Automated breast cancer detection system from breast mammogram using deep neural network

Suneetha Chittineni[1], Sai Sandeep Edara[2]
[1]Department of Computer Applications, R. V. R. and J. C. College of Engineering, Chowdavaram, Guntur, India
[2]Department of Computer Science and Engineering, R. V. R. and J. C. College of Engineering, Chowdavaram, Guntur, India

Article Info	ABSTRACT
Article history: Received Apr 8, 2021 Revised Nov 23, 2021 Accepted Nov 26, 2021 **Keywords:** Breast cancer Deep neural network Hybridization Mammography Thermography	All over the world breast cancer is a major disease which mostly affects the women and it may also cause death if it is not diagnosed in its early stage. But nowadays, several screening methods like magnetic resonance imaging (MRI), ultrasound imaging, thermography and mammography are available to detect the breast cancer. In this article mammography images are used to detect the breast cancer. In mammography image the cancerous lumps/microcalcifications are seen to be tiny with low contrast therefore it is difficult for the doctors/radiologist to detect it. Hence, to help the doctors/radiologist a novel system based on deep neural network is introduced in this article that detects the cancerous lumps/microcalcifications automatically from the mammogram images. The system acquires the mammographic images from the mammographic image analysis society (MIAS) data set. After pre-processing these images by 2D median image filter, cancerous features are extracted from the images by the hybridization of convolutional neural network with rat swarm optimization algorithm. Finally, the breast cancer patients are classified by integrating random forest with arithmetic optimization algorithm. This system identifies the breast cancer patients accurately and its performance is relatively high compared to other approaches.

Corresponding Author:

Suneetha Chittineni
Department of Computer Applications, R. V. R. and J. C. College of Engineering
Chowdavaram, Guntur, India
Email: suneethachittineni@gmail.com

1. INTRODUCTION

One of the most common diseases that affect women in recent years is the breast cancer [1]. In the latest survey taken by world health organization (WHO) it is predicted that by 2025 in the world there are 19.3 million victims affected by breast cancer. Breast cancer is a condition in which cells grow out of control, resulting in a tumour that can spread throughout the body. Although the specific causes of breast cancer are unknown, researchers believe that aberrant cell growth is caused by a combination of genes, lifestyle, environment, and hormones [2]. This breast cancer must be detected in its early stage otherwise it may cause death. Hence, there are numerous medical imaging techniques like m agnetic resonance imaging (MRI), ultrasound imaging, thermography and mammography are available to identify the breast cancer [3]. But, diagnosing breast cancer at its early stage becomes a challenging work to the medical experts like doctors and radiologists.

In magnetic resonance imaging (MRI) the breast images are captured from 3D view. It employs a non-ionizing radiation [4]. But the rate of MRI is high and it is difficult to differentiate the normal lumps

and cancerous lumps from the breast MRI. The breast ultrasound produces less accurate results for patients with dense breast [5]. The result of images is based on the expert who is taking the ultrasound for the patient. So many times it produces high false positive rate that leads to unnecessary biopsy [6]. Breast thermography utilizes the infrared cameras to capture the breast images [7]. The camera has an inbuilt infrared sensor that helps to record the temperature of the breast. Based on the variation in the temperature the breast cancer is detected. If any cold pressure was applied to the breast then the original breast temperature changes and produces false result. To avoid these limitations in this study mammographic images are used to identify the breast cancer in its early stage. Mammography is one of the commonly used and believed methods to identify the lumps in the breast.

The mammography images show the presence of cancerous lumps/ microcalcifications in the breast [8]. The cancerous lumps in the mammography images are tiny in size and its image contrast is low. So it is hard for the doctors/radiologists to detect the microcalcifications/cancerous lumps in the mammographic images. Hence, to ease the work of the doctors/radiologists, a novel system is proposed that detects the cancerous lumps in the breast from the mammographic images. The proposed deep neural network system acquires the mammography images from the mammographic image analysis society (MIAS) image dataset. The convolutional neural network (CNN) algorithm, integrated with rat swarm optimization extracts the features of breast cancer from the mammographic images. The features are extracted by tuning the parameters of CNN and there by updating the position of the rat swarm optimization. Then the extracted features are classified using the classifier random forest integrated with the arithmetic optimization algorithm. The classifier is designed by the arithmetic optimization algorithm which helps to avoid the reasoning problem occur in the output of CNN. Figure 1 describes the block diagram of the proposed system.
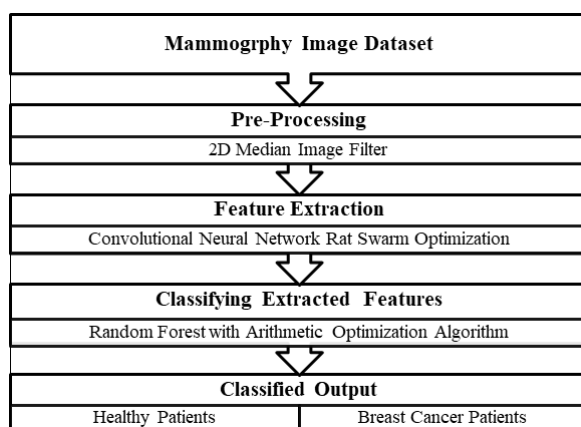

Figure 1. Block diagram of the proposed system

This article is planned as follows. Literature review based on mammography images, extracting the features and classifying the breast cancer is explained in section 2. Section 3 explains the background models used in the proposed approach. Section 4 describes the proposed method and the algorithms used in the proposed method. In section 5, experimental results with simulation are explained and discussed. The conclusion of this article and its future work is discussed in section 6.

2.   LITERATURE REVIEW

Cao *et al.* [9] introduces a novel convolutional neural network (CNN) framework to identify the breast cancer from the ultrasound images. This framework contains several object detection and classification approaches to identify the tumour. It first detects the presence of tumour in the breast ultrasound image and then classifies the type of the tumour by the CNN framework. In this, underfitting problem occurs while finding the malignant lumps and only less parameter are considered to identify the tumour. Singh and Singh [10] combine and improve several existing approaches in segmentation, feature selection, feature extraction and classification. And then applied this approaches to the thermography images. It identifies the breast cancer but can only be suitable for database having less thermography images. The accuracy of the result may

vary based on the dimension of the lump and also produce false positive rate. Chen *et al.* [4] introduced an abbreviated protocol (AP) for MRI that identifies the cancerous lumps in the breast. This protocol has two other protocols abbreviated protocol1 (AP1) and abbreviated protocol2 (AP2). Maximum intensity projection (MIP) and first post-contrast subtracted (FAST) images were grouped together to form AP1 protocol. AP2 protocol was a combination of AP1 protocol with diffusion-weighted imaging (DWI). These two protocols examine the ultrasound images and then detect the breast cancer. But this model has the limitations that it doesn't consider the past history of the patients also the small lumps in the breast are not identified in this method.

Aslam *et al.* [11] implemented an automatic deep convolutional neural network (DCNN) approach for identifying the breast cancer. This approach first gathers the data from two datasets. Then utilizes the convolutional neural network layers for training the data and then classifies the breast cancer patients. The performance of this approach was based on the number of data available for training. If the training data decreases the performance of this approach also decreases. Ibrahim *et al.* [12] uses thermal images to identify the breast cancer. The thermal images are gathered from the database for mastology research with infrared image (DMR-IR). The gathered thermal images undergo pre-processing and segmentation. After that the cancerous features are extracted from the segmented image. Then the breast cancer patients were classified from the extracted images. During this process various algorithms were used in every stage that may cause many problems like setting the k-value and the data after merging totally changed from its original size and density, which produces wrong prediction. To overcome the above limitations the mammography images are used in this article. From the mammography images it is difficult to identify the cancerous lumps for that well experienced experts are needed and they have to examine the mammography image clearly to detect the breast cancer correctly. All the time the experts are not available so to ease their work an automated system is implemented to detect the breast cancer using conventional neural network and arithmetic optimization algorithm.

3. BACKGROUND
3.1. Deep neural network

One of the sub divisions of machine learning is deep learning model. The deep learning is designed by including more hidden layers in the traditional neural networks. The hidden layers are present in-between the input layer and output layer. The deep neural network (DNN) becomes popular in medical field because it provides high performance in extracting the features from the images [13]. In order to provide good performance DNN requires huge dataset for training the model. Selecting the hyper-parameter is also an important process in DNN to extract the optimal features from the mammographic image dataset.

3.2. Convolutional neural network

There are numerous deep neural networks the most commonly used neural network by the researches are convolutional neural network (CNN) [14]. Normally a CNN consists of a set of feed forward layers, this feed forward layers executes the convolutional filter, pooling layer and fully connected layers that helps to extract the image features. By learning the input image patterns CNN allow feature extraction this is done in feature extraction layer/convolutional layer [15]. CNN is used for tuning the hyper parameters like batch size, number of epochs, activation layer and learning rate to extract the features of mammography images. So, radiologists are not needed to segment the breast cancer image features.

4. PROPOSED METHODOLOGY

In the proposed approach the chest mammographic images are obtained from the MIAS image database. The obtained images are pre-processed to make all the images in same size. Then the features are extracted from the pre-processed images by a convolutional neural network integrated with rat swarm optimization algorithm. This algorithm tunes the parameters by updating the location of the rat to extract the breast cancer features from the images. At last arithmetic optimization algorithm integrated with random forest approach classifies the normal and breast cancer affected patients from the extracted images. The reasoning problem in CNN is eliminated by the arithmetic optimization algorithm.

4.1. Pre-processing

Pre-processing is applied to the mammography image database to eliminate the unwanted noise included in the images. By pre-processing the features that are needed for detecting the breast cancer are sharp-

ened and the image quality also improved [16]. This process does not change the features of the original image it only enhances the features. The pre-processing uses 2D median image filter function that increases the mammography image quality by clearing and fading the unwanted image portions out of sight and makes the image suitable for further processing [17], [18]. The mechanism of this filter is it moves every pixel one by one and then every pixel value is altered by the median of neighbouring pixel value.

4.2. Feature extraction

In this process, the cancerous features are extracted from the pre-processed mammographic images by tuning the hyperparameters using convolutional neural network integrated with rat swarm optimization (CNN-RSO) algorithm. The rat swarm optimization algorithm is a bio inspired optimization algorithm that describes the public activities of rat and swarm [19]. Here the rat is the predator that tries to catch the swarm which is the victim. This algorithm tunes the hyperparameters batch size, number of epochs, activation layer, and learning rate by that it alters the location of the rat. The group of rats tries to hunt the swarm by chasing and fighting with it. The predator chasing the victim is mathematically modelled in (1). The information about the locality of the victim is known by the best search agent. Based on the location of the best search agent the other search agents can modify their locations.

$$\vec{L} = U.\vec{L}(x) + V.(\vec{L}_r(x) - \vec{L}_i(x)) \tag{1}$$

Here, the location of the rat is represented by $\vec{L}_i(x)$ and the ideal solution is represented by $\vec{L}_r(x)$. The values for the variables U and V are computed as,

$$U = R - x(\frac{R}{Max_{Iteration}}) \tag{2}$$

$$V = 2.rand() \tag{3}$$

where, the values of x=0,1,2,..., $Max_{Iteration}$. R and V are the random numbers that varies from 1 to 5.The aggressive fighting of the rat with the swarm to kill him is mathematically computed as follows:

$$\vec{L_i}(x+1) = |\vec{L_r}(x) - \vec{L}| \tag{4}$$

here, the next modified location of the rat is represented by $\vec{L_i}(x+1)$. Each time the location of the rat changes the best ideal solution is stored in $\vec{L_i}(x+1)$. Algorithm 1 shows the hyperparameter tuning of rat swarm optimization. Thus the hyperparameters batch size, number of epochs, activation layer and learning rate are tuned to extract the breast cancer features from the pre-processed mammography image dataset.

Algorithm 1 Hyperparameter tuning of rat swarm optimization

Input: the batch size $\vec{L}_i$ (i=1,2,...,n)
Output: the best optimal extracted image dataset
Procedure HyperparameterRSO
   Initialize the parameters U, V and R
   Compute the fitness value for each image dataset
   $\vec{L}_r \leftarrow$ the best image dataset
while $x < Max_{Iteration}$ **do**
  for $each image dataset$ **do**
    Change the location of the present image dataset by (4)
  end for
  Change the parameters U, V and R
  Verify if any image dataset goes out of the given image dataset then adjust it
  Compute the fitness value for each image dataset
  Change Lr if any better solution is found
  x = x+1
end while
Return $\vec{L}_r$

4.3.   Classification

The breast cancer patients are classified from the feature extracted dataset by arithmetic optimization algorithm integrated with random forest (AOA-RF) [20]. The AOA is a population-based algorithm so the ideal solution cannot be found in a single step. It takes much iteration to found the best ideal solution. The best ideal solution in AOA is obtained by the arithmetic operators addition (A), subtraction (S), multiplication (M), and division (D). Initialization phase, exploratory phase and exploitative phase are the three processes in the AOA approaches.

4.3.1. Initialization phase

In initialization phase, the random forest (RF) algorithm is implemented to retrieve the best obtained or the nearly optimum solution [21]. The set of candidate solutions (C) is generated from the decision trees (DT) each iteration the ideal candidate solution is treated as a best obtained or the nearly optimum solution. There are L image dataset in the decision tree, the candidate solution C is represented in (5),

$$C = c(S, \theta_i) \quad i = 1, 2, ..., L \tag{5}$$

here, ith decision tree is represented by (S, θ_i). The samples for training is S and the single tree growth is represented as θ_i.

4.3.2. Exploration phase

The operators multiplication (M) and division (D) are considered as the operators for exploration. These two operators produce high decision values which helps the exploration phase to search the near ideal solution. The exploration phase can also be used in exploitative phase to assist it to find the accurate breast cancer patients. For this process it applies two techniques: division (D) search approach and multiplication (M) search approach. This technique is represented in (6).

$$x_{i,j}(P_{iter} + 1) = \begin{cases} best(x_j)/(MOP+ \in) * ((UV_j - LV_j) * \mu + LV_j), & r2 < 0.5 \\ best(x_j) * MOP * ((UV_j - LV_j) * \mu + LV_j), & Otherwise \end{cases} \tag{6}$$

Where, r denotes the random number, UV and LV represents the upper value and lower value, the result of ith location in the next iteration is $x_{i,j}(P_{iter} + 1)$, P_{iter} represents the present iteration. The math optimizer probability (MOP) is calculated in (7). The maximum number of iteration is represented as Max_{iter}.

$$MOP(P_{Iter}) = 1 - \frac{P_{Iter}^{1/\infty}}{Max_{Iter}^{1/\infty}} \tag{7}$$

4.3.3. Exploitation phase

The operators addition (A) and subtraction (S) are considered as the operators for exploitation. These two operators produce low decision values which help the exploitation phase to choose the best ideal dataset. The exploitation phase is mathematically modelled as (8).

$$x_{i,j}(P_{iter} + 1) = \begin{cases} best(x_j)/MOP * ((UV_j - LV_j) * \mu + LV_j), & r3 < 0.5 \\ best(x_j) * MOP * ((UV_j - LV_j) * \mu + LV_j), & Otherwise \end{cases} \tag{8}$$

This exploitation phase is similar to the exploration phase but it does not jammed in any dataset while searching. The final classified breast cancer patients are obtained by (9),

$$B(C) = argmax(x_{i,j})(b_{r,b,l}(C) = i) \quad i = 1, 2, ..., N \tag{9}$$

where, B(C) represents the final classified breast cancer patients, $x_{i,j}$ is the number of near ideal dataset.

5.   RESULT AND DISCUSSION

In MATLAB R2018a software the proposed system is implemented. The Mammographic Image Analysis Society (MIAS) database consists of 322 breast mammography images is used in the proposed system for the experimentation purpose. In that 206 are normal images and 113 are breast cancer images [22]. The proposed approach is compared with other classifiers like Naive Bayes (NB) [23], k-nearest neighbor (KNN) [24], decision tree (DT) [23], support vector machine (SVM) [24] and random forest (RF) [25].

The performance metrics considered for evaluation are accuracy, sensitivity, F1-Score, precision, specificity and Kappa statistic. Figure 2 shows the accuracy and precision for various algorithms. From that it is proved that the proposed AOA-RF produce high accuracy compared to other approaches. Figure 3 shows the performance of F1-score and kappa for various algorithms.
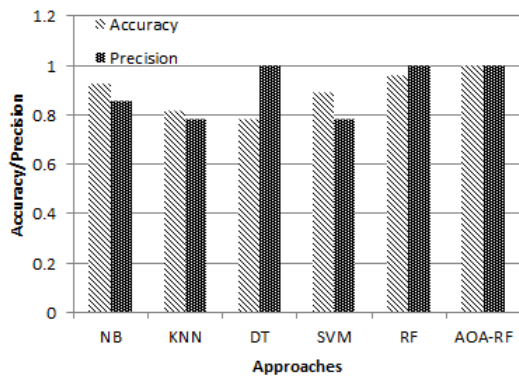
Figure 2. Performance of accuracy and precision
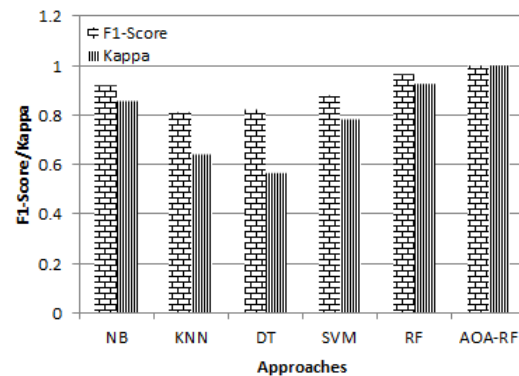
Figure 3. Performance of F1-score and kappa

Both the F1-score and kappa values are relatively high for the proposed approach. Figure 4 shows the performance of sensitivity and specificity for various algorithms. From these figures it is clear that the performance of the proposed classifier random forest integrated with arithmetic optimization algorithm is superiorly high compared to other algorithms.
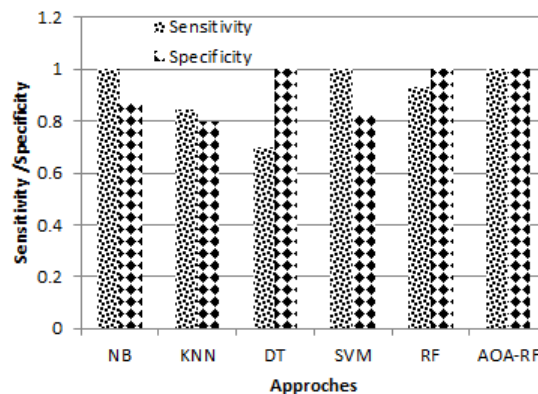

Figure 4. Performance of sensitivity and specificity

The root mean square error (RMSE) and mean absolute error (MAE) are combined into one to detect the error in the breast cancer dataset. Figure 5 represents the RMSE and MAE error for the proposed AOA-RF and for various other existing algorithms such as DT, KNN, SVM, NB, RF. From that it is evident that the proposed classifier produces less error compared to other classifiers. The proposed approach with and without rat swarm optimization (RSO) values are shown in Table 1. Figure 6 shows the proposed approach performance with RSO and without RSO.
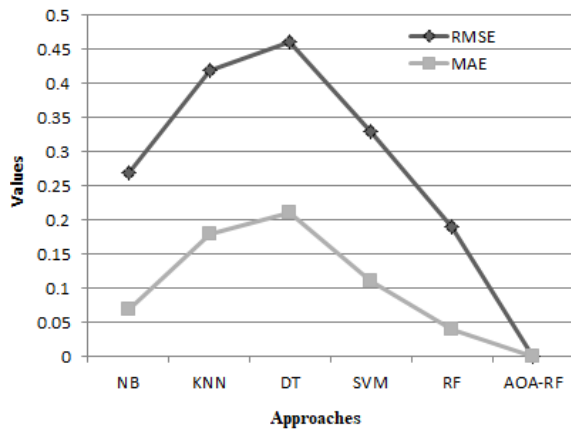
Figure 5. Comparative analysis of the proposed
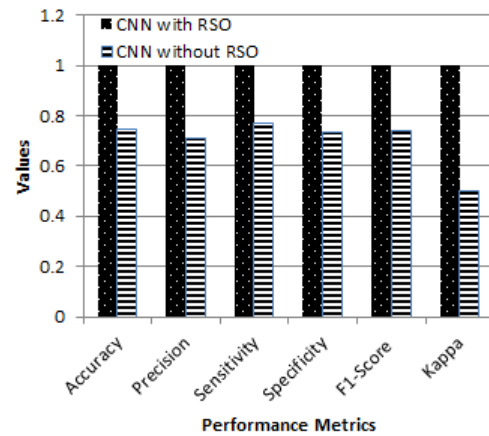AOA-RF and existing DT, KNN, SVM, NB, RF

Figure 6. Performance of proposed approach
with and without RSO

Table 1. Proposed approach with and without RSO

Proposed With RSO		Proposed Without RSO	
Parameters	Values	Parameters	Values
Specificity	1	Specificity	0.7333
FPR (1-Specificity)	0	FPR (1-Specificity)	0.2667
TPR (Sensitivity)	1	TPR (Sensitivity)	0.7692
Error	0	Error	0.2500
Precision	1	Precision	0.7143
F-measure	1	F-measure	0.7407
Accuracy	1	Accuracy	0.7500
MCC	1	MCC	0.5013
Kappa	1	Kappa	0.5000

The proposed approach with RSO produces high accuracy, precision, sensitivity, specificity, F1-score and kappa values nearly 1 compared to the proposed approach without RSO. The receiver operating characteristics (ROC) curve analysis of proposed classifier with RSO and without RSO is shown in Figure 7. The receiver operating characteristics (ROC) curve analysis of proposed classifier AOA-RF and other different classifiers with CNN as feature extraction is shown in Figure 8.
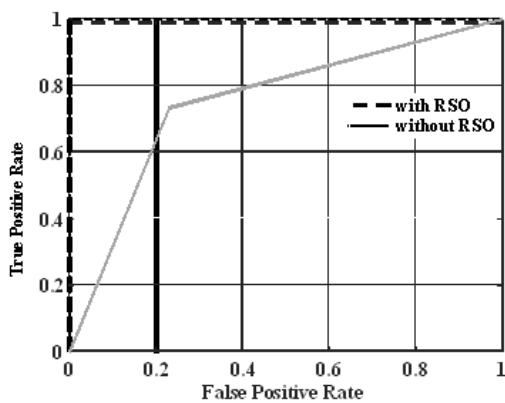

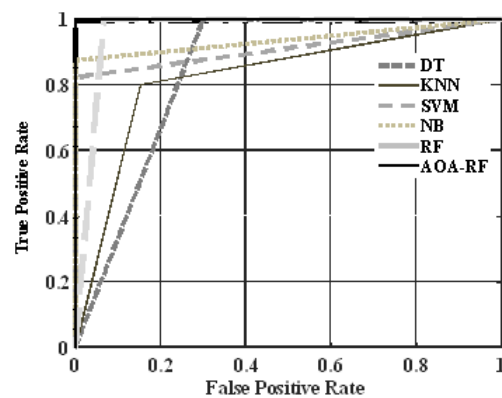Figure 7. Proposed classifier AOA-RF with and without
RSO

Figure 8. Comparison of proposed classifier
AOA-RF with other classifiers

The final result obtained is shown by the confusion matrix. The confusion matrix obtained for AOA-RF classifier without RSO is given in Figure 9. The confusion matrix obtained for AOA-RF classifier with RSO

is given in Figure 10. The proposed approach with RSO produces 100% accuracy which is relatively higher than proposed approach without RSO produces 75% accuracy.

Figure 9. Confusion matrix for AOA-RF classifier without RSO

Figure 10. Confusion matrix for AOA-RF classifier with RSO

6. CONCLUSION

In this paper, novel deep learning-based automatic breast cancer diagnosing systems from the mammographic images are developed. This system helps the doctors/radiologists to identify the breast cancer automatically. In this system, the mammographic images in MIAS dataset undergoes image pre-processing by 2D median image filter to remove the noise in the dataset. The breast cancer features from the pre-processed mammographic images are retrieved using convolutional neural network integrated with rat swarm optimization (CNN-RSO) algorithm. Finally, the arithmetic optimization algorithm integrated with random forest (AOA-RF) classifier classifies the breast cancer affected and unaffected patients. While analysing the performance of the AOA-RF classifier with other classifiers the performance of the proposed classifier AOA-RF is relatively high. The proposed system AOA-RF with CNN-RSO produces 100% accuracy in detecting the breast cancer from the mammographic images. The system can further be improved by increasing the size of the mammographic image dataset.

REFERENCES

[1] J. Wang, X. Yang, H. Cai, W. Tan, C. Jin, and L. Li, "Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning," *Scientific Reports*, vol. 6, p. 27327, 2016, doi:10.1038/srep27327.

[2] C. Aroef, Y. Rivan, and Z. Rustam, "Comparing random forest and support vector machines for breast cancer classification," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 18, no. 2, pp. 815–821, 2020, doi: 10.12928/telkomnika.v18i2.14785.

[3] M. A. Kahya, "Classification enhancement of breast cancer histopathological image using penalized logistic regression," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 13, no. 1, pp. 405–410, 2019, doi: 10.11591/ijeecs.v13.i1.pp405-410.

[4] S. Q. Chen, M. Huang, Y. Y. Shen, C. L. Liu, and C. X. Xu, "Abbreviated MRI Protocols for Detecting Breast Cancer in Women with Dense Breasts," *Korean Journal of Radiology*, vol. 18, no. 3, pp. 470–475, 2017, doi: 10.3348/kjr.2017.18.3.470.

[5] N. R. Shenoy and A. Jatti, "Ultrasound image segmentation through deep learning based improvised U-Net," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 21, no. 3, pp. 1424–1434, 2021, doi: 10.11591/ijeecs.v21.i3.pp1424-1434.

[6] M. S. Islam, N. Kaabouch and W. C. Hu, "A survey of medical imaging techniques used for breast cancer detection," *IEEE International Conference on Electro-Information Technology, EIT 2013*, 2013, pp. 1–5, doi: 10.1109/EIT.2013.6632694.

[7] J.-L. Gonzalez-Hernandez, A. N. Recinella, S. G. Kandlikar, D. Dabydeen, L. Medeiros, and P. Phatak, "Technology, application and potential of dynamic breast thermography for the detection of breast cancer," *International Journal of Heat and Mass Transfer*, vol. 131, pp. 558–573, 2019, doi: 10.1016/j.ijheatmasstransfer.2018.11.089.

[8] M. H. Alhabib and O. H. Alhabib, "Detection of partially overlapped masses in mammograms," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 18, no. 1, pp. 235–241, 2020, doi: 10.11591/ijeecs.v18.i1.pp235-241.

[9] Z. Cao, L. Duan, G. Yang, T. Yue, and Q. Chen, "An experimental study on breast lesion detection & classification from ultrasound images using deep learning architectures," *BMC Medical Imaging*, vol. 19, no. 51, 2019, doi: 10.1186/s12880-019-0349-x.

[10] D. Singh and A. K. Singh, "Role of image thermography in early breast cancer detection- Past, present and future," *Computer Methods and Programs in Biomedicine*, vol. 183, p. 105074, 2020, doi: 10.1016/j.cmpb.2019.105074.

[11] M. A. Aslam, Aslam, and D. Cui, "Breast Cancer Classification using Deep Convolutional Neural Network," *Journal of Physics: Conference Series*, vol. 1584, p. 1584, 2020, doi: 10.1088/1742-6596/1584/1/012005.

[12] A. Ibrahim, S. Mohammed, and H. A. Ali, "Breast Cancer Detection and Classification Using Thermography: A Review," *Advances in Intelligent Systems and Computing*, vol. 723, pp. 496–505, 2018, doi: 10.1007/978-3-319-74690-6_49.

[13]   W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," *Neurocomputing*, vol. 234, pp. 11–26, 2017, doi: 10.1016/j.neucom.2016.12.038.

[14]   S. Z. Ramadan, "Using Convolutional Neural Network with Cheat Sheet and Data Augmentation to Detect Breast Cancer in Mammograms," *Computational and Mathematical Methods in Medicine*, vol. 2020, pp. 1–9, 2020, doi: 10.1155/2020/9523404.

[15]   K. Kumar and A. C. S. Rao, "Breast cancer classification of image using convolutional neural network," *2018 4th International Conference on Recent Advances in Information Technology (RAIT)*, 2018, pp. 1–6, doi: 10.1109/RAIT.2018.8389034.

[16]   A. K. Jothi and P. Mohan, "A Comparison between KNN and SVM for Breast Cancer Diagnosis Using GLCM shape and LBP Features," *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2020, pp. 1058–1062, doi: 10.1109/ICSSIT48917.2020.9214235.

[17]   W. Chen, P. Chen, Y. Hsiao and S. Lin, "A Low-Cost Design of 2D Median Filter," *IEEE Access*, vol. 7, pp. 150623–150629, 2019, doi: 10.1109/ACCESS.2019.2948020.

[18]   B. John and S. Nallathambi, "Study and analysis of filters," *Advances in Computational Sciences and Technology*, vol. 10, no. 3, pp. 331–341, 2017.

[19]   G. Dhiman, M. Garg, A. Nagar, V. Kumar, and M. Dehghani, "A novel algorithm for global optimization: Rat Swarm Optimizer," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, pp. 8457–8482, 2021, doi: 10.1007/s12652-020-02580-0.

[20]   L. Abualigah, A. Diabat, S. Mirjalili, M. A. Elaziz, and A. H. Gandomi, "The Arithmetic Optimization Algorithm," *Computer Methods in Applied Mechanics and Engineering*, vol. 376, p. 113609, 2021, doi: 10.1016/j.cma.2020.113609.

[21]   F. Guo *et al.*, "Random Forest Algorithm-based Multi-Feature Vector Optimization for Fatigue Driving Vigilance Monitoring," in *2020 Chinese Control And Decision Conference (CCDC)*, 2020, pp. 3113–3118, doi: 10.1109/CCDC49329.2020.9164300.

[22]   M. Mustra and A. Stajduhar, "Segmentation Masks for the Mini-Mammographic Image Analysis Society (mini-MIAS) Database," *IEEE Consumer Electronics Magazine*, vol. 9, no. 5, pp. 28–33, 1 Sept. 2020, doi: 10.1109/MCE.2020.2986799.

[23]   B. T. Ahmed, "Data mining techniques for lung and breast cancer diagnosis: A review," *International Journal of Informatics and Communication Technology (IJ-ICT)*, vol. 10, no. 2, pp. 93–103, 2021, doi: 10.11591/ijict.v10i2.pp93-103.

[24]   A. Agrima, I. Mounir, A. Farchi, L. Elmaazouzi, and B. Mounir, "Emotion recognition from syllabic units using k-nearest-neighbor classification and energy distribution," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 6, pp. 5438–5449, 2021, doi: 10.11591/ijece.v11i6.pp5438-5449.

[25]   H. Bunyamin and T. Tunys, "A Comparison of Retweet Prediction Approaches: The Superiority of Random Forest Learning Method," *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, vol. 14, no. 3, pp. 1052–1058, 2016, doi: 10.12928/TELKOMNIKA.v14i3.3150.

BIOGRAPHIES OF AUTHORS

Suneetha Chittineni ⓘ 🔬 ᴷᶜ Ⓟ is an Associate Professor in the Department of Computer Applications at R.V.R. & J.C. College of Engineering, Guntur, India. She received a Ph.D. degree in Computer Science and Engineering with specialization in Artificial Intelligence from Acharya Nagarjuna University, Guntur, India. Her research interests include Artificial Intelligence, Machine Learning, Deep Learning, and Data Mining. Dr. Suneetha has published more than 20 research papers in various international journals. She can be contacted at email: suneethachittineni@gmail.com.
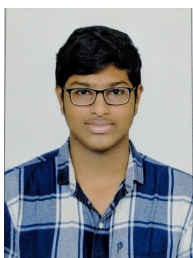
Sai Sandeep Edara ⓘ 🔬 ᴷᶜ Ⓟ is currently a B.Tech final year student pursuing Computer Science & Engineering in R.V.R. & J.C. College of Engineering, Guntur, India. His areas of interest include Data Science, Machine Learning, and Deep Learning. He can be contacted at email: esaisandeep2001@gmail.com.