An AI-Based Method for Automated Breast Cancer Detection and Localization in Mammogram and Ultrasound Images

Pradeep Kumar
Department of ECE, VNR VJIET, Bachupally, Hyderabad, 500090, Telangana, India
pradeepkumar@vnrvjiet.in

Ranjan K. Senapati
Department of ECE, VNR VJIET, Bachupally, Hyderabad, 500090, Telangana, India
ranjan_ks@vnrvjiet.in (corresponding author)

Prasanth Mankar
Department of ECE, Vasavi College of Engineering, Ibrahimbagh, Hyderabad, Telangana, India
m.prasanth@staff.vce.ac.in

Santosh Kumar Choudhary
Department of ECE, VNR VJIET, Bachupally, Hyderabad, 500090, Telangana, India
santoshkumar_ch@vnrvjiet.in

P. M. K. Prasad
Department of ECE, GVP College of Engineering for Women, Vishakapatnam, Andhra Pradesh, India
pmkp70@gmail.com

Satish Muppidi
Department of Information Technology, GMR Institute of Technology, Rajam, 532127, Andhra Pradesh, India
satish.m@gmrit.edu.in

Gandharba Swain
Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India
gswain1234@gmail.com

Received: 6 May 2025 | Revised: 2 July 2025 | Accepted: 19 July 2025

 DOI: https://doi.org/10.48084/etasr.11951

ABSTRACT

This research aims to improve the identification, classification, and segmentation of breast cancer using mammography and ultrasound images using a Refined Mask-RCNN framework. An OR operation merges the results, reducing misclassification and unnecessary biopsies. The system is deployed in the NVIDIA Jetson Nano developer kit to assist doctors and radiologists in the detection of cancerous cells. Compared to the original Mask-RCNN, the proposed method performs better in cancer identification and segmentation, showing improved metrics such as accuracy, precision, True Positive Rate (TPR), True Negative Rate (TNR), F-score, Balanced Classification Rate (BCR), Youden's index, Jaccard, and Dice coefficient, demonstrating the robustness and reliability of the combined approach in the clinical diagnosis of breast cancer.

Keywords-refined mask-RCNN; breast ultrasound; mammography; breast cancer; balanced classification rate; NVIDIA Jetson nano

I. INTRODUCTION

Early disease diagnosis reduces mortality and facilitates a quick cure with minimal human effort. In most cases, if a condition is not identified in its early stages, it progresses to a chronic level. Consequently, the global increase in mortality can be associated with inadequate early detection of the disease. Breast cancer is the leading cause of mortality for women worldwide [1]. Studies by the American Studies Society and the Indian Association of Cancer Research (IACR) reported 17.1 million cancer cases worldwide in 2018. By 2040, 27.5 million people are estimated to have cancer, with 16.3 million deaths projected. Early detection can help limit the number of deaths caused by breast cancer [2]. Breast cancer accounts for 0.8-1% of cancerous cases in men as well [3], but this study mainly focuses on breast cancer cases in women. Using breast ultrasound and mammograms individually can miss cancer cases, but when combined, they can detect cancer 82% of the time [4]. As a result, both mammograms and ultrasound images are widely utilized to diagnose cancer, as they are low-cost and highly responsive to small masses [5]. The accuracy of the diagnosis in real-world situations depends on several factors, such as the radiologist's fatigue and distraction, the breast's complex anatomy, and the early stages of the disease showing only modest symptoms [6]. In this case, Computer-Aided Diagnosis (CAD) may be more effective for detecting cancer.

In [6], Mask-RCNN was used with mammography images to detect, classify, and segment breast cancer, but with limited accuracy. In [7], a 3D CNN-based CAD system was proposed to identify lesions on Automated Breast Ultrasound (ABUS) images. In [8], a fully automated approach used a deep CNN for the segmentation of dermoscopic images. To enhance performance and avoid the need for class re-balancing, the loss function was based on the Jaccard distance. In [9], a technique for simultaneous mass detection utilized an integrated mammographic CAD based on pseudo-color mammograms. The study in [10] was based on a probabilistic neural network trained using mammogram images for the classification of breast tumors, achieving an accuracy of 80%. In [11], a mask point RCNN was used for instance segmentation. The consistency between the prediction results of segmentation and ground truth masks was enhanced by the addition of edge aggregation point auxiliary detection tasks. In addition, dense deep supervision can be utilized to improve sensitivity, using efficiently discriminative features in multiple layers. In [12], a transfer learning CNN was integrated into the screening process to enhance accuracy. This study employed a fine-tuning approach for dataset classification, training the Inception v3 model on merged datasets, and the resulting output can serve as a valuable second opinion for radiologists. In [13], a large training dataset contained both breast-level and pixel-level labels, and a neural network was proposed for breast cancer screening classification. This study employed Random Forest (RF) for classification based on maximum posterior probability, achieving an accuracy of 97.3%. In [14], an approach for automatic detection, segmentation, and classification of breast masses was proposed using ultrasound images. However, the delineation of lesion regions could be performed more accurately to classify them as malignant or benign. In [15], a CNN-based architecture was proposed to detect cancerous cells using CBIS-DDSM data, along with two segmentation methods. In [16], CNN and Decision Trees (DT) were used to classify breast masses. In [17], an automated model helped to assess patients who can develop breast cancer using mammogram analysis. In [18], a deep learning framework was proposed for the detection of breast cancer using the CBIS-DDSM dataset. Most studies on breast cancer focus mainly on the use of CAD systems to diagnose and classify breast cancer [19-27].

This study proposes a Refined Mask-RCNN to detect, classify, and segment breast cancer using two imaging modalities, ultrasound and mammography. The original Mask-RCNN [9, 20] has been used to solve the multi-instance segmentation problem. State-of-the-art instance segmentation performance can be achieved using Mask-RCNN on general images. Recent studies [10, 20, 21] have shown that Mask-RCNNs can find cancer bounding boxes, but when used for segmentation and detection of medical images, its performance was found to be lower than CNNs. This is primarily due to the original Mask-RCNN framework's semantic representation being too coarse for medical image recognition and segmentation, as it does not sufficiently focus on the edge information in breast cancer images. This study aimed to address these limitations, making the following changes in the original Mask-RCNN to improve the accuracy:

- Optimizes the value of the stride to incorporate information about the aggregation point of the breast cancer edge.
- The Region Proposal Network (RPN) employs a method for generating Regions of Interest (RoIs) that leverage multiscale semantic properties.
- Instead of using smooth L1 regularization in the bounding box loss function, the Renyi Entropy Regularization (RER) can improve the accuracy of the bounding boxes.

The contributions of this work are as follows. The RER-based Refined Mask-RCNN is applied to breast cancer images for detection, classification, and segmentation, expanding beyond the focus of most existing methods. This can help guide biopsies by identifying all cancerous regions, even when multiple sites are affected. The model is deployed in the NVIDIA Jetson Nano developer kit to assist physicians and radiologists in the analysis of mammograms and breast ultrasound images. Enhancements in the original Mask R-CNN improve performance. Extensive experiments were conducted using mammography data from CBIS-DDSM and ultrasound images.

II. MATERIALS AND METHODS

This study proposes a novel method for the detection, classification, and segmentation of breast ultrasound and

mammography images. The proposed approach varies from other existing approaches in the following ways:

1. Proposes three key improvements on the original Mask R-CNN to achieve enhanced output: (a) Optimizes the value of the stride to incorporate information on the aggregation point of the breast cancer edge, (b) Includes a procedure for forming RoIs in the RPN, which can take advantage of multiscale semantic properties, and (c) Uses RER instead of smooth L1 regularization in the loss function of the bounding box. These ensure an increase in the accuracy of bounding boxes and optimize the Mask-RCNN for better performance.
2. A method for combining the results is proposed instead of using individual results from either mammograms or breast ultrasounds. This provides better detection, classification, and segmentation, unlike providing only detection and classification, as earlier methods did.

A. The Proposed Approach

Figure 1 shows the block diagram for the proposed method. The proposed Refined Mask-RCNN is trained using breast ultrasound and mammogram images along with the relevant JSON file. Datasets [28, 29] were split into 8:1:1 for training, validation, and testing, respectively. The training and validation datasets are combined, annotated, and saved as a single file, while the testing images are stored separately without annotations. The trained model was deployed on NVIDIA Jetson Nano. The test mammogram and ultrasound images were passed to Jetson Nano for prediction, and an OR operation was applied to obtain the final result. This approach provides a final output as positive (presence of breast cancer) if any of the imaging modalities is detected positive.

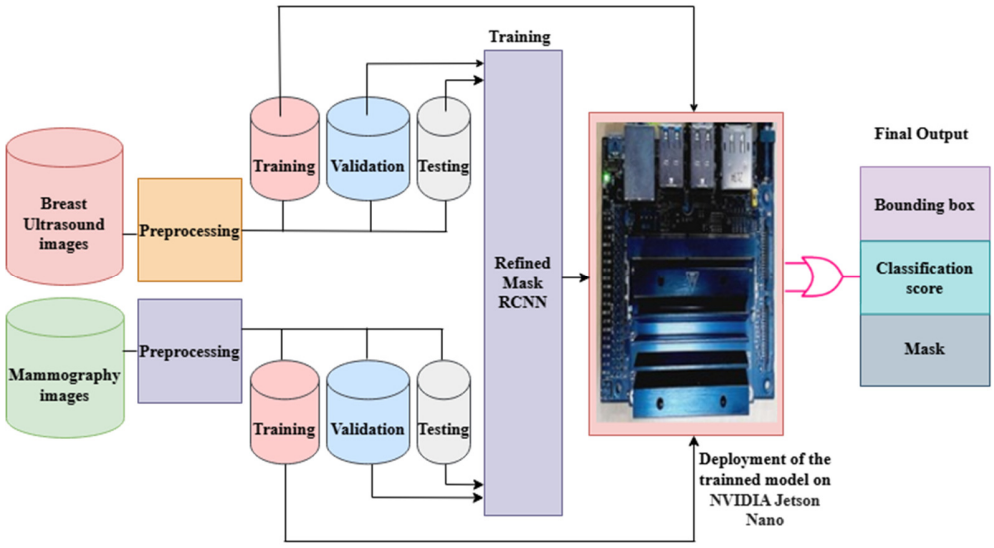

Fig. 1. Proposed classification framework with Refined Masked R-CNN.

B. Data Collection and Preprocessing

The dataset was preprocessed before training the proposed model. As seen in Figure 2, dataset preparation involved (i) collecting datasets, (ii) preprocessing, and (iii) annotation with the VGG-annotator to produce a JSON file. The proposed model was trained, validated, and tested with the Digital Database for Screening Mammograms (DDSM), and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM) for mammography, which is an improved version of DDSM. In the dataset, a positive label indicates malignancy, and a negative one indicates benign. Approximately 14% of the 55,890 images are positive, while the remaining 86% are negative. These datasets are available in the TFRecords format [28]. The TensorFlow library has a data module that was used to extract TFRecords into JPGs. For breast ultrasounds, a dataset was used, categorized into three classes: normal, benign, and malignant [29]. The dataset contains 780 images with an average size of 500×500 pixels in PNG format. After collecting the datasets, preprocessing has to be performed, which is important because random noise can lead to the corruption of medical images. Noise that occurs during the acquisition phase makes automated feature extraction and analysis of clinical data difficult. Different preprocessing techniques were applied because the dominating noise is different in both kinds of imaging modalities. In mammography images, the dominant noise is Poisson noise, and in breast ultrasound images, the dominant noise is Speckle noise. Speckle noise is multiplicative in nature, whereas Poisson noise is signal-dependent. To remove Poisson noise from mammography images, a modified Anscombe transform was used to remove the signal dependency nature of Poisson noise along with a CNN. The optimal Bayesian estimation framework was used to reduce speckle noise in breast ultrasound images [22]. In this method, a logarithmic approach is first used to change the multiplicative noise into additive noise, and then an antilogarithmic approach is used to obtain the original image. After denoising, the images are resized to 1024×1024. Then, the images were uploaded to the VGG annotator for annotation. After the annotation, the JSON file for the annotated images was downloaded.

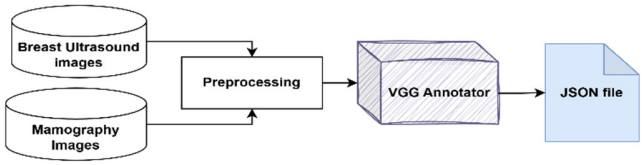

Fig. 2. Block diagram of data preparation.

C. The Proposed Network Architecture

The proposed method was developed based on an existing Mask-RCNN, which is a popular multi-instance segmentation method. Three improvements are made on the original Mask-RCNN: (a) The model incorporates the aggregation of breast cancer edge information by optimizing the stride value, (b) Multiscale semantic properties can be obtained using RPN, which includes a method for creating RoIs, and (c) RER is used instead of smooth L1 regularization in the loss function of the bounding box, increasing accuracy.

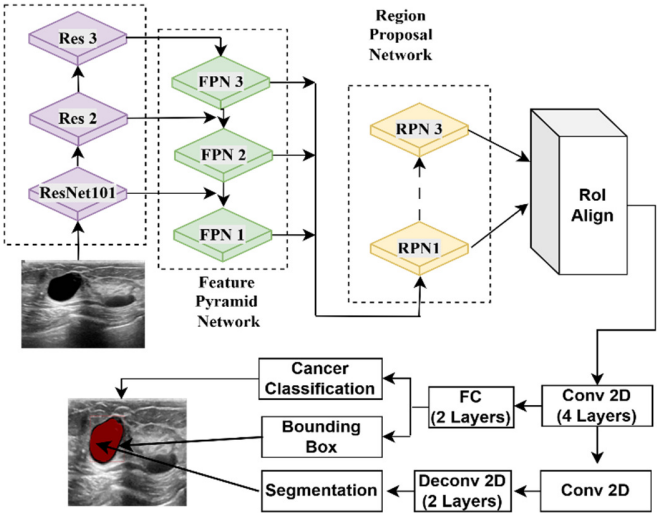

Fig. 3. The proposed network architecture.

The proposed model can be divided into three different modules. The first involves the extraction of features and the calculation of RoI, which mainly consists of ResNet101+FPN +RPN. Initially, this module generates multi-layer maps. To obtain the RoI, each feature map point is mapped onto the original image. The second module, called ROI alignment, pools the RoIs generated from the first module into a fixed size to avoid quantization errors. The third module acquires the mask. Mask generation is performed by a segmentation network that receives the fixed-size RoIs obtained from the second module. The Fully Connected (FC) layer also receives the fixed-size ROIs used for cancerous-position rectangular bounding box regression and classification of breast cancer.

All three operations, i.e., detection, classification, and segmentation, can be performed at the same time, which is the novelty of the proposed Refined Mask-RCNN.

1) Extraction of Key Features and ROI Calculation

In this step, the features of the breast cancer image are extracted, and the corresponding RoI in the feature layer is generated. First, a breast ultrasound or mammography image containing cancerous cells is fed into the ResNet101 network, which is the backbone network to extract features [13]. Then, an FPN [26] is used to fuse these multilayer features and obtain the semantic information. Fusing multiple layers can improve the accuracy of breast cancer detection. This approach provides equally sized features using dimensionality reduction and up-sampling operations. Finally, the prediction of different output layers to obtain RoIs is performed by the RPN.

2) Mask Acquisition

An FC layer receives a fixed-size RoI for classification purposes, which is also sent to a rectangular breast cancer-position bounding box. Simultaneously, a mask generation branch processes the pooled RoI, resizing it to a fixed size, using a CNN within each RoI. Figure 3 shows that this branch is made of four consecutive convolution layers and two deconvolution layers, with a kernel of 3×3 and 64 channels in each convolution layer. The output of this branch is masked to get the final breast cancer segmentation result.

The mask is created using the third loss function added based on Faster R-CNN [27], which brings the Refined Mask-RCNN framework's total loss function to:

$$L = L_{cls} + L_{box} + L_{mask} \quad (1)$$

where L_{cls} is the same as in Faster RCNN but L_{box} is modified so that it can obtain an accurate bounding box in the cancerous region of breast images. RER was used in place of smooth L1 regularization in the bounding box loss function. L_{cls} and L_{box} are defined as:

$$L_{cls} + L_{box} = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i, p_i^*) + \frac{1}{N_{box}}\sum_i p_i^* H_\alpha\,(t_i - t_i^*) \quad (2)$$

$$H_\alpha(P) = \begin{cases} \frac{1}{1-\alpha}\log(\sum_k P_k^\alpha) & \alpha > 0 \\ -\sum_k P_k \log\,(P_k) & \alpha \to 1 \\ -log max_k P_k & \alpha \to \infty \end{cases} \quad (3)$$

$$L_{cls}(p_i, p_i^*) = -p_i^* log p_i^* - (1 - p_i^*)\log\,(1 - p_i^*) \quad (4)$$

The average of the binary cross entropy is given by the loss of the mask, which can be represented as:

$$L_{mask} = \frac{-1}{n^2}\sum_{1\le i,j\le n}\left[x_{i,j} log x_{i,j}^k + (1 - x_{i,j})\log\,(1 - x_{i,j}^k)\right] \quad (5)$$

For each ROI, $p = (p_0, p_1, \ldots \ldots p_k)$, and p_i represents the probability of class i. t_i represents the predicted four parameterized translation and scaling parameters of class i. Moreover, p_i^* and t_i^* are binary representations of the corresponding parameters in the ground-truth label. Renyi entropy loss is used due its robust behavior for outlier points. N_{cls} and N_{box} represent the normalization of mini batch size and anchor locations numbers, respectively. To have a fair comparison, the same boundary loss function is used.

L_{mask} is a newly added loss function for the background segmentation branch. Once the mask prediction is obtained, for each pixel, a sigmoid function value is obtained and considered as input to L_{mask} (binary cross-entropy loss function). L_{mask} is only evaluated on positive RoI, which is similar to IoU greater than 0.5. L_{mask} and L_{cls} are similar to each other, with the only difference being that L_{cls} is calculated on the image level and L_{mask} on the pixel level.

III. RESULTS

A. Performance Evaluation

Seven evaluation parameters for classification and two for segmentation were used to measure the performance of the proposed method. The evaluation metrics used for classification were accuracy, True Positive Rate (TPR), True Negative Rate (TNR), precision, F-score, Balanced Classification Rate (BCR), and Youden's index [25, 26]. BCR is given by:

$$BCR = \frac{1}{2}[TPR + TNR] \quad (6)$$

Similarly, Youden's index is given by:

$$Youden's\ index = TPR - (1 - TNR) \quad (7)$$

The results of image segmentation, such as region overlap and boundary similarity, can be evaluated using several performance criteria [25]. This study used the Dice coefficient (DICE) [21] and Jaccard index (JAC) [26] to evaluate the overlap between the prediction and the actual cancerous regions. Considering P and Q be the areas of prediction and ground-truth, respectively, JAC and DICE are given by:

$$JAC = \frac{|P \cap Q|}{|P \cup Q|'} \quad (8)$$

$$DICE = \frac{2|P \cap Q|}{|P| + |Q|'} \quad (9)$$

Python was used along with the Keras library in the TensorFlow environment for developing the proposed Refined Mask-RCNN model. Feature extraction was performed by the ResNet101 backbone network. The Adam optimizer [24] was used with an initial learning rate of 0.001 and a batch size of 100. The weight decay for the convolutional and FC layers was set at 0.0001, and the momentum was 0.9. It is worth noting that the parameters of the proposed model were trained completely from scratch.

B. Implementation Platform

Figure 4 illustrates the NVIDIA Jetson Nano Development Kit, which is a compact yet powerful edge AI computing platform. It features a 1.43 GHz quad-core Arm Cortex-A57 CPU with a 128 CUDA core NVIDIA Maxwell GPU, delivering 472 GFLOPs (FP16) of computing power. With 4 GB of LPDDR4 RAM, it efficiently handles AI workloads while consuming as little as 5 W of power, making it ideal for robotics, IoT, and real-time AI applications. Its support for deep learning frameworks, such as TensorFlow, PyTorch, and ONNX, enables developers to deploy AI models efficiently on edge devices.

Fig. 4. Hardware implementation of the proposed algorithm.

C. Quantitative and Qualitative Analysis

The proposed model employs an end-to-end deep learning architecture where hyperparameter values are optimized for detection, classification, and segmentation. Figure 5 illustrates the performance of the proposed model. As shown in Figure 5(a), with the number of training epochs increasing, the model's accuracy improves. Figure 5(b) presents the graph of training loss versus epochs, where the model achieves a better training loss with epochs. In terms of performance evaluation, the proposed model outperformed existing models that used the same datasets [28, 29]. The proposed model achieved an accuracy of 99.20%, a TPR of 99%, a TNR of 99%, a precision of 0.99, an F-score of 0.99, a BCR of 0.99, a Youden's index of 0.98, a JAC of 0.99, and a DICE of 0.99, as shown in Table I.

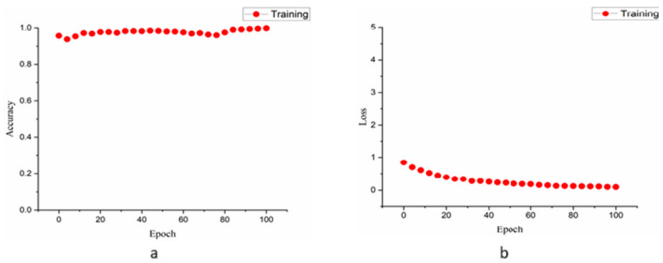

Fig. 5. (a) Training accuracy curve (b) Training loss curve.

Figures 6 and 7 illustrate some results of the proposed method in the detection, classification, and segmentation of breast cancer. Both figures list every instance of cancer detection and provide the precise localization of breast masses, which is helpful for biopsies. All of these actions were performed in parallel. The creation of a mask indicates the presence of a cancerous area in the ground truth image.

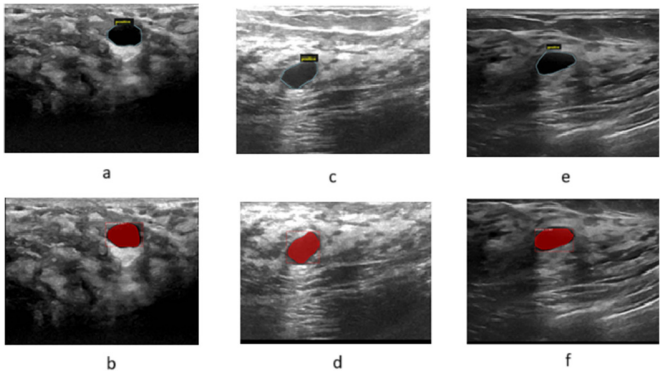

Fig. 6. Instances of cancer detection; (a), (c), (e) are ground truth images, and (b), (d), (f) are the cancer detected regions using the proposed method.

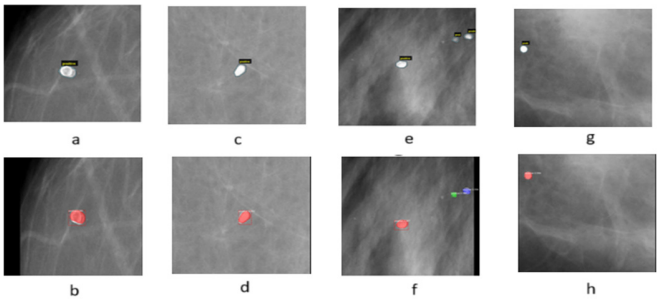

Fig. 7. Instances of cancer detection for biopsies; (a), (c), (e), (g) are ground truth images, and (b), (d), (f), (g) are the precise localization of breast mass regions detected using the proposed method.

D. Comparison with State-of-the-Art Methods

Since many studies on DL-based methods aimed at detection and classification, the results of the proposed Refined Mask-RCNN cannot be directly compared. Table I shows a comparison of the proposed method with widely used breast cancer detection, classification, and segmentation models. From this table, the following observations can be made for the detection, classification, and segmentation of breast cancer from mammogram and breast ultrasound images. It can be noticed that the proposed model for the detection, classification, and segmentation of breast cancer performs well in all measuring parameters and outperforms other standard methods in the literature.

TABLE I. CLASSIFICATION AND SEGMENTATION PERFORMANCE COMPARISON OF THE PROPOSED WITH OTHER METHODS

Methods	**Classification**							**Segmentation**	
Parameters	**Accuracy (%)**	**TPR (%)**	**TNR (%)**	**Precision**	**F-score**	**BCR**	**Youden's index**	**JAC**	**DICE**
[6]	85	80	87.36	0.75	0.78	0.84	0.68	0.75	0.78
[14]	85	91.67	89.50	0.75	0.83	0.90	0.81	0.75	0.83
[23]	93	92	94	0.94	0.93	0.93	0.86	0.94	0.93
[24]	93	96	92	0.93	0.94	0.94	0.88	0.94	0.93
Proposed	99.20	99	99	0.99	0.99	0.99	0.98	0.99	0.99

IV. CONCLUSION

This study presented a Refined Mask-RCNN framework for breast cancer detection, classification, and segmentation, which can obtain an improvement of 6% in accuracy compared to the basic network. Since it is known that using mammograms or breast ultrasound images alone can detect cancer with less than optimal efficacy, combining both screening methods can enhance breast cancer detection. Therefore, the proposed model processes both imaging modalities, and the processed values are passed to an OR operation to obtain the final result. The proposed method achieved an accuracy of 99.20%, a TPR of 99%, a TNR of 99%, a precision of 0.99, an F-score of 0.99, a balanced classification rate of 0.99, a Youden's index of 0.98, a JAC of 0.99, and a DICE of 0.99. The results show that the proposed method outperforms others across various metrics by effectively leveraging the Refined Mask-RCNN architecture.

REFERENCES

[1] *Global Cancer Facts & Figures*, 4th ed. American Cancer Society, 2018.

[2] J. B. Harford, "Breast-cancer early detection in low-income and middle-income countries: do what you can versus one size fits all," *The Lancet Oncology*, vol. 12, no. 3, pp. 306–312, Mar. 2011, https://doi.org/10.1016/S1470-2045(10)70273-4.

[3] C. Gómez-Raposo, F. Zambrana Tévar, M. Sereno Moyano, M. López Gómez, and E. Casado, "Male breast cancer," *Cancer Treatment Reviews*, vol. 36, no. 6, pp. 451–457, Oct. 2010, https://doi.org/10.1016/j.ctrv.2010.02.002.

[4] O. Akin *et al.*, "Advances in oncologic imaging," *CA: A Cancer Journal for Clinicians*, vol. 62, no. 6, pp. 364–393, 2012, https://doi.org/10.3322/caac.21156.

[5] M. G. Ertosun and D. L. Rubin, "Probabilistic visual search for masses within mammography images using deep learning," in *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Washington, DC, USA, Nov. 2015, pp. 1310–1315, https://doi.org/10.1109/BIBM.2015.7359868.

[6] S. K. Raza, S. S. Sarwar, S. M. Syed, and N. A. Khan, "Classification and Segmentation of Breast Tumor Using Mask R- CNN on Mammograms." Research Square, May 14, 2021, https://doi.org/10.21203/rs.3.rs-523546/v1.

[7] T. C. Chiang, Y. S. Huang, R. T. Chen, C. S. Huang, and R. F. Chang, "Tumor Detection in Automated Breast Ultrasound Using 3-D CNN and Prioritized Candidate Aggregation," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 240–249, Jan. 2019, https://doi.org/10.1109/TMI.2018.2860257.

[8] Y. Yuan, M. Chao, and Y. C. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks With Jaccard Distance," *IEEE Transactions on Medical Imaging*, vol. 36, no. 9, pp. 1876–1886, Sep. 2017, https://doi.org/10.1109/TMI.2017.2695227.

[9] H. Min *et al.*, "Fully Automatic Computer-aided Mass Detection and Segmentation via Pseudo-color Mammograms and Mask R-CNN," in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Iowa City, IA, USA, Apr. 2020, pp. 1111–1115, https://doi.org/10.1109/ISBI45749.2020.9098732.

[10] Y. A. Hamad, K. Simonov, and M. B. Naeem, "Breast Cancer Detection and Classification Using Artificial Neural Networks," in *2018 1st Annual International Conference on Information and Sciences (AiCIS)*, Fallujah, Iraq, Nov. 2018, pp. 51–57, https://doi.org/10.1109/AiCIS.2018.00022.

[11] Z. Wang, G. Yu, Y. Kang, Y. Zhao, and Q. Qu, "Breast tumor detection in digital mammography based on extreme learning machine," *Neurocomputing*, vol. 128, pp. 175–184, Mar. 2014, https://doi.org/10.1016/j.neucom.2013.05.053.

[12] H. Chougrad, H. Zouaki, and O. Alheyane, "Deep Convolutional Neural Networks for breast cancer screening," *Computer Methods and Programs in Biomedicine*, vol. 157, pp. 19–30, Apr. 2018, https://doi.org/10.1016/j.cmpb.2018.01.011.

[13] N. Wu *et al.*, "Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1184–1194, Apr. 2020, https://doi.org/10.1109/TMI.2019.2945514.

[14] J. Y. Chiao, K. Y. Chen, K. Y. K. Liao, P. H. Hsieh, G. Zhang, and T. C. Huang, "Detection and classification the breast tumors using mask R-CNN on sonograms," *Medicine*, vol. 98, no. 19, May 2019, Art. no. e15200, https://doi.org/10.1097/MD.0000000000015200.

[15] D. A. Ragab, M. Sharkas, S. Marshall, and J. Ren, "Breast cancer detection using deep convolutional neural networks and support vector

machines," *PeerJ*, vol. 7, Jan. 2019, Art. no. e6201, https://doi.org/10.7717/peerj.6201.

[16] Z. Jiao, X. Gao, Y. Wang, and J. Li, "A deep feature based framework for breast masses classification," *Neurocomputing*, vol. 197, pp. 221–231, Jul. 2016, https://doi.org/10.1016/j.neucom.2016.02.060.

[17] G. Carneiro, J. Nascimento, and A. P. Bradley, "Automated Analysis of Unregistered Multi-View Mammograms With Deep Learning," *IEEE Transactions on Medical Imaging*, vol. 36, no. 11, pp. 2355–2365, Aug. 2017, https://doi.org/10.1109/TMI.2017.2751523.

[18] P. Kumar, S. Srivastava, R. K. Mishra, and Y. P. Sai, "End-to-end improved convolutional neural network model for breast cancer detection using mammographic data," *The Journal of Defense Modeling and Simulation*, vol. 19, no. 3, pp. 375–384, Jul. 2022, https://doi.org/10.1177/1548512920973268.

[19] H. C. Shin *et al.*, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, Feb. 2016, https://doi.org/10.1109/TMI.2016. 2528162.

[20] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, Oct. 2017, pp. 2980–2988, https://doi.org/10.1109/ICCV.2017. 322.

[21] A. O. Vuola, S. U. Akram, and J. Kannala, "Mask-RCNN and U-Net Ensembled for Nuclei Segmentation," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, Venice, Italy, Apr. 2019, pp. 208–212, https://doi.org/10.1109/ISBI.2019.8759574.

[22] P. Kumar, S. Srivastava, Y. Padma Sai, and S. Choudhary, "Optimal Bayesian Estimation Framework for Reduction of Speckle Noise from Breast Ultrasound Images," in *Innovations in Cyber Physical Systems*, 2021, pp. 255–263, https://doi.org/10.1007/978-981-16-4149-7_22.

[23] H. Rahman, T. F. Naik Bukht, R. Ahmad, A. Almadhor, and A. R. Javed, "Efficient Breast Cancer Diagnosis from Complex Mammographic Images Using Deep Convolutional Neural Network," *Computational Intelligence and Neuroscience*, vol. 2023, no. 1, 2023, Art. no. 7717712, https://doi.org/10.1155/2023/7717712.

[24] M. M. Rahman *et al.*, "Breast Cancer Detection and Localizing the Mass Area Using Deep Learning," *Big Data and Cognitive Computing*, vol. 8, no. 7, Jul. 2024, Art. no. 80, https://doi.org/10.3390/bdcc8070080.

[25] S. M. Shaaban, M. Nawaz, Y. Said, and M. Barr, "An Efficient Breast Cancer Segmentation System based on Deep Learning Techniques," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12415–12422, Dec. 2023, https://doi.org/10.48084/etasr.6518.

[26] T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 936–944, https://doi.org/10.1109/ CVPR.2017.106.

[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *Advances in Neural Information Processing Systems*, 2015, vol. 28.

[28] "DDSM Mammography." [Online]. Available: https://www.kaggle.com/ datasets/skooch/ddsm-mammography.

[29] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images," *Data in Brief*, vol. 28, Feb. 2020, Art. no. 104863, https://doi.org/10.1016/j.dib.2019.104863.