most effective one for detecting brain abnormalities using NNs. The proposed stacked classifier model leverages the advanced technologies and demonstrates better results in contrast to the existing models.

III. PROPOSED METHODOLOGY

In case of breast tumor classification, standard MRI examination and scanning protocols are not readily available. Another challenge is that datasets including breast MRI images are not extensively available. The next challenge deals with the type of features extracted. Combination of features may confuse the clinician making it difficult to understand and interpret. Further, certain features cannot be applied directly. Hence, an ideal set of features proficient in an efficient classification of tumors should be selected.

Accurate selection of features simplifies the job of a classifier. Performance can be improved by using sophisticated classifiers like ANN, Support Vector Machine (SVM), etc., by involving several features. Unintelligible set of features can be combined with non-linear classifier.

A. Noise Removal using Gaussian Filtering

Gaussian filtering [20] is involved in blurring images and removing noise. Gaussian function is given by,

$$G(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \frac{e^{-x^2}}{2\sigma^2} \tag{1}$$

Where,

σ - Standard Deviation of distribution

The mean is initialized to ‘0’.

These filters involve 2D Gaussian distribution which is a point-spread function. It convolves distribution function with image. Discrete approximation is to be produced for Gaussian function. As Gaussian distribution is non-zero, theoretically it demands an infinitely huge convolution kernel.

B. Classification

Target class for every case in data is to be predicted accurately. Classification models are verified by comparing predicted and known target values in test data. Historical data for classification is typically split into training and testing sets. Once features are extracted, the image is classified into malignant or benign ones by using the ELM.

A systematic assessment of various pattern recognition approaches are involved in classifying breast lesions using morphological, dynamic as well as textural features. Classification is performed using ANN, SVM, Linear Discriminant Analysis (LDA) and ELM-based approaches.

Extreme Learning Machines (ELM)

Huang et. al (2011) [19] have proposed ELM for SLFN. It supports increased learning speed and offers good generalization. Hidden nodes are chosen randomly in ELM. They outperform SLFN and SVM. ELM systematically assigns weights to output nodes and supports faster training. It yields better generalization but finds difficulty in handling huge amount of data.

Accurate selection of classifier is challenging. Several classifiers are available in the literature.

- Linear classifiers are efficient but combined feature interpretation is an uphill task.
- Non-linear classifiers like ANN and SVM yield good performance, but interpretation of classification hyperplanes are demanding.
- Tree-based classifiers grounded on simple threshold-based rules face the issue of overfitting.

Features can be combined as classifiers are not appropriate for all features.

Stacked Extreme Learning Machine (S-ELM)

Algorithm for S-ELMs learning network is detailed below.

Given,
$\{(x_i, t_i) | x_i \in R_n, t_i \in R_m, i = 1, \ldots, n\}$ - A large training dataset
g(x) - Activation function
N - Quantity of hidden nodes in every layer ‘L’
L^1 - Quantity of target combined nodes
C - Regularization co-efficient
S - Total amount of layers

Step 1: Implement ELM algorithm in layer 1:

- Arbitrarily produce hidden layer factors, bias ‘b_i’ and weight ‘w_i’, where i=1,..., L
- Determine output matrix ‘H’ in hidden layer
- Determine output weight ‘β’

$$\beta = \left({}^{I}/_{C} + H^T H\right)^{-1} . H^T H \tag{5}$$

Step 2: Reduce PCA dimension of ‘β^T’

- Reduce dimension of ‘β^T’ from ‘L’ to ‘L^1’
- Record reordered Eigen vectors as $\sim V \in R^{L,L^1}$

Step 3: Learning step for layer 2→S:

for p=1 to S - 1

- Arbitrarily produce $(L - L^1)$ hidden nodes and determine conforming ‘H_{new}’
- Determine joint significant nodes ‘H^1’

$$H^1 = H^{\sim} . V \tag{6}$$

- Compute output matrix of hidden layer

$$H = [H^1, H_{new}] \tag{7}$$

- Repeat step 1(c) and step 2. Step 2 is not necessary in the last layer ‘S’ end for

Step 4:

- Compute 'F_{out}' and give it as output of S-ELMs learning network in last layer
  $$F_{out} = H.\beta \tag{8}$$

The inefficiency of ELM may be due to setting of random weights. Statistically Controlled Activation Weight Initialization (SCAWI) technique is proposed for initializing parameters.

IV. STATISTICALLY CONTROLLED ACTIVATION WEIGHT INITIALIZATION (SCAWI)

In SCAWI [21], networks are initialized with weights and impact of random initialization is reduced. Neurons in both saturated and paralyzed states are taken into consideration. A neuron is said to be saturated when output is greater than threshold, and is in paralyzed state when difference amid target as well as output are equal to or more than threshold. This demands training of the network. Paralyzed Neuron Percentage (PNP) is involved in computing number of times the neuron moves to paralyzed state. It is based on the quantity of neurons which are paralyzed and total amount of neurons.

$$PNP = \left(\frac{1}{\beta}\right).P \tag{9}$$

where,

$\left(\frac{1}{\beta}\right)$ - Probability that one output is unsuitable

P - Neuron's saturation probability

Initial weights are computed using optimal value of PNP.

$$W = \left(\frac{R.a_0}{\sqrt{\partial}}\right).r_{ji} \tag{10}$$

where,

A - Neuron activation

∂ - Expectation of random variable

R - Random number

Scale factor (R) is given by,

$$R = \frac{\sqrt{3}}{\Delta\,(PNP)} \tag{11}$$

where,

$$\Delta = erf^{-1}\left(\frac{1}{2} - \left(\frac{\beta}{2}\right).PNP\right) \tag{12}$$

In this work, segmentation and classification of tumor data using subtractive clustering and S-ELM classifier is propounded. The ELM is designed with double activation functions to have reduced influence of arbitrary initialization and offer better accuracy. The ELM-based classifier is proposed for analyzing and classifying data.

The SCAWI propounded for initializing hidden node factors of propounded ELM offers improved accuracy and training time. Average output from hidden node is obtained after activation function transformations. Activation functions include tangent (tanh) as well as hyperbolic cosine functions. They deal with non-linear data. In the proposed approach, the problem of random initialization of hidden node parameters in ELM is resolved.

Non-linear data is handled efficiently by using dual activation functions and output is got by averaging activation functions. In the proposed methodology, the ELM initializes weights as well as bias parameters by using SCAWI. The weight between input and hidden neurons are initialized based on paralyzed neuron percentage. Weight and bias are randomly chosen and fine-tuned by using Eq. (12) iteratively. In the hidden layer, following operations are performed.

- Calculate weight changes using ELM
  $$\Delta W_i = \eta\left(\pi\, h_{ji}\right) x_k \text{ for } j \neq 1 \text{ to } m \tag{13}$$
- Update weight
  $$W_i = \Delta W_i + W_i \tag{14}$$
- Calculate weight after including momentum term '∞'
  $$W_i = W_i + \infty\Delta W_i \tag{15}$$

In the hidden layer, inverse hyperbolic cosine and tangent functions are involved and average output is computed. Inverse hyperbolic cosine function is given by,

$$ar\cosh z = \ln(z + \sqrt{z^2 - 1}) \tag{16}$$

Inverse hyperbolic tangent function is determined as follows.

$$ar\tanh z = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right) \tag{17}$$

Combined activation functions aids in handling non-linearity. S-ELM uses ELM-Auto Encoder (ELM-AE) as building blocks in every layer to assist learning. It regenerates input signal with linear or non-linear activation functions. Input data is mapped to L-dimensional feature space of ELM and output is obtained.

$$F_L = \sum_{i=0}^{n} b_i\,(x) = (h(x).b) \tag{18}$$

Output weight matrix amid hidden and output nodes is given by,

$$b = \{[b_1, \ldots . b_l]\}^T \tag{19}$$

Output of hidden node for input 'x' is determined as follows.

$$h(x) = [g_1(x), \ldots\ldots g_L(x)] \tag{20}$$

$g_i(x)$ - Output of i^{th} hidden node

For 'n' training samples $\{(x_{i,}\, t_i)\}$, i=1, ELM is involved in resolving the ensuing learning problem.

$$H_b = T \tag{21}$$

Target labels are represented as shown below.

$$T = [t_1, \ldots t_n] \tag{22}$$

$$H = [h^T\,(x_1), \ldots., h^T\,(x_n)]^T \tag{23}$$

Output weight (b) is determined as follows.

$$b = H^{\dagger}.T \tag{24}$$

where,

'H†'- Moore-Penrose generalized inverse of matrix 'H'

Regularization can be done to enhance ELM performance.

$$B = \left(\left(\frac{I}{C}\right) + H^T.H\right)^{-1}.H^T.H \tag{25}$$

Unsupervised learning is performed using ELM. Orthogonalization binds parameters closer to each other. The performance of ELM-A is increased by choosing weights and biases orthogonally. As shown in Johnson-Lindenstrauss lemma, input is projected to different or equal dimension space and computed as shown below.

$$h = g(a.x + m) \tag{26}$$

$$a^T.a = I, m^T.m = I \tag{27}$$

$$a = [a_{1,\ldots\ldots}a_T] \tag{28}$$

Eq. (28) represents orthogonal random weights.

Orthogonal random biases amid input as well as hidden nodes is determined by,

$$m = [m_{1,\ldots\ldots}m_L] \tag{29}$$

Transformation from feature space to input is learnt by using output weight of ELM-AE. For representations of sparse and compressed ELM-AE, weights (outputs) are determined as shown below.

$$B = \left(\left(\frac{I}{C}\right) + H^T.H\right)^{-1}.H^T.X \tag{30}$$

The hidden layer outputs of ELM-AE are determined as shown below.

$$H = [H_1, \ldots.H_n] \tag{31}$$

Input and output data are given by,

$$X = [x_1, \ldots.x_n] \tag{32}$$

V. Results And Discussion

Clinical data and MRI images of different types of breast tumor are gathered from domains which are publicly available. The propounded tumor detection scheme is applied to Breast Cancer Wisconsin (Diagnostic) [22] with 569 instances and 30 features. It is seen that the proposed scheme offers better sensitivity, specificity and accuracy in contrast to other methods taken under consideration.

The SCAWI is used for random initialization of parameters as it is the default option in the ANN. The proposed scheme is implemented using MATLAB. True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) are determined based on which sensitivity, specificity, and accuracy are calculated.

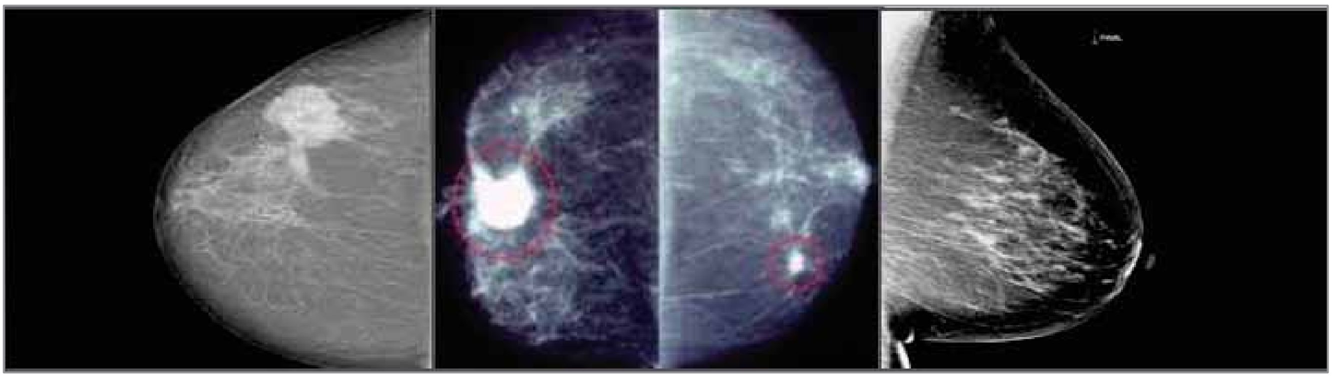

Figure 1. MRI Images of Breast

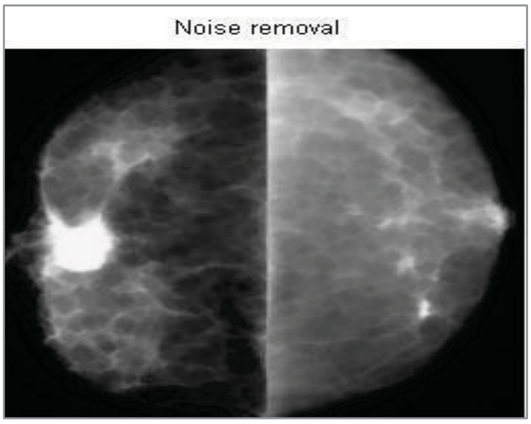

Figure 2. MRI Images of Breast after Noise Removal

Performance is evaluated by varying the hidden neurons from 50 to 200 in steps of 25. Performance is improved by combining both dynamic and morphological features.

The proposed method is tested with clinical data and MRI images of breast tumors. Figure 1 shows MRI images of breast tumors. Figure 2 and Figure 3 show MRI images after noise removal and segmentation. Figure 4 shows the images after clustering.

Figure 5 and Figure 6 show the sensitivity and specificity in terms of dynamic, morphological, textural, combined dynamic, and morphological and combined dynamic, textural and morphological features for the data of breast cancer using the SCAWI-ELM. Both Sensitivity and Specificity are high for Dynamic + Morphological features.

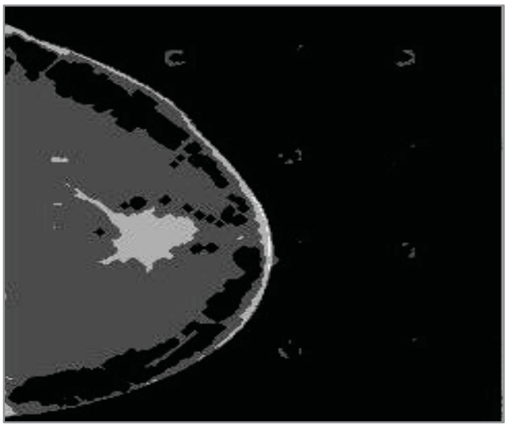

Figure 3. MRI Images of Breast after Segmentation

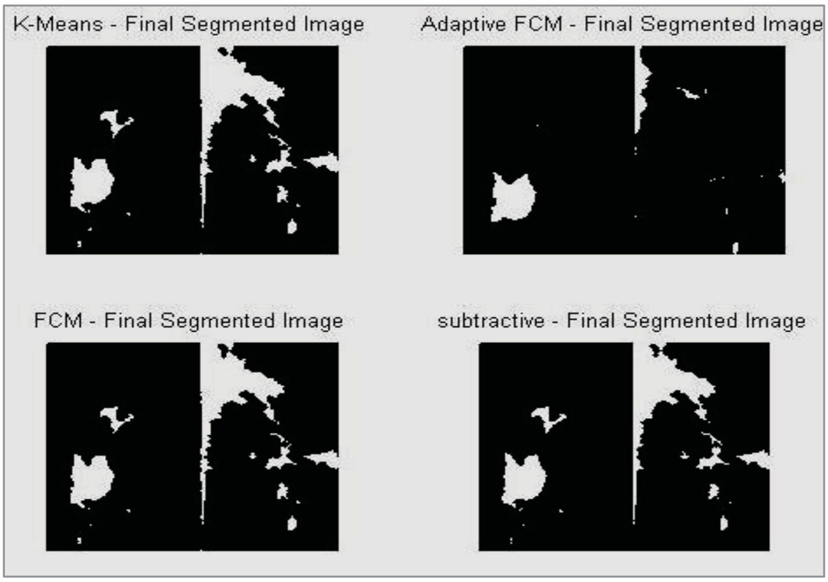

Figure 4. MRI Images of Breast after Clustering

Figure 7 to Figure 9 show the performance of classifiers. The number of statistical samples taken by classifiers is larger, thus yielding accurate analysis. Furthermore, the classifiers perform well when dynamic and morphological features are used simultaneously. From the results, it is evident that S-ELM combined with SCAWI yields better performance when compared to S-ELM in terms of accuracy, training, and testing time. The propounded scheme is compared with basic ELM and ML-ELM. It is seen that ML-ELM yields better classification results in contrast to the ELM. ML-based algorithms depend on training data and features extracted from data. Sensitivity and specificity cannot be directly compared. Based on the results obtained, average sensitivity and specificity values are computed for every classifier.

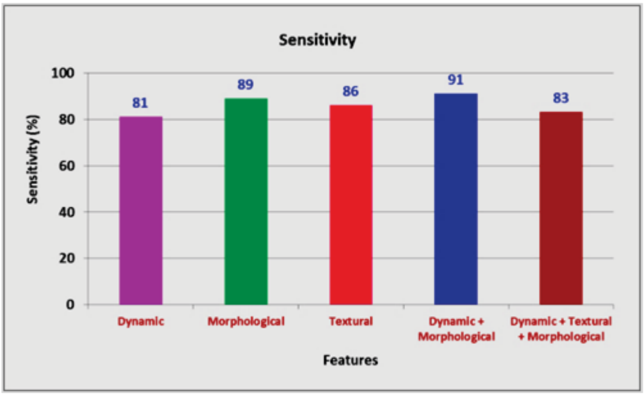

Figure 5. Sensitivity for Diverse Features

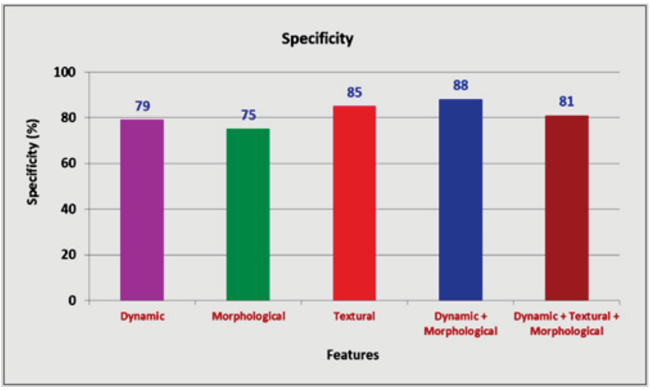

Figure 6. Specificity for Diverse Features

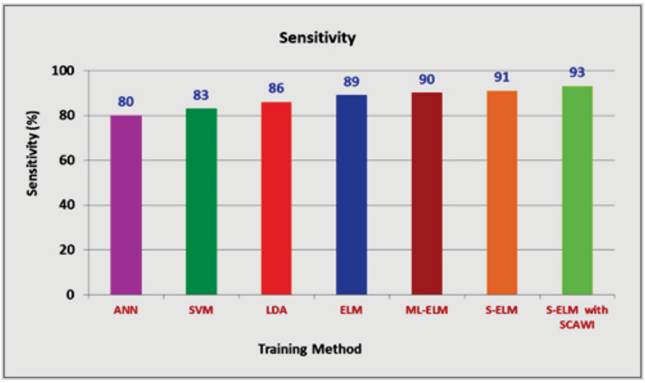

Figure 7. Sensitivity offered by Diverse Training Methods

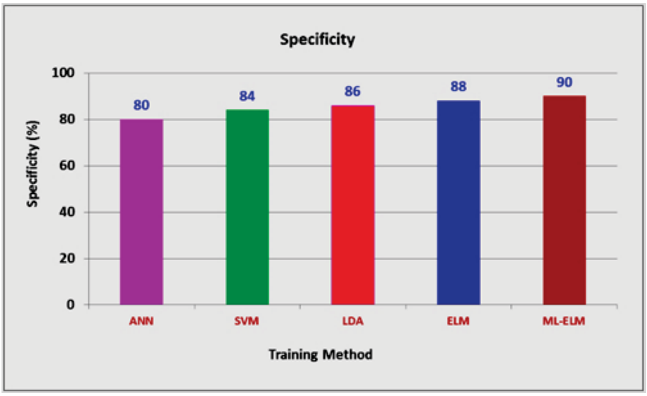

Figure 8. Specificity offered by Diverse Training Methods

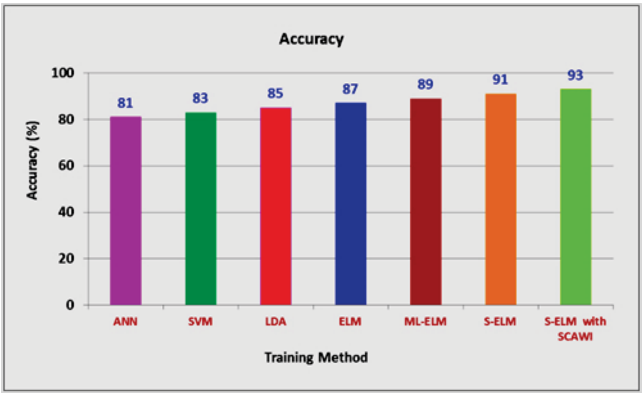

Figure 9. Accuracy offered by Diverse Training Methods

VI. CONCLUSION

In this paper, breast tumors are classified using Subtractive clustering and S-ELM classifier. The image is segmented through clustering and classified by using the ELM classifier. The SCAWI is involved in setting the initial parameters. The proposed scheme is assessed in terms of sensitivity, specificity, and accuracy. From the results, it is evident that S-ELM with SCAWI yields better results in contrast to ML-ELM and S-ELM. In the future, the proposed scheme can be applied to diverse datasets and performance can be analyzed.

REFERENCES

[1] Saslow, D., Boetes, C., Burke, W., Harms, S., Leach, M. O., Lehman, C. D., ... & American Cancer Society Breast Cancer Advisory Group. (2007). American Cancer Society guidelines for breast screening with MRI as an adjunct to mammography. CA: a cancer journal for clinicians, 57(2), 75-89.

[2] Corsetti, V., Houssami, N., Ghirardi, M., Ferrari, A., Speziani, M., Bellarosa, S., ... & Ciatto, S. (2011). Evidence of the effect of adjunct ultrasound screening in women with mammography-negative dense breasts: interval breast cancers at 1 year follow-up. European journal of cancer, 47(7), 1021-1026.

[3] Zhou, H., Huang, G. B., Lin, Z., Wang, H., & Soh, Y. C. (2014). Stacked extreme learning machines. IEEE transactions on cybernetics, 45(9), 2013-2025.

[4] Kaur, R., Roul, R. K., & Batra, S. (2023). Multilayer extreme learning machine: a systematic review. Multimedia Tools and Applications, 82(26), 40269-40307.

[5] Sandhu, D. S., Sandhu, S., Karwasra, R. K., & Marwah, S. (2010). Profile of breast cancer patients at a tertiary care hospital in north India. Indian journal of cancer, 47(1), 16-22.

[6] Huang, G. B., Wang, D. H., & Lan, Y. (2011). Extreme learning machines: a survey. International journal of machine learning and cybernetics, 2, 107-122.

[7] Aličković, E., & Subasi, A. (2017). Breast cancer diagnosis using GA feature selection and Rotation Forest. Neural Computing and applications, 28, 753-763.

[8] Mangasarian, O. L., Street, W. N., & Wolberg, W. H. (1995). Breast cancer diagnosis and prognosis via linear programming. Operations research, 43(4), 570-577.

[9] Assiri, A. S., Nazir, S., & Velastin, S. A. (2020). Breast tumor classification using an ensemble machine learning method. Journal of Imaging, 6(6), 39.

[10] Yadav, S. S., & Jadhav, S. M. (2022). Thermal infrared imaging based breast cancer diagnosis using machine learning techniques. Multimedia Tools and Applications, 1-19.

[11] Dewangan, K. K., Dewangan, D. K., Sahu, S. P., & Janghel, R. (2022). Breast cancer diagnosis in an early stage using novel deep learning with hybrid optimization technique. Multimedia Tools and Applications, 81(10), 13935-13960.

[12] Mohamed, A., Amer, E., Eldin, N., Hossam, M., Elmasry, N., & Adnan, G. T. (2022). The impact of data processing and ensemble on breast cancer detection using deep learning. Journal of Computing and Communication, 1(1), 27-37.

[13] Dar, R. A., Rasool, M., & Assad, A. (2022). Breast cancer detection using deep learning: Datasets, methods, and challenges ahead. Computers in biology and medicine, 149, 106073.

[14] Abunasser, B. S., AL-Hiealy, M. R. J., Zaqout, I. S., & Abu-Naser, S. S. (2022). Breast cancer detection and classification using deep learning Xception algorithm. International Journal of Advanced Computer Science and Applications, 13(7).

[15] Abunasser, B. S., Al-Hiealy, M. R. J., Zaqout, I. S., & Abu-Naser, S. S. (2023). Convolution neural network for breast cancer detection and classification using deep learning. Asian Pacific journal of cancer prevention: APJCP, 24(2), 531.

[16] Labcharoenwongs, P., Vongmansup, S., Chunhapran, O., Noolek, D., & Yampaka, T. (2023). An automatic breast tumor detection and classification including automatic tumor volume estimation using deep learning technique. Asian Pacific Journal of Cancer Prevention: APJCP, 24(3), 1081.

[17] Ranjbarzadeh, R., Dorosti, S., Ghoushchi, S. J., Caputo, A., Tirkolaee, E. B., Ali, S. S., ... & Bendechache, M. (2023). Breast tumor localization and segmentation using machine learning techniques: Overview of datasets, findings, and methods. Computers in Biology and Medicine, 152, 106443.

[18] Ali, S., & Agrawal, J. (2024). Automated segmentation of brain tumour images using deep learning-based model VGG19 and ResNet 101. Multimedia Tools and Applications, 83(11), 33351-33370.

[19] Huang, G. B., Wang, D. H., & Lan, Y. (2011). Extreme learning machines: a survey. International journal of machine learning and cybernetics, 2, 107-122.

[20] Ito, K., & Xiong, K. (2000). Gaussian filters for nonlinear filtering problems. IEEE transactions on automatic control, 45(5), 910-927.

[21] Drago, G. P., & Ridella, S. (1992). Statistically controlled activation weight initialization (SCAWI). IEEE Transactions on Neural Networks, 3(4), 627-631.

[22] https://www.kaggle.com/datasets/uciml/breast-cancer-wisconsin-data