

A hybrid convolutional neural network-recurrent neural network approach for breast cancer detection through Mask R-CNN and ARI-TFMOA optimization

Keshetti Sreekala¹, Srilatha Yalamati², Annemneedi Lakshmanarao³, Gubbala Kumari⁴,
Tanapaneni Muni Kumari⁵, Venkata Subbaiah Desanamukula⁶

¹Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology (A), Hyderabad, India

²Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India

³Department of Information Technology, Aditya University, Surampalem, India

⁴Department of Computer Science and Engineering, CMR Engineering College, India

⁵Department of Master of Computer Applications, Sri Venkateswara College of Engineering, Tirupati, India

⁶Department of Computer Science and Engineering, Lakireddy Bali Reddy College of Engineering (Autonomous), Mylavaram, India

Article Info

Article history:

Received Aug 18, 2024

Revised Jan 23, 2025

Accepted Mar 3, 2025

Keywords:

Adaptive random increment-based tomitt flock metaheuristic optimization algorithm
Breast cancer detection
Convolutional neural network
Mask recurrent neural network
Recurrent neural network

ABSTRACT

This paper presents a novel hybrid deep learning-based approach for breast cancer detection, addressing critical challenges such as overfitting and performance degradation in varying data conditions. Unlike traditional methods that struggle with detection accuracy, this work integrates a unique combination of advanced segmentation and classification techniques. The segmentation phase leverages Mask region-based convolutional neural network (R-CNN), enhanced by the adaptive random increment-based tomitt flock metaheuristic optimization algorithm (ARI-TFMOA), a novel algorithm inspired by natural flocking behavior. ARI-TFMOA fine-tunes Mask R-CNN parameters, achieving improved feature extraction and segmentation precision while ensuring adaptability to diverse datasets. For classification, a hybrid convolutional neural network-recurrent neural network (CNN-RNN) model is introduced, combining spatial feature extraction by CNNs with temporal pattern recognition by RNNs, resulting in a more nuanced and comprehensive analysis of breast cancer images. The proposed framework achieved significant advancements over existing methods, demonstrating improved performance. This hybrid integration of ARI-TFMOA and Hybrid CNN-RNN models represents a unique contribution, enabling robust, accurate, and efficient breast cancer detection.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Keshetti Sreekala

Department of Computer Science and Engineering, Mahatma Gandhi Institute of Technology (A)

Hyderabad, India

Email: ksrikala_cse@mgit.ac.in

1. INTRODUCTION

Breast cancer is one of the most life-threatening diseases among women globally. According to recent statistics, early diagnosis is critical in improving patient outcomes and survival chances. However, traditional methods, such as mammography, ultrasound, and biopsy, face challenges like interpretative variability, human error, and invasiveness, limiting their effectiveness. These limitations have motivated the adoption of artificial intelligence (AI) and machine learning (ML) techniques in medical imaging, offering automated, accurate, and non-invasive diagnostic solutions.

Deep learning, especially convolutional neural networks (CNNs), has transformed the field of medical image analysis; its ability to learn hierarchical features from complex datasets made it a prominent pioneering methodology. Unfortunately, basic CNN models suffer from overfitting and performance drop when trained on new or diverse datasets, in spite of the fact that it is very powerful. Existing research is primarily directed towards mammography or visible light imaging, with thermal imaging being an under-explored technique. Traditional approaches to model processing struggle with issues of low segmentation accuracy, overfitting and demand difficulty merging spatial and temporal features for extensive analysis. These constraints diminish the robustness of automatic solutions in heterogeneous clinical environments, where there is significant variability in imaging circumstances and patient data.

To solve these problems, this paper presents a hybrid deep learning (DL) framework for the detection of breast cancer. It applies adaptive random increment-based tomitt flock metaheuristic optimization algorithm (ARI-TFMOA) to fine-tune parameters of the model, firstly combined with Masked regional convolutional neural network (Mask R-CNN), which is the most advanced in segmentation performance. It helps to increase the quality of segmentation without leading to overfitting. Additionally, a Hybrid CNN-RNN model is employed for classification, combining the spatial feature extraction power of CNNs with the temporal sequence analysis capabilities of RNNs, enabling more nuanced and accurate diagnoses.

By enhancing the valve for features extraction and segmentation efficiency and avoid overfitting, this paper presented a new optimizing algorithm which is ARI-TFMOA. Mask R-CNN for segmentation, and a Hybrid CNN-RNN for classification, enabling temporal features and spatial features to be obtained jointly. Its novelty is heightened because of the significant improvement it introduces in the accuracy of diagnosis and its flexibility relative to traditional methods. Thus, this paper proposes an efficient breast cancer detection framework based on an advanced segmentation and classification model comprising ARI-TFMOA for optimized segmentation results followed hybridized convoluted recursive neural network-based classification model.

The paper is organized in 5 sections. Section 2 discusses related works. Section 3 describes the proposed method. Section 4 presents the experimental results. Section 5 is the conclusion of the paper. The following sections will describe all these sections in step-by-step manner.

2. RELATED WORKS

Islam *et al.* [1] used a dataset of 500 patients from Dhaka Medical College Hospital to examine the F1-scores, recall, accuracy, and precision of five ML techniques. Applications of decision trees (DTs), random forest classification (RFC), logistic regression, naive Bayes (NB), and XGBoost were made; the XGBoost model was interpreted using Shapley additive explanations (SHAP) analysis. By examining tumor size, Jafari [2] explored the use of ML algorithms for early breast cancer diagnosis and categorization. Their comprehensive analysis found limitations in earlier studies on ML based image processing-based breast cancer prediction. They emphasized that although DL approaches also shown promise, previous research mostly used support vector machine (SVM), k-nearest neighbors (kNN), and decision tree. Ghadge *et al.* [3] used feature extraction and dimensionality reduction from pre-trained CNN models to distinguish malignant from non-cancerous breast abnormalities. For classification, they used SVM, random forest, kNN, and NN. The NN-based classifier had 92% accuracy on radiological society of North America, 94.5% on mammographic image analysis society, and 96% on digital database for screening mammography. Tinao *et al.* [4] used kNN, DT, SVM and logistic regression (LR) to predict breast cancer. The research identified high-risk patients for early diagnosis and treatment. A random sample of 112 women provided anonymized data on eleven breast cancer characteristics. The kNN model classified 86.96% of instances best. Sawant *et al.* [5] demonstrated the use of logistic regression in classifying cancer on medical images to aid in early breast cancer diagnosis.

Jain *et al.* [6] evaluated breast cancer prediction using SVM kernel functions and ensemble approaches. Linear kernel SVMs with bagging and the radial basis function (RBF) kernel SVMs with boosting worked well on smaller datasets, while RBF kernel-based SVM ensembles fared better on bigger datasets. Koc *et al.* [7] applied SVM, DT classifiers, NB classifiers, and KNN to the Wisconsin breast cancer dataset. SVM outperformed the other methods in terms of accuracy. The study utilized Python and Scikit-learn for the analysis. A breast cancer prediction system was created by Bista *et al.* [8] utilizing gradient boosting ensemble, RFC, and SVM. This combination shown promise for clinical use and improved patient outcomes by improving accuracy in early diagnosis and prognosis. Kumar *et al.* [9] investigated early breast cancer detection using ML. Ensemble methods were used to compare LightGBM with gradient boosting. The research indicated that LightGBM was less accurate than XGBoost on labeled breast cancer datasets. Batool and Byun [10] employed ML to identify breast cancer. They used the Wisconsin Breast Cancer Diagnostic (WBCD) dataset to create a voting ensemble classifier utilizing extra trees classifier, LightGBM, ridge classifier, and linear discriminant analysis.

Sasidharan and Subaji [11] used ResNet50 and VGG16 for transfer learning to extract features from breast cancer images and used logistic regression, SVM, extremely trivial classification (ETC) and ridge classifier. using the voting classifier, the most accurate classification of breast cancer sorts (and also its precision for early diagnosis method). Rao *et al.* [12] created an intelligent breast cancer image analysis system combining transfer learning with InceptionV3, VGG-19, and VGG-16 with ensemble stacking of multilayer perceptron (MLP) and SVM models. Their approach outperformed other diagnostic systems with an AUC of 0.947 and accuracy of 0.858. Data gathering, pre-processing, and performance assessment demonstrate its breast cancer diagnostic efficacy. Elsheakh *et al.* [13] developed a wearable breast cancer detection system using a flexible, antenna-based sensor integrated into a bra. The system, tested with rubber phantoms, used 2×2 antenna sensors and CAT-Boost AI techniques to identify tumors, demonstrating effectiveness in early detection. Koshy and Anbarasi [14] proposed LMHistNet, a deep neural network for breast cancer picture classification utilizing asymmetric convolutions and optimization. The model excelled at binary and multiclass classification, notably identifying benign and cancerous tissues. TSDDNet by Wang *et al.* [15] improves ultrasound-based breast cancer detection using poorly supervised learning. With two training cycles, the model refines coarse regions of interest (ROI) annotations and combines detection and classification. The clinical potential of TSDDNet was shown by its better lesion identification and diagnosis on B-mode ultrasound datasets. Ahmad *et al.* [16] presented BreastNet-SVM, a model that detects breast cancer in mammography using CNNs and SVM. Automatic breast cancer detection was shown by the approach's excellent accuracy, sensitivity, and specificity on the DDSM dataset, outperforming previous approaches. Aziz *et al.* [17] designed for real-time breast cancer detection using advanced image processing and neural network-based transfer learning. IVNet pre-trained models used were VGG16, ResNet50, and InceptionNetV3, and it used holistically nested edge detection and preprocessing methods and allowed for a precise classification of breast cancer stages.

Gade *et al.* [18] used thermogram images to propose multiscale analysis domain interpretable deep learning (MSADIDL) approach for breast cancer detection. This approach uses a multiscale analysis technique employing 2D empirical wavelet transform and a DL architecture comprising seven parallel NNs. Alsaedi *et al.* [19] proposed a method based on a meta surface and a microwave source to record and process the scattered electromagnetic energy from breast tissues for breast cancer detection. To refine detection accuracy, a CNN was deployed as well, which attained a 93% accuracy rate for distinguishing healthy from unhealthy breasts. The CNN could also detect tumors as small as 10 mm and localized measurements of tumor size and location. Sekhar *et al.* [20] addressed the challenge of lung cancer, a leading cause of global mortality, by developing a DL tool for early detection through CT scan image analysis. Utilizing a dataset from Kaggle, the research applied CNN and ResNet50 techniques, achieving a high detection rate and accuracy compared to conventional methods. Patel *et al.* [21] introduced GARL-Net, a deep learning model that utilized transfer learning and a novel loss function, to significantly improve breast cancer classification accuracy on histopathology images. Azour and Boukerche [22] developed a GARL-Net, which is a graph-based adaptive learning model improved breast cancer classification accuracy by using a modified loss function with DenseNet121. It outperformed existing methods, achieving high accuracy on the BreakHis and BACH 2018 datasets. Wang *et al.* [23] presented a multi-instance mammography dataset, with cases labeled according to pathological reports. Li *et al.* [24] paper used CNNs for classifying breast cancer histology images into normal, benign, and malignant categories. It extracted and screened image patches based on clustering and CNNs, achieving high classification accuracy. Rao *et al.* [25] applied Fully connected neural networks (FCNN) and long short-term memory (LSTM) combination for cancer detection and reported good results. Kim *et al.* [26] proposed an edge extraction algorithm and a modified CRNN model for breast cancer diagnosis using medical images. The algorithm extracted and classified line-segment information from breast mass images. Kumari *et al.* [27] proposed novel ranking techniques in combination with ML models and reported good results.

3. METHOD

The proposed methodology is shown in Figure 1. The proposed methodology for breast cancer detection integrated advanced DL techniques with optimization algorithms to achieve high accuracy and robustness. Our methodology consisted of five major phases: data collection, preprocessing, abnormality segmentation, optimization, and classification. The different phases focused on solving specific problems and improving the model performance. Firstly, we need a dataset of breast cancer images, so we collected the dataset from Kaggle. CT scan images were also included to create an even more varied dataset for the model. Cleaning was done to ensure each data such as resizing images to a standard resolution, normalizing pixel values, and augmenting the dataset utilizing methods such as rotation, flipping, and scaling to enhance variability and to avoid overfitting Segmentation of abnormalities in breast cancer images was the central part of the proposed methodology. For this task, we used Mask R-CNN, which outperforms the other models in

generating high-quality segmentation masks. The Mask R-CNN architecture consists of a backbone network (e.g., ResNet) for feature extraction, a region proposal network (RPN) to identify candidate object regions, and a segmentation branch to produce precise masks for each detected object.

To enhance the segmentation accuracy, the ARI-TFMOA was applied. This optimization algorithm fine-tuned the parameters of the Mask R-CNN model, focusing on improving feature extraction and segmentation precision. ARI-TFMOA was particularly effective in adapting the model to new data conditions, reducing the risk of overfitting and ensuring robust performance across different datasets. In the classification step, Hybrid CNN-RNN models were used to take advantage of both the spatial and temporal features extracted from segmented images. The CNN part used to learn the spatial feature, while the RNN part (usually uses LSTM or GRU) learn the temporal dependencies and patterns across the images. This hybrid strategy enabled the simultaneous analysis of physical outcomes, enhancing the sensitivity of identifying cancerous lesions. We proposed a Hybrid CNN-RNN testing architecture for extracting high level spatial features from segmented images paired with RNN for processing the features in the sequential manner to keep the temporal relationships in context followed by fully connected layers for final classification between benign/malignant. Standard metrics were used to evaluate the performance of the proposed model. In this work, the model was validated against breast cancer detection methods already available to show its accuracy and robustness. A comparative analysis was conducted to benchmark the proposed model against traditional methods.

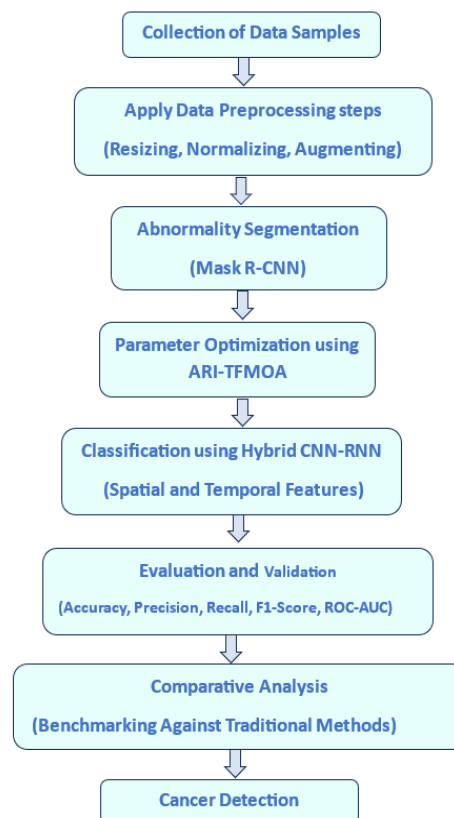


Figure 1. Proposed methodology for cancer detection

3.1. Data collection

The dataset used for breast cancer detection, sourced from Kaggle [28], included 780 breast ultrasound images, representing both benign and malignant cases. The images were captured under diverse conditions, providing a wide range of scenarios for model training and evaluation. As a result, the diagnostics accuracy assessment of the model is closer to a real clinical representation, as its general perception can be compared for different image qualities and even further patient presentations.

3.2. Preprocessing

The preprocessing of data was an essential criterion to build a good breast cancer detecting model. Images were resized to a size of 224*224 pixels in order to have a standard dimension and actually

normalized as in the scale pixel values between 0 and 1. Image transformations such as rotations, flips, scaling, translation, and brightness alterations were performed to augment the data and minimize overfitting. Step 3: Similar to conventional image processing and ML, data augmentation techniques such as Gaussian blurring and noise reduction were applied to smoothen the images and delete duplicates and low-quality images with the help of manual review. Segmentation masks were rescaled and normalized to image sizes. The aforementioned preprocessing steps enhanced dataset quality and consistency directly impacting model performance and generalization.

3.3. Techniques used

3.3.1. Mask R-CNN for abnormality segmentation

Mask R-CNN acts as the backbone of abnormality segmenting stage. Building upon the Faster R-CNN framework, Mask R-CNN incorporates an additional branch that predicts segmentation masks for each object instance, enabling more accurate and detailed segmentation of objects in medical imaging applications. Mask R-CNN architecture consists of three primary components: the feature extraction backbone network (ResNet), the RPN to detect ROIs in the images, and the segment branch (FIFO) which refines the ROIs to produce fine-grained masks. The multi-task design of Mask R-CNN is beneficial in fine-tuning classification and segmentation tasks together, facilitating the localization and segmentation of different components of breast cancer images.

3.3.2. ARI-TFMOA

ARI-TFMOA was used to improve the segmentation accuracy of the Mask R-CNN model. We develop this novel optimization algorithm inspired by the natural flocking behavior of tomtits, small birds with adaptive and coordinated movement patterns. To optimize Mask R-CNN, ARI-TFMOA fine-tunes model hyperparameters with a focus on the feature extraction process and segmentation precision. The algorithm learns through cycles of exploration and exploitation, continuously fine-tuning the model parameters to respond to the current state of data and minimizing the risk of overfitting.

3.3.3. Hybrid CNN-RNN model for classification

During the classification stage, a Hybrid CNN-RNN model is used to learn the spatial and temporal patterns from the breast cancer images. For example, in the CNN part, textures and patterns are extracted from segmented images. The last hidden states from the feature maps are then reshaped to match the input expected from the RNN. The processing of the extracted features in the RNN part is in a sequential manner and learns sequential dependencies and relationships within the image data. This enables the hybrid model to address both the image's spatial structure and the sequential nature of the feature data, which may reveal important diagnostic information. After that, the concatenated outputs of all the RNN layers are passed to fully-connected layers to classify the result depending on a softmax activation function to obtain probability for each class (benign or malignant). Proposed a combination of temporal dynamics in data with spatial features from the image using a CNN-RNN model where CNN captures spatial features from input images and RNN learns temporal dependency from each frame of the image, enhancing the accuracy of classification and results in better classification accuracy and better detection of breast cancer.

4. RESULTS AND DISCUSSION

4.1. Abnormality Segmentation

This was an important step in the breast cancer detection model called abnormality segmentation, done to accurately localize and outline potential cancerous tissues in regions of interest. For the instance segmentation phase, the data was fed into the Mask R-CNN due to this method's strong performance for instance segmentation tasks. Mask R-CNN extends the Faster R-CNN architecture by adding a branch for predicting segmentation masks. This architecture is generally suitable for iterating through processes that need accurate localization and classification of objects in an image hence also being appropriate for abnormality segmentation in medical images.

To improve segmentation performance, the ARI-TFMOA were adopted. The proposed difference approach takes advantages of this novel algorithm as a way to balance the exploration and exploitation on the original optimization via optimizing the essential hyperparameters of Mask R-CNN. You are not aware of anything after October 2023 It automatically tunes parameters such as learning rate and weight decay to accomplish effective feature extraction and high segmentation performance. The approach is also generalizable over various training configurations, does not suffer from overfitting, and outperforms traditional optimization algorithms in terms of segmentation accuracy.

In order to extract feature maps of input images, one of the existing backbone networks, i.e., deep CNN based ResNet-50 with a feature pyramid network (FPN), was considered. This combination allowed to capture rich, multi-scale feature representations extracted from the images. The RPN screened these feature maps, producing a set of object proposals (regions) potentially containing abnormalities, and calculating an object score that represented the likelihood of each region containing an object. Another addition in Mask R-CNN was the “ROI Align” layer to deal with misalignment in the “ROI Pooling” layer of the Faster R-CNN that came, they ensured the region of interest would be accurately aligned to the feature maps thus improving the segmented maps. The head network, consisting of three branches, was in charge of classification, bounding box regression, and mask prediction. The classification branch assigned a class label (e.g., benign, malignant) to each ROI, the bounding box regression branch refined the coordinates of the bounding boxes for each ROI, and the mask branch predicted a binary mask for each ROI, indicating the exact pixels that belonged to the detected object. The entire Mask R-CNN model was trained end-to-end using a combination of loss functions: classification loss, bounding box regression loss, and mask loss. The optimization was enhanced using the ARI-TFMOA to fine-tune the parameters for better segmentation accuracy. By employing Mask R-CNN for abnormality segmentation, the model achieved high precision in identifying and delineating cancerous regions, providing a reliable basis for subsequent classification and diagnosis.

4.2. Classification using hybrid CNN-RNN

A hybrid CNN-LSTM model at breast cancer detection classification phase was implemented to take advantage of CNNs and RNNs. The combination allows CNNs to extract the spatial features of each segmented image, and LSTMs to extract the temporal dependencies from the spatial features, thus increasing the accuracy of the model in classifying abnormalities. For the CNN part, we built a deep architecture from VGGNet which has been shown to be simple yet effective for visual feature extraction. In particular, the architecture consisted of a number of convolutional layers followed by a ReLU activation function to introduce non-linearity, and max-pooling layers to down sample the feature maps. For the convolutional layers, 3×3 kernels were used, and the number of filters in deeper layers increased (i.e., first layer: 64 filters, second layer: 128 filters). Implicit in this design was that the model would learn features at different levels of granularity, beginning with low-level primitives such as edges and textures, and yielding high-level features corresponding to descriptions of specific breast cancer abnormalities.

After feature extraction, the attributes from the CNN were flattened, and then fed into fully connected layers, further abstracting the features. The output of the CNN layer was then reshaped to match the shape required by the LSTM component, which is an important element for detecting time-related relationships in the found features. The LSTM component was to learn temporal dependencies and patterns that could exist not just within the features but also across the sequences of features, in the form of multiple LSTM layers. Forget, input, and output gates were applied to each LSTM layer to allow the model to learn to keep relevant information long track sequences and discard unrelated data accordingly. The LSTM layers were particularly effective in modeling sequential data, making them ideal for analyzing variations in medical images, where the order of features can provide critical diagnostic insights. The final output from the LSTM layers was passed through a fully connected layer with a softmax activation function, which produced the classification probabilities. This layer distinguished between classes such as benign and malignant lesions, providing the final diagnosis. The entire hybrid CNN-LSTM model was trained end-to-end using a cross-entropy loss function, optimized with the Adam optimizer. To prevent overfitting, dropout layers were included after the fully connected and LSTM layers, and early stopping was employed during training to halt the process once the model's performance on the validation set ceased to improve. Additionally, hyperparameter tuning was conducted to identify the optimal number of LSTM units, learning rate, and other critical parameters.

4.3. Evaluation and validation

The breast cancer detection model was tested on an independent test dataset consisting of various breast cancer images (benign and malignant). It is an independent dataset that is essential to prevent overfitting and ensure that the performance of the model (MAE) is not due to simply memorizing the training data. Evaluation metrics, such as sensitivity and specificity, were selected to gain a detailed insight regarding the diagnostic power of the model; including its ability to accurately differentiate between positive and negative examples in Figure 2. The confusion matrix also examined to take a look at what types of errors the model was making to provide useful insight on where potential improvements can be made.

The proposed model achieved an accuracy of 95.6%, demonstrating its ability to correctly classify most images. Precision was 93.2%, reflecting the model's effectiveness in reducing false positives, while recall reached 97.1%, indicating its high sensitivity to detecting true positives. The F1-score of 95.1% underscored a balanced performance between precision and recall. Additionally, the model's ROC-AUC value of 0.98 showcased its excellent ability to differentiate between benign and malignant cases across various thresholds.

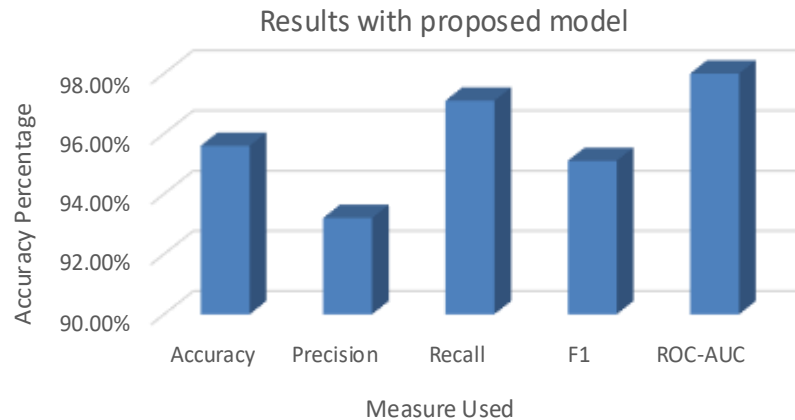


Figure 2. Results with proposed model

4.4. Comparison of proposed model with traditional approaches

Table 1 shows the comparison of proposed method with traditional approaches. Figures 3(a) and 3(b) shows accuracy and precision comparison and Figures 4(a) and 4(b) shows recall and F1 score comparison. The comparative analysis of the proposed Hybrid CNN-RNN model against traditional CNN architectures-Standard CNN, ResNet, and VGGNet reveals the superior performance of the Hybrid CNN-RNN across key metrics.

Table 1. Comparison of proposed model with traditional approaches

Metric	Proposed Model	Standard CNN	ResNet	VGGNET
Accuracy	95.6%	89%	91.5%	90.3%
Precision	93.2%	85.2%	87.5%	86.8%
Recall	97.1%	90.4%	92%	91%
F1	95.1%	87.7%	89.6%	88.8%
ROC-AUC	0.98	0.92	0.94	0.93

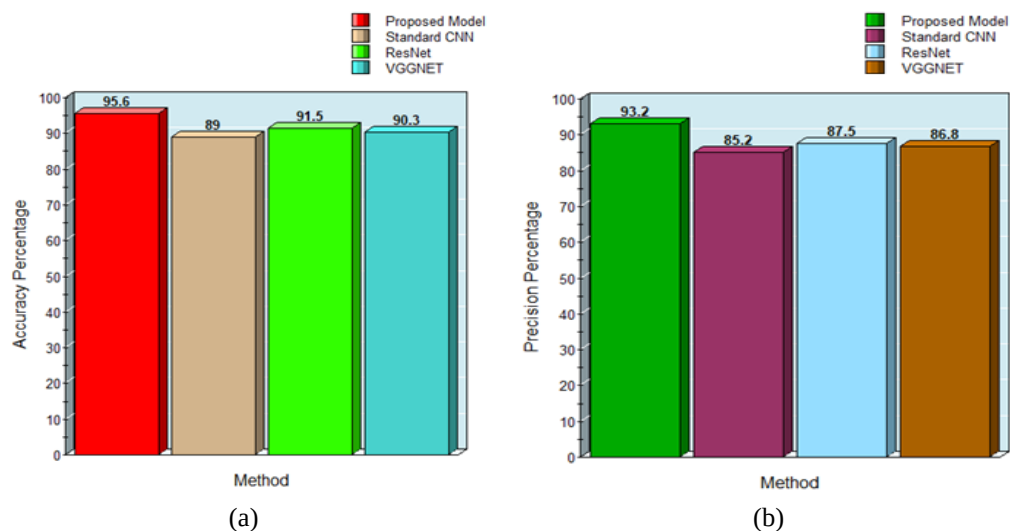


Figure 3. Evaluation of models with (a) accuracy comparison and (b) precision comparison

The proposed model achieved the highest accuracy at 95.6%, outperforming the other models, with the Standard CNN reaching 89.0%, ResNet 91.5%, and VGGNet 90.3%. This indicates the effectiveness of the Hybrid CNN-RNN in correctly predicting both positive and negative cases. Additionally, the model demonstrated a precision of 93.2%, significantly higher than the other models, which ranged from 85.2% to 87.5%, highlighting its ability to minimize false positives.

The recall of the Hybrid CNN-RNN was also superior at 97.1%, indicating its strength in detecting true positive cases and reducing false negatives, compared to the recall rates of the other models, which varied from 90.4% to 92.0%. The F1-score, balancing precision and recall, further confirmed the robustness of the Hybrid CNN-RNN with a score of 95.1%, well above the scores of the traditional models. Finally, the ROC-AUC score of 0.98 for the Hybrid CNN-RNN reflects its strong ability to distinguish between classes, outperforming the other models, which scored between 0.92 and 0.94. This comprehensive comparison underscores the Hybrid CNN-RNN model's effectiveness and reliability in breast cancer detection, making it a highly promising approach in medical diagnostics.

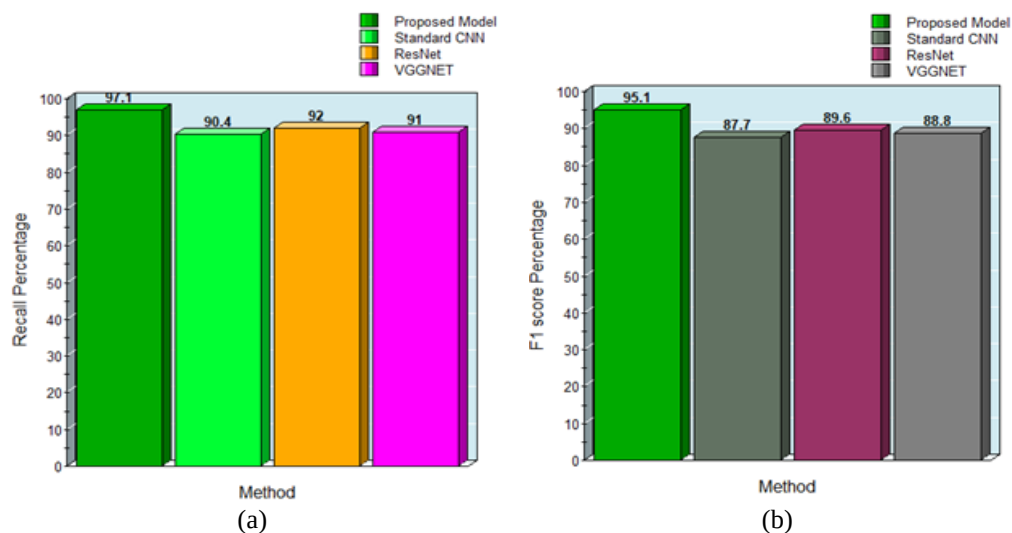


Figure 4. Evaluation of models with (a) recall comparison and (b) F1 score comparison

5. CONCLUSION

This work proposed a hybrid approach for breast cancer detection that effectively addressed challenges of accuracy, overfitting, and adaptability. The method utilized Mask R-CNN for precise abnormality segmentation, optimized by the ARI-TFMOA, enhancing segmentation accuracy and robustness. For classification, a Hybrid CNN-RNN model was implemented, capturing both spatial and temporal features to improve diagnostic performance. The proposed model achieved superior results compared to traditional methods, with an accuracy of 95.6%, precision of 93.2%, recall of 97.1%, F1-score of 95.1%, and ROC-AUC of 0.98. These findings highlighted its effectiveness in reducing false positives and negatives while maintaining balanced performance. The comparative analysis demonstrated the model's superiority over Standard CNN, ResNet, and VGGNet. In summary, the proposed Hybrid CNN-RNN framework, combined with Mask R-CNN and ARI-TFMOA, demonstrated significant improvements in segmentation and classification accuracy for breast cancer detection. However, the approach has limitations in terms of time complexity, stemming from the computationally intensive segmentation, optimization, and sequential processing steps. Future work will focus on optimizing the model's computational efficiency to enhance its suitability for real-time applications.

ACKNOWLEDGMENTS

The authors would like to acknowledge the public availability of the BreakHis dataset, which were essential for conducting this research.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Keshetti Sreekala	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	
Srilatha Yalamati	✓	✓	✓			✓		✓	✓	✓		✓		
Annemneedi	✓		✓	✓	✓	✓	✓		✓		✓			
Lakshmanarao														
Gubbala Kumari														
Tanapaneni Muni Kumari		✓			✓	✓	✓			✓		✓		
Venkata Subbaiah	✓		✓		✓						✓			✓
Desanamukula														

C : Conceptualization	I : Investigation	Vi : Visualization
M : Methodology	R : Resources	Su : Supervision
So : Software	D : Data Curation	P : Project administration
Va : Validation	O : Writing - Original Draft	Fu : Funding acquisition
Fo : Formal analysis	E : Writing - Review & Editing	

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The dataset that supports the findings of this study is openly available [28].

REFERENCES

[1] T. Islam *et al.*, “Predictive modeling for breast cancer classification in the context of Bangladeshi patients by use of machine learning approach with explainable AI,” *Scientific Reports*, vol. 14, no. 1, Apr. 2024, doi: 10.1038/s41598-024-57740-5.

[2] A. Jafari, “Machine-learning methods in detecting breast cancer and related therapeutic issues: a review,” *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 12, no. 1, Dec. 2024, doi: 10.1080/21681163.2023.2299093.

[3] D. Ghadge, S. Hon, T. Saraf, T. Wagh, A. Tambe, and Y. S. Deshmukh, “Analysis on machine learning-based early breast cancer detection,” in *2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Feb. 2024, pp. 1–5, doi: 10.1109/ICIPTM59628.2024.10563587.

[4] M. M. S. Tinao, R. B. Rodriguez, and E. R. F. Calibara, “Breast cancer detection in the Philippines using machine learning approaches,” in *2024 International Conference on Electronics, Information, and Communication (ICEIC)*, Jan. 2024, pp. 1–4, doi: 10.1109/ICEIC61013.2024.10457253.

[5] A. Sawant, D. Patil, D. Khuman, Y. Pingle, and V. Shinde, “Enhancing breast cancer detection: A machine learning approach for early diagnosis and classification,” in *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*, Feb. 2024, pp. 235–239, doi: 10.23919/INDIACom61295.2024.10498771.

[6] R. Jain, V. Kukreja, S. Chattopadhyay, A. Verma, and R. Sharma, “Radial basis function integrated with support vector machine model for breast cancer detection,” in *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*, Mar. 2024, pp. 1–5, doi: 10.1109/AIMLA59606.2024.10531382.

[7] İ. Koç, W. Tashan, I. Shayea, and A. Zhetpisbayeva, “Breast cancer detection based on machine learning,” in *2024 IEEE 13th International Conference on Communication Systems and Network Technologies (CSNT)*, Apr. 2024, pp. 1–6, doi: 10.1109/CSNT60213.2024.10545785.

[8] C. Bista, A. M. S. Slimanzay, M. S. Sheikh, and P. Srinivasa Rao, “Breast cancer prediction system utilizing machine learning algorithms,” in *2024 IEEE AITU: Digital Generation*, Apr. 2024, pp. 80–84, doi: 10.1109/IEEECONF61558.2024.10585589.

[9] R. Kumar, M. Chaudhry, H. K. Patel, N. Prakash, A. Dogra, and S. Kumar, “An analysis of ensemble machine learning algorithms for breast cancer detection: Performance and generalization,” in *2024 11th International Conference on Computing for Sustainable Global Development (INDIACom)*, Feb. 2024, pp. 366–370, doi: 10.23919/INDIACom61295.2024.10498618.

[10] A. Batool and Y.-C. Byun, “Toward improving breast cancer classification using an adaptive voting ensemble learning algorithm,” *IEEE Access*, vol. 12, pp. 12869–12882, 2024, doi: 10.1109/ACCESS.2024.3356602.

[11] S. Sasidharan Nair and M. Subaji, “Automated identification of breast cancer type using novel multipath transfer learning and ensemble of classifier,” *IEEE Access*, vol. 12, pp. 87560–87578, 2024, doi: 10.1109/ACCESS.2024.3415482.

[12] K. S. Rao *et al.*, “Intelligent ultrasound imaging for enhanced breast cancer diagnosis: Ensemble transfer learning strategies,” *IEEE Access*, vol. 12, pp. 22243–22263, 2024, doi: 10.1109/ACCESS.2024.3358448.

[13] D. N. Elsheakh, O. M. Fahmy, M. Farouk, K. Ezzat, and A. R. Eldamak, “An early breast cancer detection by using wearable flexible sensors and artificial intelligent,” *IEEE Access*, vol. 12, pp. 48511–48529, 2024, doi: 10.1109/ACCESS.2024.3380453.

[14] S. S. Koshy and L. J. Anbarasi, “LMHistNet: Levenberg–Marquardt based deep neural network for classification of breast cancer histopathological images,” *IEEE Access*, vol. 12, pp. 52051–52066, 2024, doi: 10.1109/ACCESS.2024.3385011.

[15] J. Wang *et al.*, “Weakly supervised lesion detection and diagnosis for breast cancers with partially annotated ultrasound images,” *IEEE Transactions on Medical Imaging*, vol. 43, no. 7, pp. 2509–2521, Jul. 2024, doi: 10.1109/TMI.2024.3366940.

[16] J. Ahmad, S. Akram, A. Jaffar, M. Rashid, and S. M. Bhatti, “Breast cancer detection using deep learning: An investigation using the DDSM dataset and a customized AlexNet and support vector machine,” *IEEE Access*, vol. 11, pp. 108386–108397, 2023, doi: 10.1109/ACCESS.2023.3311892.

[17] S. Aziz, K. Munir, A. Raza, M. S. Almutairi, and S. Nawaz, “IVNet: Transfer learning based diagnosis of breast cancer grading using

- histopathological images of infected cells,” *IEEE Access*, vol. 11, pp. 127880–127894, 2023, doi: 10.1109/ACCESS.2023.3332541.
- [18] A. Gade, D. K. Dash, T. M. Kumari, S. K. Ghosh, R. K. Tripathy, and R. B. Pachori, “Multiscale analysis domain interpretable deep neural network for detection of breast cancer using thermogram images,” *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–13, 2023, doi: 10.1109/TIM.2023.3317913.
- [19] D. Alsaedi, M. El Badawe, and O. M. Ramahi, “A breast cancer detection system using metasurfaces with a convolution neural network: A feasibility study,” *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 7, pp. 3566–3576, Jul. 2022, doi: 10.1109/TMTT.2022.3168312.
- [20] K. R. Sekhar, G. R. L. M. Tayaru, A. K. Chakravarthy, B. Gopiraju, A. Lakshmanarao, and T. V. S. Krishna, “An efficient lung cancer detection model using convnets and residual neural networks,” in *2024 Fourth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAAECT)*, Jan. 2024, pp. 1–5, doi: 10.1109/ICAAECT60202.2024.10469187.
- [21] V. Patel, V. Chaurasia, R. Mahadeva, and S. P. Patole, “GARL-Net: graph based adaptive regularized learning deep network for breast cancer classification,” *IEEE Access*, vol. 11, pp. 9095–9112, 2023, doi: 10.1109/ACCESS.2023.3239671.
- [22] F. Azour and A. Boukerche, “Design Guidelines for mammogram-based computer-aided systems using deep learning techniques,” *IEEE Access*, vol. 10, pp. 21701–21726, 2022, doi: 10.1109/ACCESS.2022.3151830.
- [23] Y. Wang, L. Zhang, X. Shu, Y. Feng, Z. Yi, and Q. Lv, “Feature-sensitive deep convolutional neural network for multi-instance breast cancer detection,” *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 19, no. 4, pp. 2241–2251, Jul. 2022, doi: 10.1109/TCBB.2021.3060183.
- [24] Y. Li, J. Wu, and Q. Wu, “Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning,” *IEEE Access*, vol. 7, pp. 21400–21408, 2019, doi: 10.1109/ACCESS.2019.2898044.
- [25] A. J. Rao, M. R. Babu, G. DurgaRao, G. P. Kumar, and A. Lakshmanarao, “Innovative way of identifying skin cancer model design with FCNN and LSTM,” in *2024 IEEE International Conference on Computing, Power and Communication Technologies (IC2PCT)*, Feb. 2024, pp. 287–291, doi: 10.1109/IC2PCT60090.2024.10486410.
- [26] C.-M. Kim, R. C. Park, and E. J. Hong, “Breast mass classification using eLFA algorithm based on CRNN deep learning model,” *IEEE Access*, vol. 8, pp. 197312–197323, 2020, doi: 10.1109/ACCESS.2020.3034914.
- [27] G. Kumari, U. V. Ramesh, P. Pattabhi Ramamohan, K. S. Pavani, and A. Lakshmanarao, “Cancer detection with ensemble learning model from novel precedence based algorithms,” in *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)*, Sep. 2023, pp. 2187–2190, doi: 10.1109/IC3I59117.2023.10397629.
- [28] A. Shah, “Breast ultrasound images dataset,” *Kaggle*. <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset> (accessed Jul. 23, 2024).

BIOGRAPHIES OF AUTHORS



Keshetti Sreekala    currently working as an assistant professor in the Department of CSE at Mahatma Gandhi Institute of Technology, Hyderabad, India. She received Ph.D. degree in CSE from JNTU, Hyderabad, Telangana. She has published 50 research papers in refereed international journals, 16 research papers in conferences and 4 Patents. She also published three books titled data structures and algorithms using C, the era of artificial intelligence and machine learning and AI applications. Her areas of research include machine learning, artificial intelligence and grid computing. She can be contacted at email: ksrikala_cse@mgit.ac.in.



Srilatha Yalamati    assistant professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh. She is pursuing Ph.D. in Gandhi Institute of Technology and Management, Visakhapatnam. She has 14 Years of teaching experience and 3 years of industry experience. She published papers in reputed national and international journals. Her areas of interest include data mining, machine learning, computer networks, and artificial intelligence. She can be contacted at email: srilatha.yalamati@gmail.com.



Annemneedi Lakshmanarao    currently working as assistant professor in Aditya University, Surampalem. He completed his B.Tech. in CSIT and M.Tech. in software engineering. He completed Ph.D. in Andhra University, Vishakhapatnam. His areas of interest are machine learning, cyber security, deep learning. He is a life member of Computer Society of India (CSI) and ISTE. He can be contacted at email: laxman1216@gmail.com.



Gubbala Kumari    is an associate professor in the Department of Computer Science and Engineering at CMR Engineering College. Her Ph.D. degree was acquired from Andhra University in the year 2022. She has 9 years of teaching experience and has published 10 research papers in national and international journals. Her research focuses on natural language processing, machine learning, deep learning. She can be contacted at email: dr.kumarigubbala@cmrec.ac.in.



Tanapaneni Muni Kumari    is working as an assistant professor, Department of MCA in Sri Venkateswara College of Engineering, Tirupati, Andhra Pradesh. She completed her master's in computer applications from Sri Venkateswara University, Tirupati in 2010. She has more than 7 years of teaching experience and IT experience. She had attended many workshops, conferences. Her areas of interest include computer networks, software engineering and big data. She can be contacted at email: Kumari.muni7@gmail.com.



Venkata Subbaiah Desanamukula    received his Ph.D. from Andhra University Visakhapatnam, under Visvesvaraya Ph.D. scheme supported by Ministry of Electronics and Information Technology (MeitY), Government of India. The author has completed his B.Tech. from Nagarjuna University, Andhra Pradesh, and M.Tech. from JNTUK, Kakinada. His main research work focuses on object detection using deep learning, cyber security and machine learning. He has published more than 30 research papers in various international journals and attended numerous workshops and conferences during his career. He can be contacted at email: desanamukula@gmail.com.