



# Breast Cancer Semantic Segmentation for Accurate Breast Cancer Detection with an Ensemble Deep Neural Network

T. Nagalakshmi<sup>1</sup>

Accepted: 19 April 2022 / Published online: 20 May 2022  
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

Breast tumors are the major malignancy in females and diagnostic systems using artificial intelligence algorithms for breast imaging have shown promising results. Among many algorithms, a deep convolutional neural network (DCNN) using K-means clustering and a multiclass support vector machine model enhance the precision of categorizing breast tumors from mammogram images. Nonetheless, effective breast tumor identification is still difficult without partitioning the pectoral muscle (PM) boundary from the remaining breast area. Therefore, this article proposes an Ensemble-Net model by ensembling the transfer learning model with different pre-trained CNN structures for partitioning the PM boundary from the remaining breast area in the mammographic scans. The segmentation process has 2 phases. In the initial phase, different region-of-interests are generated that include the object according to the input images. In the secondary phase, the object class is predicted after the areas of bounding boxes are refined and a pixel-range mask is created for the entity. These 2 different phases are associated with the backbone structure which creates the pyramid hierarchy of DCNN to acquire the features from the raw images. Moreover, it employs global average pooling followed by the softmax classification to recognize the normal, benign and malignant cases. Finally, the experimental outcomes demonstrate that the Ensemble-Net achieves 96.72% accuracy than the other classical classifiers.

**Keywords** Breast tumor · DCNN · K-means clustering · MSVM · Mammogram scans · Pectoral muscle · Transfer learning · Pixel-level mask

## 1 Introduction

A Breast tumor is one of the tumor categories that form and progresses in the breast tissues. Around 99% of all incidences of breast tumors occur in women. The World Health Organization (WHO) statistics show that breast tumors are the most common malignancy among females and that incidences are growing in virtually every major part of the country [1]. It

---

✉ T. Nagalakshmi  
nagalakshmi.vital@gmail.com

<sup>1</sup> Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India

is detected at an advanced stage in developing and undeveloped regions, even though few threats can be reduced by preventive measures. To improve clinical outcomes, it is important to diagnose the breast tumor as quickly as possible.

Early identification and prevention of breast tumors depend on effective and rapid prognosis during radiography assessments [2]. There are different screening methods available to recognize and diagnose breast tumors. Amongst others, mammography serves an essential function in dependable early recognition and treatment of breast tumors. It is commonly used to test for breast tumors via mammography, which utilizes low-dose X-rays for inspecting the breast. Mammographic examinations are useful for recognizing abnormalities and groups of lesions, both of which are significant features of breast tumors [3]. Generally, mammograms are suggested for different conditions: (1) treatment of breast abnormalities, (2) mammography that was taken after a prior negative mammogram and (3) monitoring of tumors or deformities.

When using categorization algorithms to diagnose breast tumors, the pre-processing phase is intended to highlight abnormalities in mammography scans that have poor resolution [4]. It increases the resolution, highlighting abnormal areas of the image. In the enhancing phase, a specific description of the unseen visual is displayed. It is a diagnostic aid that improves the quality of the image.

Then, a segmentation phase is applied to extract the region-of-interests (ROI) or the cancerous area from the actual image. Region-based, edge-based, clustering and thresholding are the categories of segmentation algorithms [5]. The desired characteristics, like shape, texture, etc., are extracted and chosen from the mammogram scans. Finally, the categorization phase is executed to distinguish benign and malignant lesions from the mammogram scan.

There are several categorization algorithms developed, including machine and deep learning [6]. Based on the datasets, characteristics, methods, findings and variables, the best algorithm is decided for categorization. The machine learning algorithms are highly complex for radiologists to execute the diagnostic task in less time while increasing the number of images. To solve this problem, many deep learning algorithms have been suggested which identify the incidence of breast tumors rapidly and precisely.

The deep convolutional neural network (DCNN) model containing pre-processing and feature mining by the K-means clustering is developed and explained in [7]. A novel layer has been included at the categorization level which conducts a fraction of 70% learning to 30% testing of the DCNN and multiclass support vector machine (MSVM). On the other hand, precise breast tumor identification from mammographic scans is still challenging since the pectoral muscle (PM) boundary is not effectively partitioned from the remaining breast area. To combat this issue, Mask Regional CNN (RCNN) [8] has been suggested which uses a Residual Network (ResNet101) as a backbone structure to extract object bounding boxes, masks, and keypoints. But the accuracy is not satisfactory since different scales of feature maps semantic details are neglected.

Hence, in this paper, the Ensemble-Net model is developed to segment the PM boundary from the remaining breast area. This new Ensemble-Net is built by ensembling the transfer learning (TL) model with different pre-trained structures such as VGG16, ResNet50 and Xception65. It involves 2 different phases: (1). the first phase creates various ROIs which include the entity depending on the entry scans and (2). The secondary phase estimates the object label once it processes the area bounding boxes and generates a pixel-range mask for the entity.

These phases are linked to the backbone configuration called the feature pyramid net that forms the pyramid hierarchy of DCNNs for extracting the features from the raw images. Also, it uses the global average pooling to improve the training of the multiple-scaled version of

the input images and softmax classification for categorizing the malignant cases. Thus, it enhances the identification rate of breast tumors through partitioning the PM boundary from the mammogram scans.

The rest of the paper is prepared as the following: Sect. 2 discusses the work associated with the breast tumor categorization. Section 3 explains the design of Ensemble-Net model and Sect. 4 demonstrates its efficiency. Section 5 summarizes this paper and suggests future improvements.

## 2 Literature Survey

Al-Masni et al. [9] presented the ROI-based CNN called YOLO to concurrently identify and categorize breast masses in digital mammograms. This model has different phases: preprocessing, feature extraction, mass identification with belief and tumor categorization using Fully Connected (FC) neural networks. But, this model was not suitable for a limited number of datasets and the tumor area in mammograms was not partitioned. To categorize mammography mass lesions, Chougrad et al. [10] designed the DCNN to categorize and to learn the CNN model. Once the ROIs extracted from the entire mammograms were pre-processed and normalized, they were combined to create a unified huge dataset and applied for updating CNN. But, it does not eliminate the noise or muscle which leads to miscategorization.

Wu et al. [11] suggested a new 2-step DCNN structure for forecasting breast tumors. In this model, a high-capacity path-level network was utilized for training from pixel-level classes along with network training from macroscopic breast-level classes. A classic ResNet-based network was employed as the basic unit wherein the trade-off between depth and width was fine-tuned. Also, multiple-input views were merged optimally among the number of promising probabilities. Still, the accuracy of prediction was not greatly increased.

Mohanty et al. [12] presented a block-based discrete wavelet conversion that was applied for extracting the different features and the principal component analysis was utilized for selecting the discriminating features. Also, a fine-tuned kernel extreme learning machine utilizing a weighted chaotic swarm strategy was employed to categorize digital mammograms. But, it has a high computation burden and does not deal with the automated partitioning of ROIs.

To address click feature prediction from visual data, Yu et al. [13] discussed a Hierarchical Deep Word Embedding (HDWE) model that incorporates sparse restrictions and an enhanced RELU operator. HDWE is a coarse-to-fine click feature predictor that is trained using an auxiliary picture dataset comprising click data. Feature selection aids in the automatic selection of better, more reliable feature elements, resulting in improved identification accuracy.

Agarwal et al. [14] designed a completely automated model called Faster-RCNN for identifying masses in full-field digital mammograms. Also, TL was applied to the learned Faster-RCNN for detecting the masses in the limited datasets of digital mammograms. But, it lacks partition and categorization of the masses into different categories since it deals with only mass identification. To extract the attributes from the mammogram scans, Dabass et al. [15] designed a technique by the Mamta-Hanman entropy factor. Such attributes were fed to the Hanman conversion and the hesitancy based Hanman conversion classifications for categorization. But, it needs to extend the information set via adding the concept of non-membership factor to increase efficiency.

Sapate et al. [16] developed a fusion-based model wherein the width of the annular area was chosen dynamically depending on the dimension of the nodules and the breast for

matching the respective couples of malignant nodules from craniocaudal and medio-lateral oblique views. Then, single-view features and 2-view correspondence scores were merged to differentiate malignant and benign lesions using the SVM classifier. But its accuracy was not effective and the learning time was high.

Xie et al. [17] presented a computerized multi-scale end-to-end deep neural network system for categorizing mammogram images. This system has breast area partition, feature extraction, multi-scale feature and categorization units. First, the raw mammogram images were partitioned according to their features. Then, the multi-scale unit was applied to create the feature maps at 3 different levels that offers the classifier with the data of global breast and confined tumors for categorization. But, it lacks ROI annotation, which leads to less classifier efficiency.

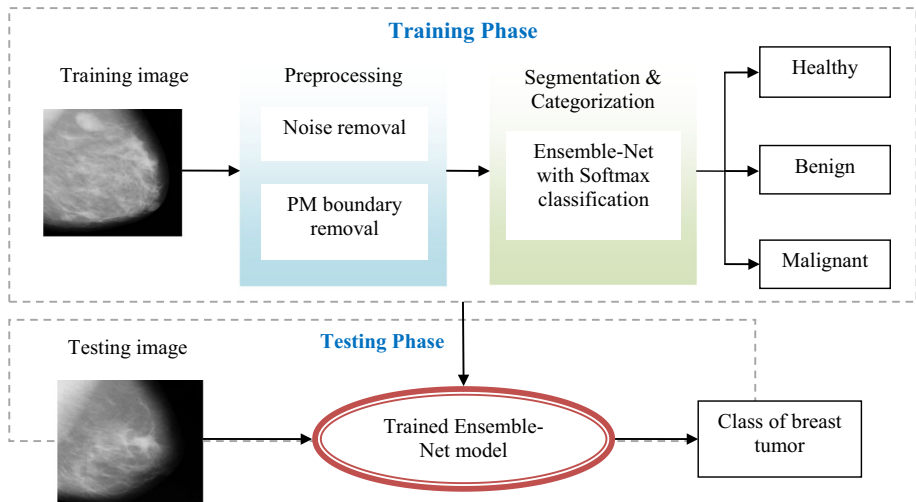
Shu et al. [18] developed 2 pooling architectures and an end-to-end structure depending on DCNN. In this system, 3 phases were performed: feature extraction, special pooling and categorization. First, the mammographic images were split into areas based on the extracted feature maps and the areas with the higher chance of containing malignant were chosen to acquire the resultant feature. This process was achieved by area-based and global group-max aggregation. Then, the concatenated features were used for categorization, which determines the diagnostic outcome. But, its efficiency depends on the amount of learning images.

Salama and Aly [19] designed a Modified U-Net (MUNet) model for partitioning breast region from the mammogram scans. Also, they used TL as data augmentation method. After partition, InceptionV3, DenseNet121, Resnet50, VGG16 and MobileNetV2 structures were recommended for categorizing mammogram scans. But, it does not consider the partition the PM boundaries which impacts the categorization of tumor cases. But Yu et al. [20] proposed single-pixel reconstruction net (SPRNet), which introduces a single-pixel reconstruction (SPR) detector to perform efficient instance segmentation. The newly added SPR branch reconstructs the pixel-level mask directly from each pixel in the convolution feature map. SPRNet achieves faster inference speed using the same ResNet-50 backbone.

### 3 Proposed Methodology

In this section, the Ensemble-Net model is explained briefly. The schematic overview of the Ensemble-Net-based breast tumor segmentation and categorization model is illustrated in Fig. 1. The processes in this model are the following:

1. Initially, the mammogram scans are gathered and preprocessing using an adaptive median filter is performed to eliminate the noise from raw scans.
2. Then, the PM area is eliminated by using different mechanisms such as binarization, histogram normalization and morphological functions.
3. After that, the pre-processed scan is split into equal dimensions. The other small patches are padded at the edge to generate an equal-dimension database.
4. Moreover, this database is utilized for learning Ensemble-Net for breast tumor partition and categorization.
5. Then, the learned Ensemble-Net model is used to classify the test scans into different statuses of breast tumors, e.g., healthy, benign and malignant.



**Fig. 1** Schematic representation of Ensemble-Net model for breast tumor segmentation and categorization

### 3.1 Image Acquisition

Initially, different categories of breast mammogram scans include 208 healthy scans, 62 benign scans and 52 malignant tumor scans have been gathered from the Mini-Mammographic Image Analysis Society (Mini-MIAS) database. This is accessible at <https://www.kaggle.com/kmader/mias-mammography> and contains 161 patient's details having an overall of 322 digitized scans (left and right breast). Each of the 322 scans is of dimension  $1024 \times 1024$  pixels and has a resolution of 200 microns. Also, it provides various details including MIAS database source index, nature of environment noise, category of irregularity, severity of illness,  $x, y$  scan-coordinates of axis of irregularity and estimated radius (in pixels) of a loop covering the irregularity.

### 3.2 Pre-processing

After collecting the breast mammogram scans, image pre-processing is applied for decreasing the image redundancy. The image denoising technique based on adaptive median filter is used for obtaining noise-free unique quality scans by enhancing their visual aspects.

It also eliminates unwanted data by insight and creates a scan highly suitable for the consecutive processes. After eliminating the unwanted distortions (e.g., speckle noise), it is observed that a few scans contain white strips in the top and bottom areas of the scans. Such areas are chosen and the pixel range is altered with zero about 1% of the top and bottom areas.

Typically, the database comprises both left and right-part scans of the breast. To eliminate the muscle area, several mechanisms are applied to right-side scans, whereas the left-part scan is identified with the pixel rate and sliding horizontally to create a similarity to the right-part. Once the muscle area is removed, the scan is sliding reverse to the left-part. The final scan is binarized subsequent to the histogram normalization wherein merely the bright areas are defined. The bright area is further improved using the morphological functions.

Afterwards, it is observed that the muscle area appeared in the top area. So, the processed scan is analysed from the left to the right-part horizontally until the white pixel. All white pixel locations are accumulated as  $(x, y)$  coordinates. It maintains analysis until  $1/3$  height of the scan or the white part endings. As well, perpendicularly scanning is not analysed repeatedly. It skips 10 pixels perpendicular behind all parallel loops. The points provide the course of the muscle area. The contour formula is a set of the chosen points for expanding the line on all sides. The point is extended to the right base and top left. It provides a top and right base pixel position, whereas the top right location is already identified. The 3 points are applied for producing a triangular mask that is considered for cropping muscle areas.

From the enhanced scans without noise and muscles, the breast area is extracted by the maximum linked elements. Additionally, the breast boundaries are smoothened since they have irregular boundaries. It supports generating the binary masks which are used for extracting the breast area from the actual scans.

### 3.3 Transfer Learning

In this model, the Ensemble-Net is constructed based on 3 different pre-trained backbone structures, such as VGG16, ResNet50 and Xception65. The primary and final layers of respected structures are modified to apply the strategy of TL. The FC convolutional layers are layered consecutively. The scans include various details and are being extracted by the convolutional layers. By fine-tuning numerous pre-trained models, the transfer learning method is being utilised to distinguish between malignant and benign breast cancer.

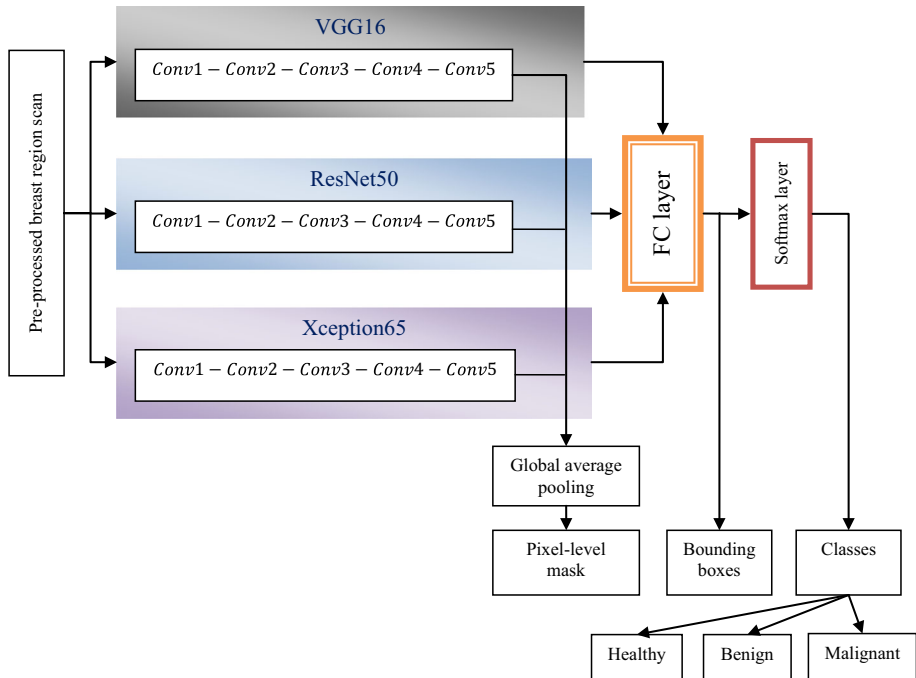
The main intention of this TL is to modify Ensemble-Net using test images without altering the classical structures. The model may need to be adapted or refined on the input–output pair data available for the task of interest. The task of modification involves the fine-tuning of the primary and final layers. The final unit is substituted using the special label FC layers. Also, the back-propagation scheme is met and fine-tuning the novel categorization issue is also addressed.

### 3.4 Ensemble Network-Based Segmentation

The Ensemble-Net allocates bounding boxes, labels and masks to the considered scan employing the deep learner. It has 2 phases: (1) During the initial phase, various areas which involve the object depending on the input scans are created and (2) The object label is predicted in the secondary phase once the area bounding boxes are refined, and a pixel-range mask is created for the entity. These phases are linked to the backbone configuration which is a feature pyramid network which generates the pyramid hierarchy of DCNN for extracting the features from the raw scans. This network structure has 3 kinds of links: bottom-up link, top–bottom link and lateral links. The structure of Ensemble-Net is shown in Fig. 2.

The attribute miner unit is called a bottom-up path which may be a few CNN. In this study, 3 pre-trained structures are employed: VGG16, ResNet50 and Xception65. The VGG16, ResNet50 and Xception65 structures involve convolutional layers, pooling layers, and FC layers. These structures use the strategy of skip links and normalizing the mined attributes batch-wise. The skip links strategy is established using the gate elements and so it has less complexity.

The attribute pyramid map is created by the top–bottom path. It utilizes tangential links to merge low-resolution, robust and fragile attributes semantically. This net consists of semantically robust features at different resolution scales. Also, global average pooling is applied



**Fig. 2** Structure of Ensemble-Net

to aggregate the extracted attributes at various scales from different structures. The bilinear sampling adjusts the dimensions of the scale features. Thus, it provides the feature maps that are used by the regional proposal network for scanning and localizing entities. The searching strategy employs the idea of anchors for minimizing the search difficulty and mapping the mined features to an unprocessed scan.

The anchors are a collection of prescribed boxes with predefined positions in the scan. The anchor's scales are assigned comparative to the unprocessed scans. The boundary boxes and ground truth labels are set as the anchors by the Intersection-Over-Union (IOU) ranges. The anchors are connected with the various stages of the attribute map, a regional proposal net is learned for extracting the object by the feature map and the dimension of the bounding box is assigned, respectively. The training strategy of a convolutional layer, downsampling and upsampling maintains the trained features in a similar relative position as the object in actual scans.

During the secondary phase, it considers the entry from the area projected by the regional proposal net, allocates the attribute map stage to the unprocessed scan positions and observes the final region. Then, it performs 3 processes, i.e., allocates the objects multiple labels, creates bounding boxes and marks pixel-level mask with those bounding boxes. It utilizes a 2-phase strategy called ROI rather than an anchor trick. The ROI extracts various areas from the feature map. After that, such areas are applied to create a pixel-range mask for every entity that exists in the projected areas. Consecutively, a softmax classification layer is added to categorize the partitioned ROIs into normal, benign, and malignant. Thus, this Ensemble-Net can segment the breast tumor regions in input scans effectively by eliminating muscle areas and categorize the tumor status.

**Algorithm for Ensemble-Net with Noise and Muscle Removal Processes:****Input:** Raw breast mammogram scans**Output:** Breast tumor status (healthy or benign or malignant)**Begin**Acquire the input scan ( $I$ ) and apply the adaptive median filtering to remove the noise;

Create the binary mask and crop the extracted breast scan;

**if**(scanside( $I$ ) == right) // do normalization and morphological functions. $J \leftarrow \text{Histogram\_normalize}(I);$  $J \leftarrow \text{Binarize}(J);$  $J \leftarrow \text{Morpological\_function}(J);$ **while**( $\frac{1}{3}$  heightofscan)Accumulate all white pixels as  $(x, y)$  coordinates.**endwhile** $P \leftarrow \text{Polyfit}(x, y);$  $P \leftarrow \text{Expand\_line}(P, \text{width}(I));$  $\text{Binarymask} \leftarrow \text{Poly2Mask}(P, \text{size}(I));$  $\text{Binarymask} \leftarrow \text{Invert}(\text{Binarymask});$ 

Crop the breast area. // removal of PM boundary.

**if**(scanside( $I$ ) == left) $\text{Breastarea} \leftarrow \text{Fliphorizontal}(\text{Breastarea});$ **endif****endif**

Obtain the pre-processed breast area scan without noise and muscle.

Train the Ensemble-Net model.

Partition the tumor region in the considered scan.

Perform softmaxclassification.

Classify the test scans (healthy, benign, and malignant scans).

**End**



## 4 Experimental Results

In this section, the efficiency of Ensemble-Net is analyzed by implementing it in MATLAB 2017b and compared with the classical models regarding precision, recall, f-measure, accuracy and IOU. The considered classical models are DCNN [7], Mask RCNN [8] and InceptionV3-MUNet [19]. From the considered database, 70% of scans have been reported for learning and the residual 30% have been reported for testing. Figure 3 portrays the outcomes of preprocessing for removing the PM areas from the breast scans (Fig. 4).

- **Accuracy:** It is the fraction of the properly categorized scan from the overall number of scans.

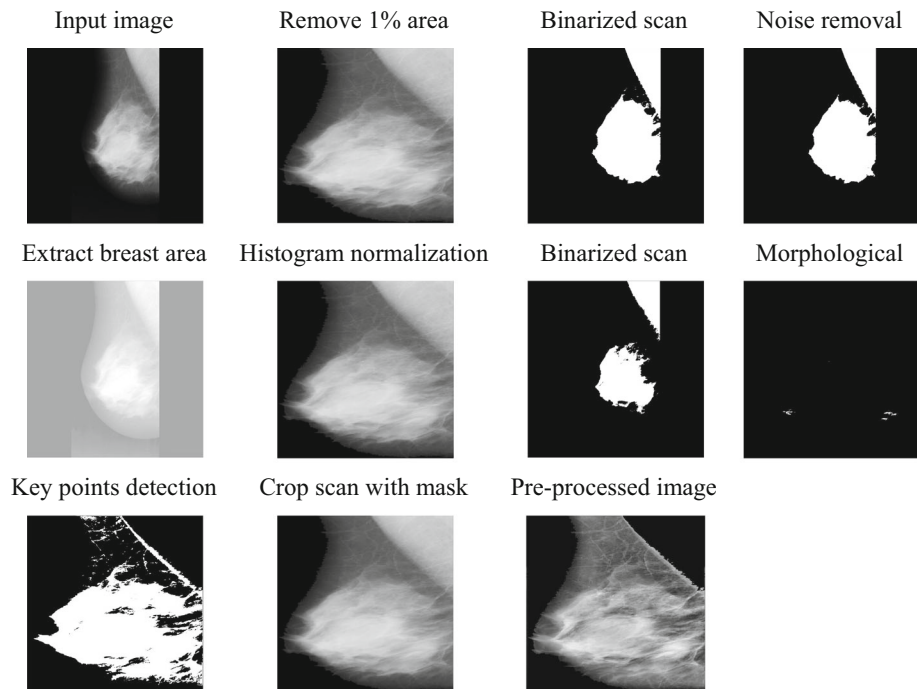
$$Accuracy = \frac{\text{Amount of perfectly categorized scan}}{\text{Total scans}} \times 100$$

- **Precision:** It is the fraction of True Positives (TP) of the scans that are categorized properly, i.e., True Positive (TP) and False Positive (FP).

$$Precision = \frac{TP}{TP + FP} \times 100$$

- **Recall:** It is the fraction of overall relevant scans properly categorized by the model.

$$Recall = \frac{TP}{TP + \text{False Negative (FN)}} \times 100$$



**Fig. 3** Step-by-step outcomes of preprocessing for PM removal

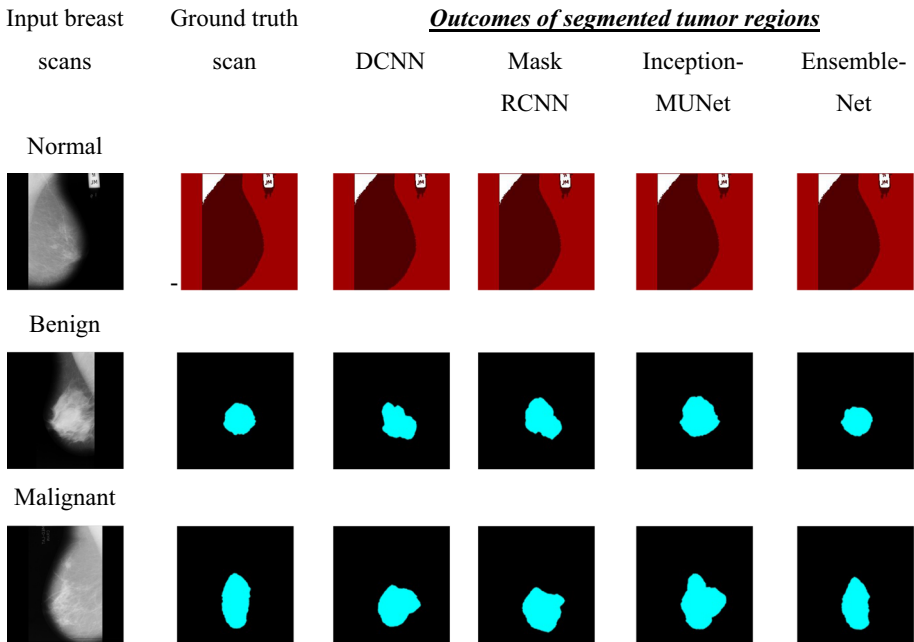


Fig. 4 Illustrates the segmented tumor region outcomes obtained from different models

- *F-measure*: The harmonic average of precision and recall is called f-measure.

$$F\text{-measure} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100$$

- *IOU*: It is the area of overlapped or predicted labels from the region of the union. It calculates the overlap involving 2 edges. It is utilized for determining how much the expected edge covers with the ground truth.

$$IOU = \frac{\text{Region of overlap}}{\text{Region of union}} \times 100$$

- *Root mean square error (RMSE)*: It is used to measure the segmentation accurateness. It is determined by taking the square root of MSE value as:

$$RMSE = \sqrt{\frac{1}{N} \sum_i \sum_j (S_{ij} - I_{ij})^2} \times 100$$

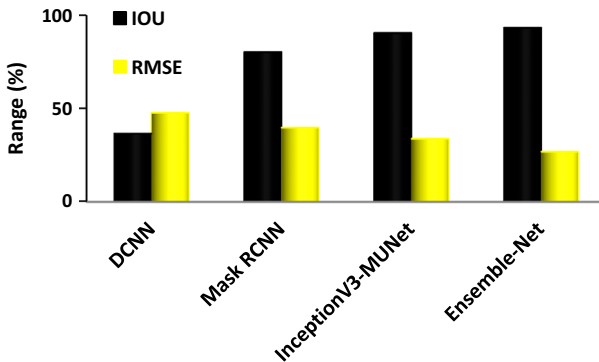
Here,  $N$  is the total amount of scans,  $S$  is the segmented scan,  $A$  is an actual scan and  $i, j$  are pixels in the scans.

Table 1 presents the outcomes of IOU and RMSE for DCNN, Mask RCNN, InceptionV3-MUNet and Ensemble-Net. From this scrutiny, it observes that the presented Ensemble-Net has higher IOU and lower RMSE compared to all other models for segmenting tumor regions from breast images.

Figure 5 displays the statistical analysis of segmentation efficiency for DCNN, Mask RCNN, InceptionV3-MUNet and Ensemble-Net models implemented on Mini-MIAS

**Table 1** Comparison of segmentation efficiency for proposed and existing models

Metrics/Models	DCNN	Mask RCNN	InceptionV3-MUNet	Ensemble-Net
IOU (%)	36.74	80.49	90.78	93.53
RMSE (%)	48	40	34	27

**Fig. 5** Statistical analysis of segmentation performance

dataset. It indicates that the Ensemble-Net model accomplishes greater efficiency in partitioning and categorizing breast tumor scans successfully compared to all other existing models.

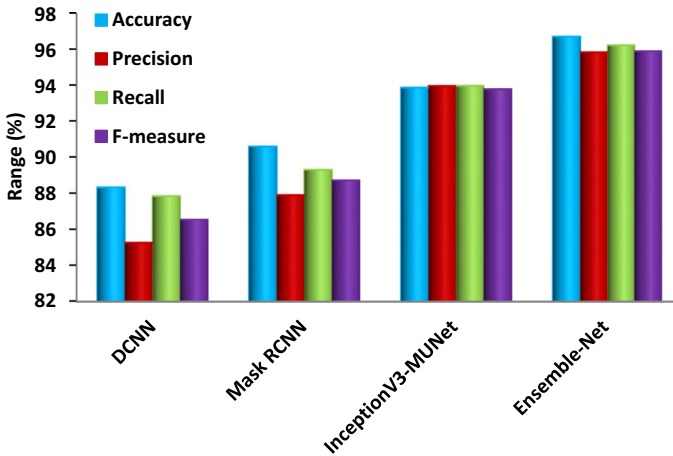
The IOU of Ensemble-Net is 154.57% larger than the DCNN, 16.2% larger than the Mask RCNN and 3.03% larger than the InceptionV3-MUNet. Likewise, the RMSE of Ensemble-Net is 43.75% less than the DCNN, 32.5% less than the Mask RCNN and 20.59% less than the InceptionV3-MUNet.

Table 2 presents the outcomes achieved by classification of breast tumor types using different models. This analysis indicates that the presented Ensemble-Net achieves better efficiency for classifying the breast tumor types compared to all other models.

Figure 6 displays the statistical analysis of categorization efficiency for DCNN, Mask RCNN, InceptionV3-MUNet and Ensemble-Net models implemented on Mini-MIAS dataset. It indicates that the Ensemble-Net model accomplishes greater efficiency in partitioning and categorizing breast tumor scans successfully compared to all other existing models. The accuracy of Ensemble-Net is 9.47% greater than the DCNN, 6.74% greater

**Table 2** Comparison of classification efficiency for proposed and existing models

Metrics/Models	DCNN	Mask RCNN	InceptionV3-MUNet	Ensemble-Net
Accuracy (%)	88.35	90.61	93.88	96.72
Precision (%)	85.27	87.92	93.97	95.85
Recall (%)	87.86	89.33	93.98	96.24
F-measure (%)	86.55	88.74	93.79	95.91



**Fig. 6** Statistical analysis of categorization performance

than the Mask RCNN and 3.03% greater than the InceptionV3-MUNet. The precision of Ensemble-Net is 12.41% higher than the DCNN, 9.02% higher than the Mask RCNN and 2% higher than the InceptionV3-MUNet. Also, the recall of Ensemble-Net is 9.54% greater than the DCNN, 7.74% greater than the Mask RCNN and 2.4% greater than the InceptionV3-MUNet. Similarly, the f-measure of Ensemble-Net is 10.81% increased than the DCNN, 8.08% increased than the Mask RCNN and 2.26% increased than the InceptionV3-MUNet. Thus, it summarizes the Ensemble-Net with softmax classification outperforms than the other models for both partitioning and categorization of breast tumor classes.

## 5 Conclusion

In this paper, the Ensemble-Net structure was presented for partitioning the PM boundary from the remaining breast area in the mammographic scans for categorizing the breast tumor. First, this model focused on mammogram preprocessing and normalization into the deep learner framework. The preprocessing was achieved to eliminate unwanted noise, artifacts and muscle regions. Then, the TL-based Ensemble-Net model was applied to the pre-processed images to segment the breast tumor region accurately. Further, the softmax classifier was employed for categorizing the tumor status into healthy, benevolent, and malevolent. So, this model can be trained for mining the sample relying on the mammogram, categorizing the breast tumor and partitioning the tumor area. To conclude, the testing findings revealed that the proposed model has 96.72% accuracy compared to the other classical segmentation and categorization models. On the other hand, it needs to consider the different subjective characteristics related to physician experience to increase the accuracy of breast tumor classification. So, future work will concentrate on extracting more features for breast tumor classification.

**Funding** Not applicable.

**Availability of data and materials** Data sharing is not applicable to this article as no new data were created or analyzed in this study.

## Declarations

**Conflict of interest** The authors declare that they have no conflict of interest.

**Informed consent** Informed consent does not apply as this was a retrospective review with no identifying patient information.

**Human and animal rights** This article does not contain any studies with human or animal subjects performed by any of the authors.

## References

1. Akram M, Iqbal M, Daniyal M, Khan AU (2017) Awareness and current knowledge of breast cancer. *Biol Res* 50(1):1–23
2. Ginsburg O, Yip CH, Brooks A, Cabanes A, Caleffi M, Dunstan Yataco JA, Gyawali B, McCormack V, McLaughlin de Anderson M, Mehrotra R, Mohar A (2020) Breast cancer early detection: a phased approach to implementation. *Cancer* 126:2379–2393
3. Mendes J, Matela N (2021) Breast cancer risk assessment: a review on mammography-based approaches. *J Imaging* 7(6):1–20
4. Meenalochini G, Ramkumar S (2021) Survey of machine learning algorithms for breast cancer detection using mammogram images. *Mater Today Proc* 37:2738–2743
5. Michael E, Ma H, Li H, Kulwa F, Li J (2021) Breast cancer segmentation methods: current status and future potentials. *BioMed Res Int* 2021:1–29
6. Pang T, Wong JHD, Ng WL, Chan CS (2020) Deep learning radiomics in breast cancer with different modalities: overview and future. *Expert Syst Appl* 158:1–15
7. Kaur P, Singh G, Kaur P (2019) Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification. *Inform Med Unlocked* 16:1–15
8. He K, Gkioxari G, Dollár P, Girshick R (2017) Mask RCNN. In: *Proceedings of the IEEE international conference on computer vision*, pp 2961–2969
9. Al-Masni MA, Al-Antari MA, Park JM, Gi G, Kim TY, Rivera P, Valarezo E, Choi MT, Han SM, Kim TS (2018) Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. *Comput Methods Programs Biomed* 157:85–94
10. Chougrad H, Zouaki H, Alheyane O (2018) Deep convolutional neural networks for breast cancer screening. *Comput Methods Programs Biomed* 157:19–30
11. Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M, Jastrzębski S, Févry T, Katsnelson J, Kim E, Wolfson S (2019) Deep neural networks improve radiologists' performance in breast cancer screening. *IEEE Trans Med Imaging* 39(4):1184–1194
12. Mohanty F, Rup S, Dash B, Majhi B, Swamy MNS (2020) An improved scheme for digital mammogram classification using weighted chaotic salp swarm algorithm-based kernel extreme learning machine. *Appl Soft Comput* 91:1–16
13. Yu J, Tan M, Zhang H, Rui Y, Tao D (2022) Hierarchical deep click feature prediction for fine-grained image recognition. *IEEE Trans Pattern Anal Mach Intell* 44(2):563–578. <https://doi.org/10.1109/TPAMI.2019.2932058>
14. Agarwal R, Díaz O, Yap MH, Llado X, Martí R (2020) Deep learning for mass detection in full field digital mammograms. *Comput Biol Med* 121:1–10
15. Dabass J, Hanmandlu M, Vig R (2020) Classification of digital mammograms using information set features and Hanman Transform based classifiers. *Inform Med Unlocked* 20:1–8
16. Sapate S, Talbar S, Mahajan A, Sable N, Desai S, Thakur M (2020) Breast cancer diagnosis using abnormalities on ipsilateral views of digital mammograms. *Biocybern Biomed Eng* 40(1):290–305
17. Xie L, Zhang L, Hu T, Huang H, Yi Z (2020) Neural networks model based on an automated multi-scale method for mammogram classification. *Knowl Based Syst* 208:1–9
18. Shu X, Zhang L, Wang Z, Lv Q, Yi Z (2020) Deep neural networks with region-based pooling structures for mammographic image classification. *IEEE Trans Med Imaging* 39(6):2246–2255
19. Salama WM, Aly MH (2021) Deep learning in mammography images segmentation and classification: automated CNN approach. *Alex Eng J* 60(5):4701–4709

20. Yu J, Yao J, Zhang J, Yu Z, Tao D (2020) SPRNet: single-pixel reconstruction for one-stage instance segmentation. *IEEE Trans Cybern* 51:1731–1742

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.