

Automatika

Journal for Control, Measurement, Electronics, Computing and Communications



ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/taut20

A modified recurrent neural network (MRNN) model for and breast cancer classification system

A. Abdul Hayum, J. Jaya, B. Paulchamy & R. Sivakumar

To cite this article: A. Abdul Hayum, J. Jaya, B. Paulchamy & R. Sivakumar (2023) A modified recurrent neural network (MRNN) model for and breast cancer classification system, *Automatika*, 64:4, 1193-1203, DOI: [10.1080/00051144.2023.2253064](https://doi.org/10.1080/00051144.2023.2253064)

To link to this article: <https://doi.org/10.1080/00051144.2023.2253064>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 10 Sep 2023.



Submit your article to this journal [↗](#)



Article views: 425



View related articles [↗](#)



View Crossmark data [↗](#)



A modified recurrent neural network (MRNN) model for and breast cancer classification system

A. Abdul Hayum^a, J. Jaya^b, B. Paulchamy^a and R. Sivakumar^c

^aDepartment of Electronics and Communication Engineering, Hindusthan Institute of Technology, Coimbatore, India; ^bDepartment of Electronics and Communication Engineering, Hindusthan College of Engineering and Technology, Coimbatore, India; ^cSchool of Electronics Engineering, Vellore Institute of Technology, Vellore, India

ABSTRACT

Breast cancer is most dangerous cancer among women. Image processing techniques are used for Breast cancer detection. A Block-based cross diagonal texture matrix (BCDTM) method is used first to extract Haralick's features from each mammography ROI. Likewise, wrapper method is utilized to choose the crucial features from the condensed feature vector. There are lot of factors that affects the quality of the images such as salt or pepper noise. As a result, this is less precise and more prone to mistakes because of human. In order to address the problems, input breast image is first pre-processed via median filtering to reduce noise. ROI segmentation is done using weighted K means clustering. Feature extraction, texture and form descriptors based on Centroid Distance Functions (CDF) and BCDTM are used. Kernel Principal Component Analysis (KPCA) is used as dimensionality reduction on the extracted features. Improved Cuckoo Search Optimization (ICSO) is used to compute relevant feature selection. Modified Recurrent Neural Network (MRNN) is utilized to classify breast cancer into benign and malignant. Results show that the suggested model achieved highest accuracy, precision and recall values compared with other state-of-the-art approaches.

ARTICLE HISTORY

Received 20 June 2023
Accepted 23 August 2023

KEYWORDS

Breast cancer; medical imaging; benign; detection; malignant and Recurrent Neural Network

1. Introduction

A tumour is an abnormal growth of body tissue. It is an uncontrolled cell division that sets off cancer. A tumour might be cancerous or benign. Malignant tumour expands quickly and damages the tissues in its vicinity by invading them [1]. Breast lump, breast shape and size alterations, variations in the colour of the breast skin, breast pains and gene changes are certain signs of breast cancer. Over 8% of women could be affected with breast cancer, which making it the second-leading cause of mortality of women worldwide. Over 500,000 women would die each year and almost 1,000,000 women would receive a new breast cancer diagnosis, according to a World Health Organization report [2,3].

According to the world health organization (WHO) records, as the environment continues to deteriorate, the occurrence of this disease will rise. There were 40,480 fatalities and 182,460 recently diagnosed cases reported in the US in 2021. Earlier detection is essential to lowering the death rate (by 40% or more) because the origins of breast cancer are still unknown. Cancers can be treated more effectively the earlier they are discovered. Early detection, though, necessitates a precise and trustworthy diagnostic that can tell benign from malignant tumours apart. Low False Positive (FP) and False Negative (FN) rates are indicators of a successful detection strategy [4,5].

The most effective tool for identifying and diagnosing breast cancer is thought to be mammography. Despite an increase in breast cancer incidence over the previous 10 years, mortality rates have decreased for women of all ages. This encouraging trend in mortality decrease could be attributed to advancements in the medication of cancer and the widespread use of mammography screening. The fact that skilled radiologists sometimes miss a sizable number of abnormalities is well known. In addition, a significant portion of mammographic abnormalities revealed by biopsy turn out to be benign. Conventional techniques of disease monitoring and diagnosis rely on a human observer identifying specific signal qualities as present. Over the past 10 years, several computer-aided-diagnosis (CAD) methods have emerged to help intensive care unit keep track the patients more effectively [6]. For instance, a recent study employed feature selection and Bayes classifiers in combination to identify breast cancer, whereas a different study used an artificial neural network and a breast ultrasound image to predict breast cancer. These efforts do not yield results with adequate accuracy.

Recent research aimed at producing a powerful CAD model that can categorize digital mammograms as benign, malignant, in order to overcome the above-mentioned issues. In order to extract Haralick's features

CONTACT J. Jaya ✉ drjjaya2011@gmail.com 📠 Department of Electronics and Communication Engineering, Hindusthan College of Engineering and Technology, Coimbatore, India

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

from each mammography ROI, a block-based CDTM method was used first. The KPCA method is then used to decrease the size of the generated feature vector. The grasshopper optimization concept utilized to obtain the ideal values of the kernel extreme learning machine (KELM) parameters. The key attributes from the shortened feature vector are also chosen using the proposed wrapper strategy. However, there are numerous elements that aim to degrade image quality, like salt or pepper noise. As a result, this is less precise and more prone to mistakes because of numerous human factors.

In order to address the problems, the breast image is first pre-processed through median filtering to reduce noise. After the noise has been eliminated, ROI segmentation is carried out using weighted K means clustering. For feature extraction, shape descriptor approaches based on Centroid Distance Functions (CDF) and block-based Cross Diagonal Texture Matrix (CDTM) are used. Extracted features are provided for dimensionality reduction with Kernel Principal Component Analysis in this study (KPCA). a relevant feature selection phase that is calculated using Improved Cuckoo Search Optimization (ICSO). At the end, the Modified Recurrent Neural Network (MRNN) is utilized to categorize breast cancer into benign and malignant forms, and the parameters of the MRNN are adjusted utilizing optimization in this study for feature selection.

Major contribution to the research: Developing a more accurate framework for identifying and categorizing breast cancer. MRNN algorithm is utilized to categorize breast cancer into benign and malignant forms.

2. Literature review

Francis et al [7] created a revolutionary technique. Rotational thermography is used to get over the drawbacks of traditional breast thermography. From the standpoint of a cold challenge, this technique promises for automated diagnosis of breast anomalies. Prior to and following the administration of the cold trial, rotating thermogram series are used to extract texture features in the spatial domain. With the use of these attributes and a support vector machine, normal and cancerous breasts may be automatically classified with an accuracy of 83.3%. Analysis of principal components has been utilized to reduce the number of features. It has been investigated whether this method can find the abnormality as a unique endeavour. The study research shows that rotational thermography has a lot of potential as a breast cancer screening method.

Carvalho et al [8] devised a system that categorizes histological breast images in four groups: normal tissue, benign lesion, *in situ* carcinoma and invasive carcinoma. This method uses phylogenetic diversity indices to define images. According to the literature at hand, the classifiers utilized were Multilayer Perceptron, Random

Forest, Support Vector Machine and XGBoost. In addition, the material image recovery was carried out to validate the categorizing outcomes and offer a ordering for collections of unlabelled images. The outcomes were remarkably reliable and supported the design of a CADx system to support specialists in sizable medical facilities.

Vijayarajeswari et al [9] provided mammograms classification using Hough transform features. Two dimensions are involved in the Hough transform. Masses and miniature size classification are the two most critical threat signs, and their automated identification is important for the earliest detection of breast cancer. Computerized mass location and layout is quite challenging since masses are frequently unclear from the surrounding parenchymal. The methodologies for feature extraction and categorization are discussed in this research. Here, the SVM is utilized to categorize the mammography picture after the Hough transform has detected its features. SVM classifier use increases classification accuracy. 95 mammograms images were gathered and categorized using SVM to test this strategy. The outcome demonstrates that the suggested strategy successfully categorizes the problematic classes of mammograms.

Hassanien et al [10] described a hybrid strategy syndicates the benefits of fuzzy sets, multilayer perceptron neural networks (MLPNN) and ant-based clustering classifiers with statistical-based feature extraction. The ability and accuracy of a hybridization system to categorize images into two results: malignant or benign. The technique used by the hybrid system to improve the contrast of the input photos is based on type-II fuzzy sets. The adaptive ant-based clustering algorithm, a modern take on the outdated ant-based clustering technique, is used to classify target items using an enhanced methodology. After that, more than 20 statistically based features are retrieved and standardized. To establish whether the tumour is benign or malignant, an MLPNN classifier was used to assess the lesion descriptors capacity to discriminate between various regions of interest. Tests on various breast MRI images will be presented to measure the efficiency of the presented technique. The outcomes demonstrate the advantage of adaptive ant-based segmentation over traditional method and the utilized hybrid techniques total accuracy demonstrates the high efficiency of the hybrid system.

Muduli et al [11] suggested a revised CAD model for categorizing breast masses as abnormal, normal, malignant and benign. The region of interest mammography images to extract the features using lifting wavelet transform (LWT). The scale of the feature vectors is instead decreased by combining LDA and PCA techniques. In conclusion, a moth flame optimization strategy combined with an Extreme Learning Machine (ELM) is used to complete the classification.

The MFO-ELM technique uses MFO to improve ELM hidden node settings. To enhance the generalization performance of the model, five-fold stratified cross-validation is also executed. On the DDSM and MIAS-standard datasets, the suggested model is assessed. The results of the experiment show that the proposed CAD model obtains accuracy rates of 99.76% and 98.80% when used to the DDSM dataset. This model also gets the best result for the MIAS dataset. The experimental products demonstrate that the suggested model, with a noticeably smaller number of characteristics, outperforms other cutting-edge models in regard to classification accuracy.

3. Breast cancer classification model

The proposed breast cancer recognition technique is discussed in this section. The process has six steps as follows: In the first step, the noise is eliminated using median filtering. In the second step, ROI is segmented using weighted K means clustering. In the third step, features are extracted using Centroid Distance Function (CDF) and Block-based CDTM. In the fourth step, dimensionality is reduced using KPCA. The fifth step involves feature selection via ICSO. The sixth step involves the Modified Recurrent Neural Network (MRNN), which is used to categorize breast cancer into malignant and benign forms. The suggested model's overall architecture is described in Figure 1.

3.1. Input

In this work, two publicly accessible datasets namely, DDSM and the Mammographic Image Analysis Society (MIAS) are employed. Images and labels for

mammography scans make up the MIAS data. (<https://www.kaggle.com/code/aditi02/breastcancer-cnn-mias>). The original 350 photos can be found in the MIAS collection of digital mammograms. Around 1200 photos make up the DDSM dataset (<https://www.kaggle.com/datasets/cheddad/miniddsm2>). The entire dataset is split into 70% testing and 30% training.

3.2. Noise removal using median filter

To filter out the noise in the input, breast images are required. This technique removes noise using a median filtering technique. Median filtering technique effectively reduces interference pulses while preserving signal characteristics [12]. It is frequently employed as a pre-processing method. When the signal length is N , the filtering windows length is described as n . The function provides the filters output:

$$med(a_i) = \left\{ \begin{array}{l} a_{k+1} \quad n = 2k + 1 (odd) \\ \frac{[a_k + a_{k+1}]}{2} \quad n = 2k (even) \end{array} \right\} \quad (1)$$

where $a_1, a_2, a_3 \dots a_k$ are the observed data and a_k is the k -th maximum. The median filters ability to remove pulse noise while preserving local features is what makes it stand out.

3.3. Region of interest (ROI) extraction using weighted K means clustering

After the noise has been removed, the images must be extracted for ROI. This work uses weighted K means clustering to compute ROI extraction. The boundaries of the items will be defined clearly and segmented, with the objective of image segmentation to separate the area

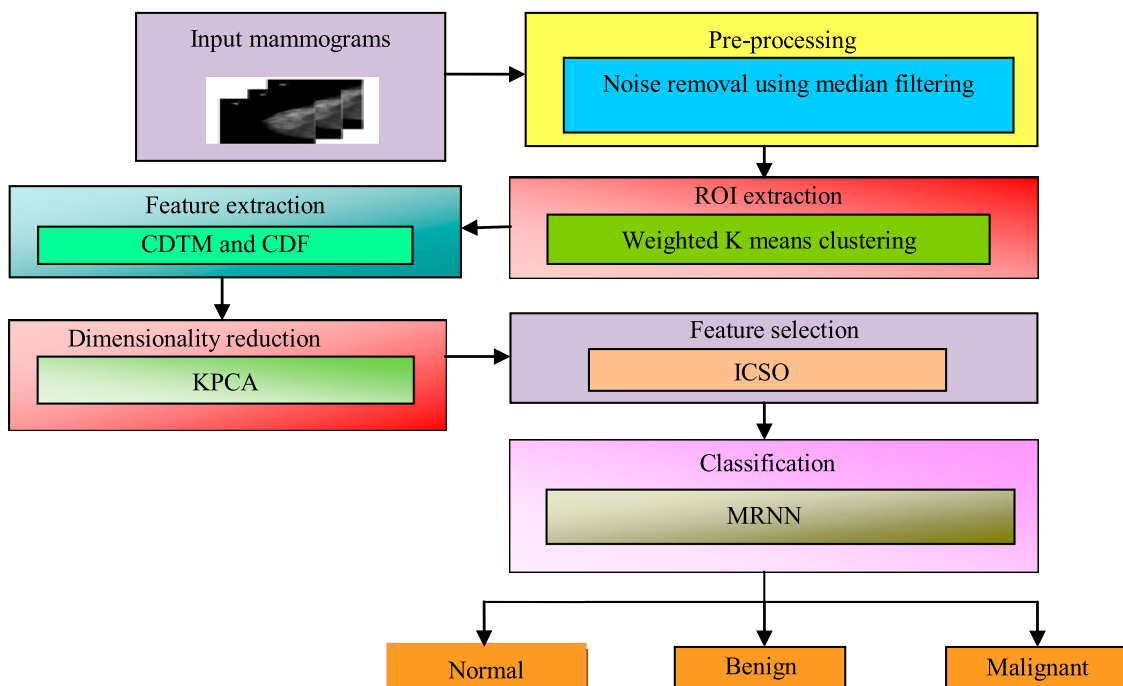


Figure 1. Overall architecture of the proposed model.

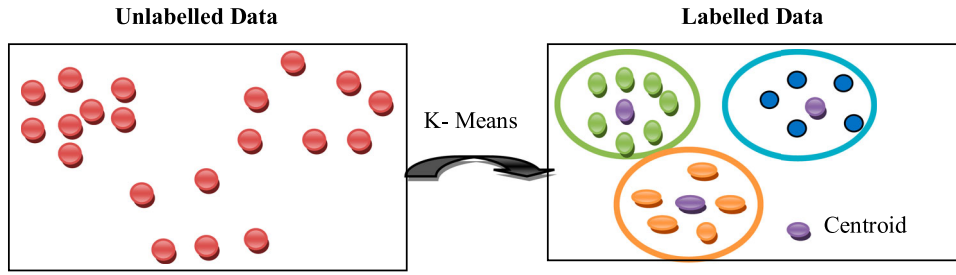


Figure 2. K means clustering.

of the object from the background to facilitate analysis. As seen in Figure 2, the segmentation method known as “k-means clustering” divides a set of data into k groups. The following steps are iterated by the algorithm. Calculate the average for each cluster or group. Then examine the distance between each data point and the cluster centre [13]. Next, assign each data point to the nearest cluster based on the calculated distance. At completion of the allocation, the cluster centre is calculated, and the distance vector is calculated using that centre. Consider the case when k clusters are needed to separate an $x \times y$ picture, with C_k acting as the cluster centres. The distance between each data point in the image and the cluster centre may be calculated using Equation (2).

$$d = p(x, y) - C_k \quad (2)$$

Here p refers pixel and d refers Euclidean distance. Assign each data point into a cluster depending on the closest distance after computing distance. Once allocation is complete, use Equation (3) to recalculate the new centre for each cluster:

$$C_k = \frac{1}{k} \sum_{x \in c_k} \sum_{y \in c_k} p(x, y) \quad (3)$$

The typical K-means clustering algorithm cannot accurately capture the interclass structure of data set because it does not consider the significance of each characteristic attribute in the data set when determining the first clustering centre point. As a result, a technique for weighted K-means clustering is suggested.

Weighted K means clustering

In the suggested weighted K-means clustering, the clustering technique is replaced with attribute synthesis weights β_j . The ratio of the i -th image data at the j -th pixel is derived for that first need:

$$P_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad (4)$$

According to the following Equation (5), E_j of the j th pixel entropy is equal to:

$$E_j = -k \cdot \sum_{i=1}^m p_{ij} \cdot \ln p_{ij}, \text{ where, } k = \frac{1}{\ln m} \quad (5)$$

Input: Dataset

Output: Segmented foreground image

Step 1: Established the number of clusters, k and its centre.

Step 2: Estimate the weighted Euclidean distance among the cluster centre and each pixel in the image shown as:

$$d(m, n) = \sqrt{\sum_{i=1}^m \beta_j (x_{mj} - x_{nj})^2} \quad (8)$$

Step 3: Depending on the closest distance, group all of the data points into clusters.

Step 4: After allotment, evaluated the novel centre for every cluster expressed as:

$$C_k = \frac{1}{k} \sum_{x \in c_k} \sum_{y \in c_k} p(x, y) \quad (9)$$

Step 5: Redo the processtill it converges.

Step 6: Assigning data points in the appropriate manner and rebuild the image.

According to the following equation, the entropy weight W_j of the j th pixel is equal to:

$$W_j = \frac{1 - E_j}{\sum_{j=1}^n (1 - E_j)} \quad (6)$$

It is decided what each pixel overall weight β_j . A preference weight a_j ($j = 1, 2 \dots n$) can be adjusted for each pixel because the significance of each pixel in each sample differs. A complete weight β_j of each pixel can be calculated by adding the entropy weight w_i and artificial preference weight α_i :

$$\beta_j = \frac{\alpha_i \cdot w_i}{\sum_{i=1}^n \alpha_i \cdot w_i} \quad (7)$$

The weighted k-means clustering technique algorithm in the following equation:

3.4. Feature extraction using block-based CDTM and centroid distance function (CDF)

The characteristics should be extracted once the ROI has been extracted. This work uses the Centroid Distance Function (CDF) and Block-based CDTM to extract texture and shape data.

CDTM

The GLCM and Texture Spectrum (TS) attributes are part of the cross-diagonal texture matrix. The structural relationship between a pixel and its neighbours is represented by the GLCM method of TS at the distance and given angle. The TS method of textural research gives

textural data as all 8 of the centre pixel surrounding pixels.

Centroid distance function (CDF)

The second contour model that will be utilized to extract form features is the centroid distance function. It is defined as the distance between its centroid (x_c, y_c) , and its boundary points (Equation (10))

$$R(S) = \sqrt{(x_s - x_c)^2 + (y_s - y_c)^2} \quad (10)$$

3.5. Dimensionality reduction using Kernel Principle Component Analysis

The proficiency of the algorithms can be enhanced by transforming the high-dimensional input set effectively. Overall, there is a link amongst the reduced dimensions and the input data allows constricting transforming the whole input space to a smaller subspace. The size of the high-dimensional feature space F is reduced using KPCA.

3.6. Feature selection using ICSO

This work uses a feature selection model based on the ICSO method to reduce time consumption and boost accuracy. The input database could have more features, which take time for results. A brand-new meta-heuristic model is Cuckoo Search method was motivated by some cuckoo required parasitism, the females of which deposit their eggs in the nests of the birds. Some host nests are capable of direct combat [14,15].

If the host bird realizes the foreign eggs are not its own, it will either throw them away or just abandon the nest and build a new one. Female cuckoos are quite skilled at copying the shades and appearance of the eggs of a select group of host bird. As a result, their reproductive potential is increased and the likelihood of their eggs would be abandoned decreases. On the other hand, it demonstrated that the flight patterns of numerous animals and insects closely resemble those of Levy flights. These nesting and flight behaviours were taken into consideration when the authors presented the CS algorithm [16]. The three main idealized rules are observed by the CS algorithm:

- (1) Every cuckoo produces one egg and places in a nest, selected at arbitrary.
- (2) High-quality eggs from the best nests will be passed down to next generations.
- (3) The host bird has a chance of finding a certain number of host nests and $p_a \in [0, 1]$ of finding a cuckoo egg. The host bird has two alternatives in this situation: either throw the egg away or to leave the nest and build a new one.

The proportion p_a to reach this ultimate postulate, one might use the n nests that are replaced with fresh nests.

The neutral function can be the fitness solution for a feature selection problem. Each egg in a nest symbolizes a result in this algorithm, and a cuckoo egg indicates a result. The aim is to utilize the novel and maybe represent a potential (cuckoos) to replace any less-than-ideal solutions in the nests. The fundamental process of the Cuckoo Search (CS) could be distilled down to these three rules [17]. The process for locating a new bird's nest in the three idealized guidelines above is as follows:

$$h_i^{(t+1)} = h_i + \alpha \oplus \text{Levy}(\lambda); i = 1, 2, \dots, n \quad (11)$$

The location of the t generation's i -th bird nest is indicated by the symbol $h_i^{(t+1)}$

α is the step size control, $\alpha > 0$, usually, $\alpha = 1$. Levy (λ) is the expression of Levi random search path shown in following equation:

$$\text{Levy}(\lambda) = t^{-\lambda}; 1 < \lambda < 3 \quad (12)$$

Regular cuckoo search has the drawback that discovery probability of the CS algorithm and the step size are fixed at the initiation. The search accuracy is decreased when the step size is set to a high value. If the step size is too tiny, the search speed is decreased and dropped into the local optimal. An enhanced CS algorithm was presented to address those problems.

In the ICSO method, step size and iteration count are combined, with a higher step length set at the start of the iteration and a smaller step size decreasing as the iteration goes on. In the final iteration of the method, achieve local optimization with reduced step sizes and improved search precision. The enhanced formula is

$$\alpha_i = a_{max} * \frac{1}{\left(\frac{a_{max}}{a_{min}}\right)^{\frac{t}{T}}} * \text{ran}_i * 0.01 \quad (13)$$

a_{min} is the minimum step size, a_{max} , is the maximum step size. T represents the overall sum of iterations. T stands for the presently occurring iteration. The i th component of the data sets ran_i is its scope. The flowchart for enhanced cuckoo search optimization is shown in Figure 3.

3.7. Classification using modified recurrent neural network

Modified Recurrent Neural Networks should be used for classification after feature selection to determine the class label. A type of artificial neural network known as RNN expands on the regular feed-forward neural network [18]. An RNN may handle consecutive inputs, whose firing at each step reflects that of the prior phase. The system can display dynamic temporal behaviour in this way.

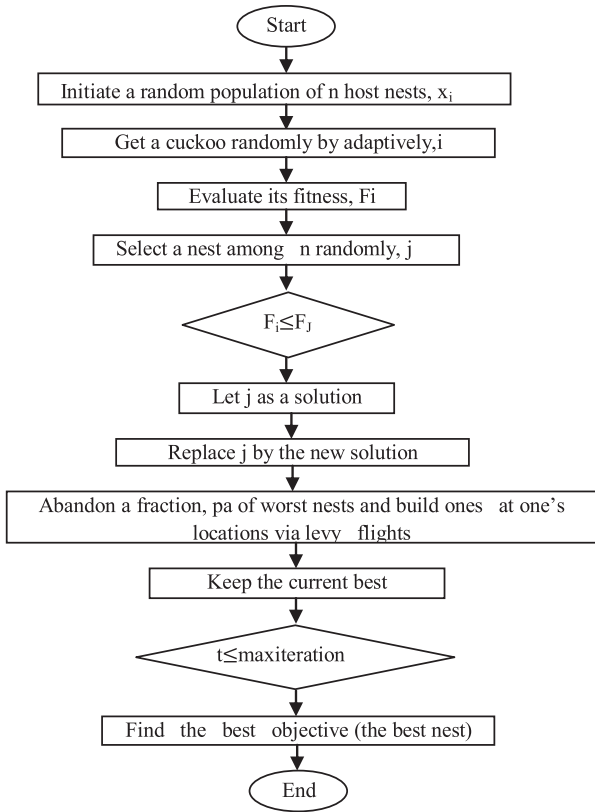


Figure 3. Flow chart of improved cuckoo search optimization.

Algorithm for ICSSO

INPUT: Extracted features
 OUTPUT: Optimal features

1. Generate primary population of N host nest $x_i \forall i, i = 1, \dots, n$.
2. **while** $t < \text{Max Generation}$ or (stop criterion) **do**
3. Get a cuckoo arbitrarily by adaptive manner and assess its fitness, F_i .
4. Select arbitrarily a nest j amongst n .
5. **if** $F_i > F_j$ **then**
6. Swap j by the new result.
7. **end if**
8. A fraction (p_a) of worse nest is deserted and novel ones is constructed.
9. Retain the top results (or nest with quality solutions).
10. Rank the results and find the best.
11. **end while**

An RNN model typically contains three layers: an output layer, input layer and hidden layer (see Figure 4). The hidden layer contains processing elements linked together by weighted synapses. RNNs, as a significant member of the deep learning family, have lately demonstrated effective outcomes in computer vision tasks and machine learning [19].

Yet, it was discovered that training RNNs to deal with long-term sequential data is challenging since inclines begin to evaporate. One popular technique to address this issue is to construct a more complicated recurrent unit.

Modified recurrent neural network

Long terrible memory is a sort long-term dependence using a recurrent hidden unit. Before applying

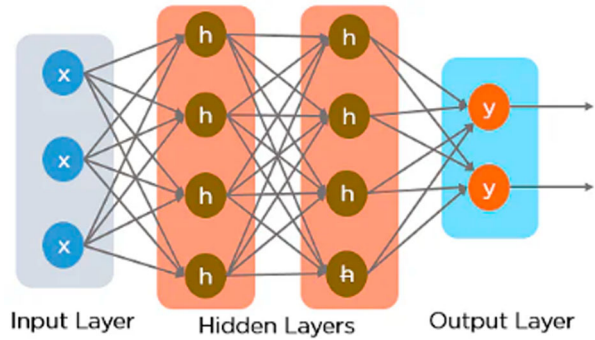


Figure 4. Recurrent Neural Network.

a nonlinear function, a traditional recurrent layer calculates a weighted linear average of the inputs. A memory cell is also formed by a recurrent layer with long short-term memory.

Provided to the input layer a hyperspectral pixel sequence $x = (x^1, x^2, \dots, x^K)$ for a hyperspectral image classification task. At step t , a memory cell is created by a recurrent layer based on LSTM c_t . The activation of processing layers can be calculated using.

$$h_t = o_t \tanh(c_t) \tag{14}$$

Here $o_t \tanh(\bullet)$ is the hyperbolic tangent function, and is the output gate that determines which portion of the memory material will be shown. The output layer is altered by following equation

$$o_t = \sigma(W_{oi}x_t + W_{oh}h_{t-1} + W_{OCCT}) \tag{15}$$

Here W means weight matrices: e.g. W_{oi} is the input-output weight matrix and W_{OC} signifies the memory-output weight matrix. $\sigma(\bullet)$ is a logistic sigmoid function. Memory cell c_t is modified by inserting new memory cell (\bar{c}_t) content and deleting some of the existing memory material.

$$c_t = i_t \odot \bar{c}_t + f_t \otimes c_{t-1} \tag{16}$$

Here \odot is element-wise multiplication as well as the novel memory cell content \bar{c}_t is derived by

$$\bar{c}_t = \tanh(W_{cix_t} + W_{ch}h_{t-1}) \tag{17}$$

Modified recurrent neural network parameter optimization using ICSSO

This phase simultaneously modifies network variables such as the activation function, number of neurones, regularization rate, loss function, optimizer and so on. This work employs ICSSO to optimize the parameters of a Modified Recurrent Neural Network.

4. Results and discussion

This segment examines the outcomes of the suggested technique experiments. This model is executed using

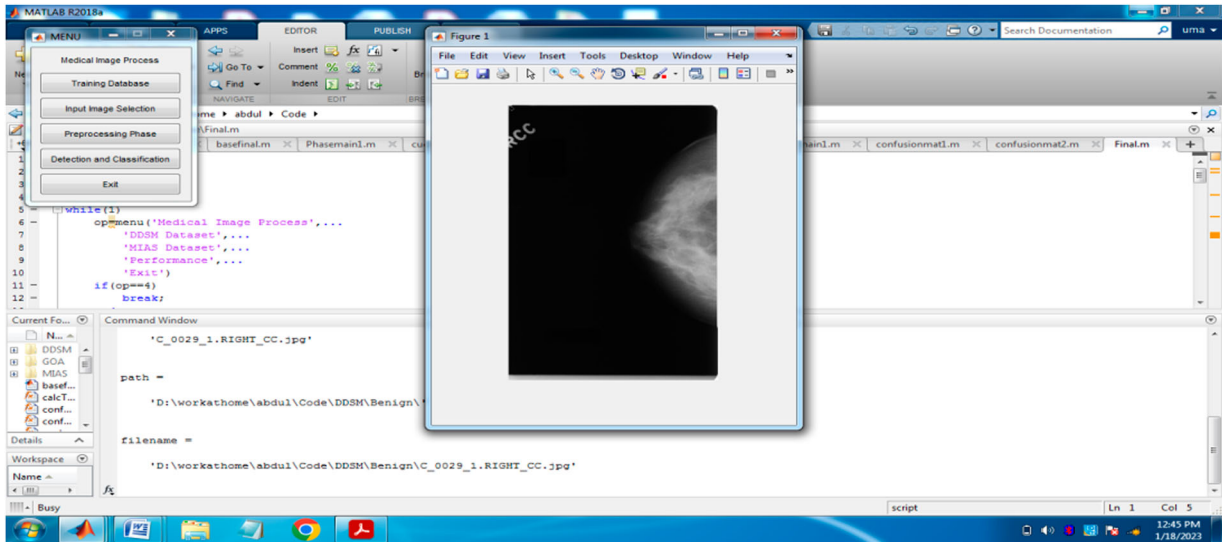


Figure 5. Input image.

MATLAB. For the DDSM and MIAS databases, the recall, precision, F-measure and accuracy of the proposed MRNN model are compared with those of the existing ELM and KELM.

Figure 5 displays the results Figure 6 displays the outcomes of the pre-processing for the chosen input picture, and Figure 7 displays the ROI extraction results.

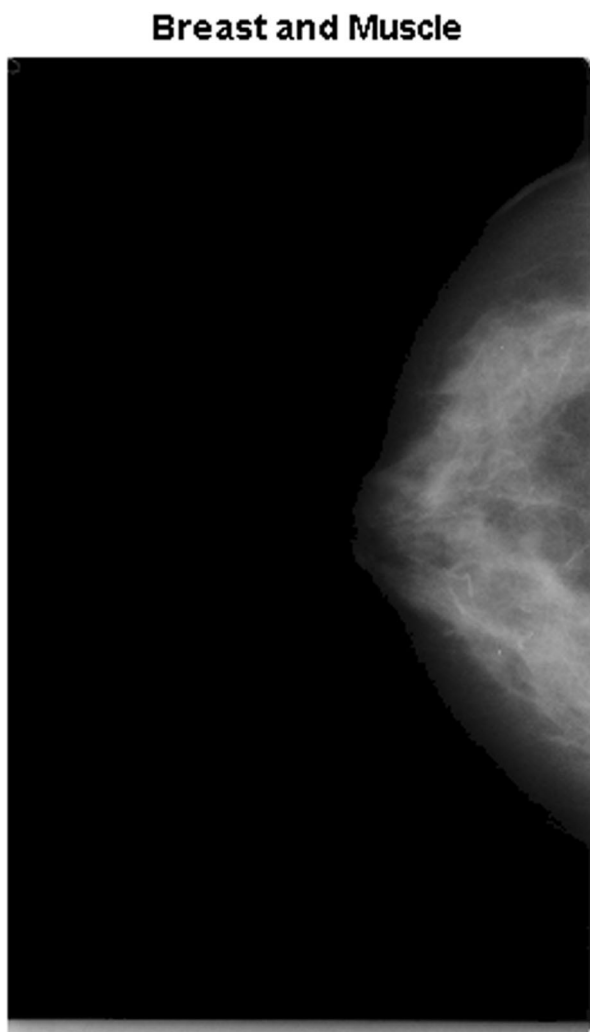


Figure 6. Pre processing.

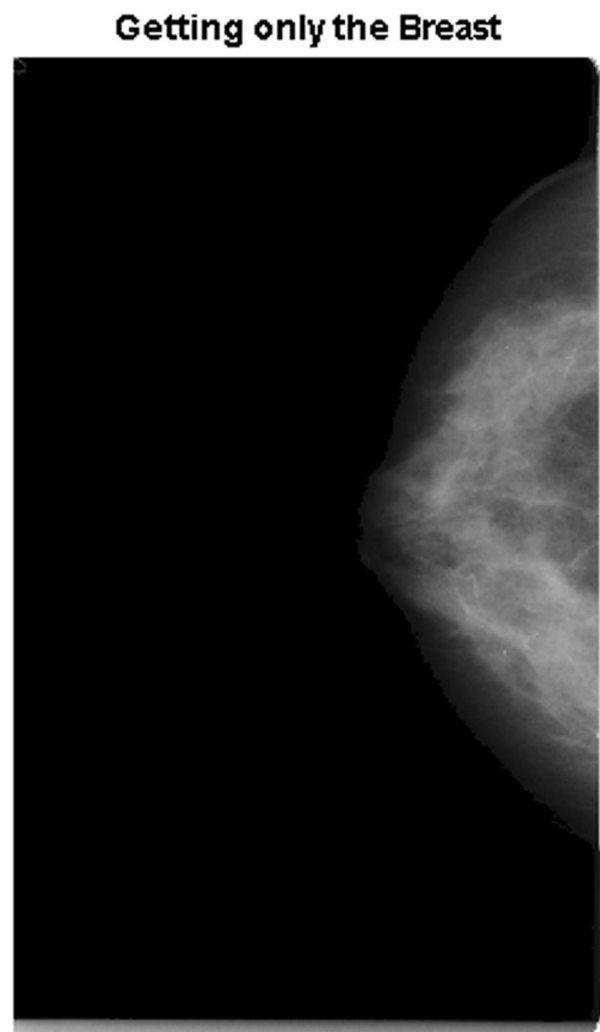


Figure 7. ROI extraction.

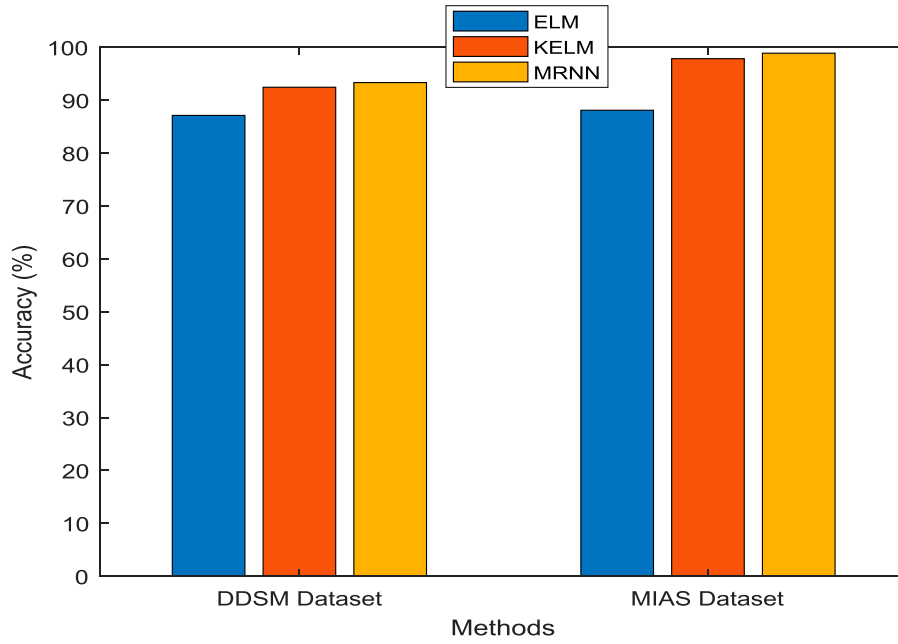


Figure 8. Accuracy results vs. classification methods.

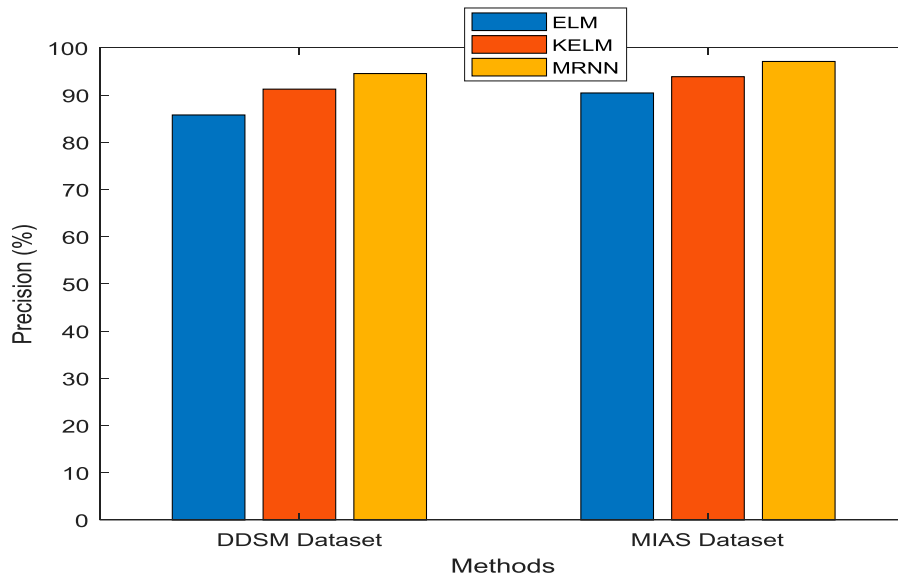


Figure 9. Precision vs. classification methods.

Performancemetrics

(1) Precision

Precision is labelled as the percentage of relevant outcomes that are obtained

$$\text{Precision} = \frac{\text{Truepositive}}{\text{truepositive} + \text{falsepositive}} \quad (18)$$

(2) Recall

Recall is labelled as a percentage of overall related results classified by the suggested algorithm.

$$\text{Recall} = \frac{\text{Truepositive}}{\text{truepositive} + \text{FalseNegative}} \quad (19)$$

(3) Accuracy

Accuracy is the proportion of correct predictions made by this model, defined as follows:

$$\text{Accuracy} = \frac{\text{Truepositive} + \text{TrueNegative}}{\text{Total}} \quad (20)$$

(4) Errorrate

The error rate is a measure of a model prediction error regarding the real model, and derived as.

$$\text{Error rate} = 100 - \text{Accuracy} \quad (21)$$

Figure 8 compares the suggested MRNN with the current ELM and KELM approaches in terms of accuracy

performance metrics. The x -axis in the graph above reflects the various approaches, while the y -axis shows the accuracy numbers. Founded results of the recommended MRNN technique has the greatest Accuracy results of 93.33%, whereas the existing KELM and ELM models produce only 92.45% and 87.12% for the DDSM dataset, respectively (Table 1).

Figure 9 represents precision comparison between the suggested MRNN and the previous ELM and KELM approaches. The x -axis in the graph above reflects the

Table 1. Performance comparison results.

Database	Metrics	Methods		
		ELM	KELM	MRNN
DDSM	Accuracy	87.12	92.45	93.33
	Precision	85.78	91.26	94.56
	Recall	86.65	90.34	90.13
	Error rate	14.87	7.55	6.66
MIAS	Accuracy	88.1	97.85	98.88
	Precision	90.45	93.09	97.12
	Recall	91.23	94.54	95.68
	Error rate	11.9	2.15	1.11

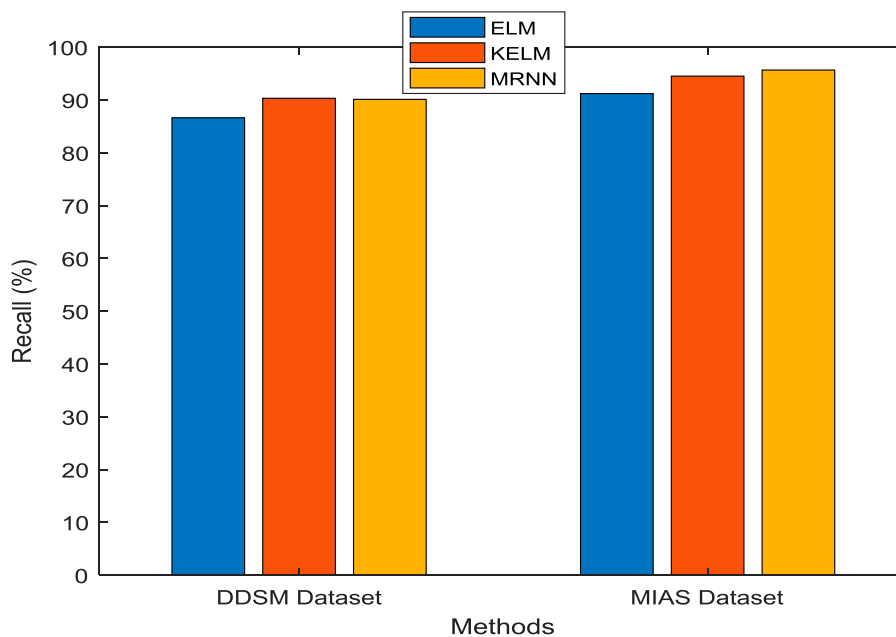


Figure 10. Recall results vs. classification methods.

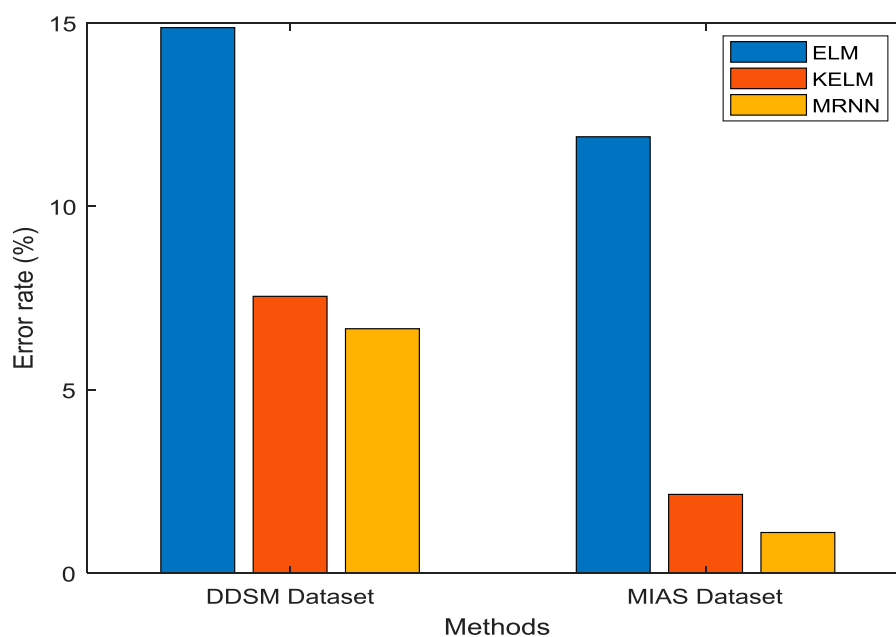


Figure 11. Error rate results vs. classification methods.

various approaches, while the y -axis shows the precision values. Based on the results, it is determined that the suggested MRNN model yields greater precision results of 94.56%, while the present KELM and ELM models yield only 91.26% and 85.78% for the DDSM dataset, respectively.

Figure 10 depicts recall evaluation between the suggested MRNN and the present KELM and ELM techniques. The x -axis in the graph above reflects the various approaches, whereas the y -axis indicates the recall values. According to the results, the suggested MRNN model has a higher value of 90.13%, while the present KELM and ELM and models only have 90.34% and 86.65% for the DDSM dataset.

Figure 11 compares the error rates of classification including MRNN and the conventional KELM and ELM algorithms. The x -axis in the graph above reflects the various approaches, whereas the y -axis indicates the recall values. According to the results, the suggested MRNN model has a lower error rate of 6.66%, while the conventional KELM and ELM approaches had error rates of 7.55% and 14.87%, respectively, for the DDSM dataset.

5. Conclusion and future work

This paper is focused on developing a more accurate framework for identifying and categorizing breast cancer. To eliminate noise from the input breast image, pre-processing is performed utilizing median filtering. Weighted K means that clustering is used for ROI segmentation. Features are obtained using a Block-based Centroid Distance Function (CDF) and CDTM. Kernel Principle Component Analysis (KPCA) is used to reduce dimensionality. Depending on ICSO, relevant feature selection is calculated. Furthermore, the ICSO is utilized to fine-tune the properties of the Modified Recurrent Neural Network (MRNN), which is utilized to classify breast cancer as amalignant or benign. According to experimental findings, the suggested model yields 86% accuracy, which is superior to other models currently used. The enormous number of parameters in the suggested deep learning model makes computation more difficult. The proposed algorithm achieved only 93.33% accuracy. In future, swarm-intelligence-based optimization algorithms will be used to tune the weight and bias of classification algorithms to improve classification accuracy.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

[1] Yassin NI, Omran S, El Houby EM, et al. Machine learning techniques for breast cancer computer aided

- diagnosis using different image modalities: a systematic review. *Comp Meth Prog Biom.* 2018;156:25–45. doi:10.1016/j.cmpb.2017.12.012
- [2] Wang X, Qin T, Witte RS, et al. Computational feasibility study of contrast-enhanced thermoacoustic imaging for breast cancer detection using realistic numerical breast phantoms. *IEEE Trans Microw Theor Tech.* 2015;63(5):1489–1501. doi:10.1109/TMTT.2015.2417866
- [3] Veta M, Pluim JP, Van Diest PJ, et al. Breast cancer histopathology image analysis: a review. *IEEE Trans Biomed Eng.* 2014;61:1400–1411. doi:10.1109/TBME.2014.2303852
- [4] George K, Faziludeen S, Sankaran P. Breast cancer detection from biopsy images using nucleus guided transfer learning and belief based fusion. *Comput Biol Med.* 2020;124:103954–20. doi:10.1016/j.compbimed.2020.103954
- [5] Wan T, Cao J, Chen J, et al. Automated grading of breast cancer histopathology using cascaded ensemble with combination of multi-level image features. *Neuro Comp.* 2017;229:34–44.
- [6] Kumar A, Singh SK, Saxena S, et al. Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer. *Inf Sci (Ny).* 2020;508:405–421. doi:10.1016/j.ins.2019.08.072
- [7] Francis SV, Sasikala M, Bharathi GB, et al. Breast cancer detection in rotational thermography images using texture features. *Infra Phys Technol.* 2014;67:490–496. doi:10.1016/j.infrared.2014.08.019
- [8] Carvalho ED, Antonio Filho OC, Silva RR, et al. Breast cancer diagnosis from histopathological images using textural features and CBIR. *Artif Intell Med.* 2020;105; doi:10.1016/j.artmed.2020.101845
- [9] Vijayarajeswari R, Parthasarathy P, Vivekanandan S, et al. Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. *Measurement.* 2019;146:800–805. doi:10.1016/j.measurement.2019.05.083.
- [10] Hassanien AE, Moftah HM, Azar AT, et al. MRI breast cancer diagnosis hybrid approach using adaptive ant-based segmentation and multilayer perceptron neural networks classifier. *Appl Soft Comput.* 2014;14:62–71. doi:10.1016/j.asoc.2013.08.011
- [11] Muduli D, Dash R, Majhi B. Automated breast cancer detection in digital mammograms: a moth flame optimization based ELM approach. *Biomed Sign Proc Cont.* 2020;59:101912–12. doi:10.1016/j.bspc.2020.101912
- [12] Varatharajan R, Manogaran G, Priyan MK. Retracted article: a big data classification approach using LDA with an enhanced SVM method for ECG signals in cloud computing. *Multimed Tools Appl.* 2018;77(8): 10195–10215. doi:10.1007/s11042-017-5318-1
- [13] Celebi ME, Kingravi HA, Vela PA. A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Syst Appl.* 2013;40(1):200–210. doi:10.1016/j.eswa.2012.07.021
- [14] Usman AM, Yusof UK, Naim S. Filter-based multi-objective feature selection using NSGA III and cuckoo optimization algorithm. *IEEE Access.* 2020;8:76333–76356. doi:10.1109/ACCESS.2020.2987057
- [15] Raeisi A, Akbarizadeh G, Mahmoudi A. Combined method of an efficient cuckoo search algorithm and nonnegative matrix factorization of different zernike moment features for discrimination between oil spills and lookalikes in SAR images. *IEEE J. Selec Topic Appl*

- Earth Observ Remot Sens. 2018;11(11):4193–4205. doi:[10.1109/JSTARS.2018.2841503](https://doi.org/10.1109/JSTARS.2018.2841503)
- [16] Shetty RP, Sathyabhama A, Pai PS. An efficient online sequential extreme learning machine model based on feature selection and parameter optimization using cuckoo search algorithm for multi-step wind speed forecasting. *Soft Comput.* 2021;25(2):1277–1295. doi:[10.1007/s00500-020-05222-x](https://doi.org/10.1007/s00500-020-05222-x)
- [17] Yildirim S, Kaya Y, Kılıç F. A modified feature selection method based on metaheuristic algorithms for speech emotion recognition. *Appl Acoust.* 2021;173:107721–11. doi:[10.1016/j.apacoust.2020.107721](https://doi.org/10.1016/j.apacoust.2020.107721)
- [18] Mesnil G, Dauphin Y, Yao K, et al. Using recurrent neural networks for slot filling in spoken language understanding. *IEEE/ACM Trans Audio Speech Lang Process.* 2015;23(3):530–539. doi:[10.1109/TASLP.2014.2383614](https://doi.org/10.1109/TASLP.2014.2383614)
- [19] Shriram S, Nagaraj B, Jaya J, et al. Deep learning-based real-time AI virtual mouse system using computer vision to avoid COVID-19 spread. *J Healthc Eng.* 2021: 1–8. doi:[10.1155/2021/8133076](https://doi.org/10.1155/2021/8133076)