



Automated abnormalities detection in mammography using deep learning

Ghada M. El-Banby¹ · Nourhan S. Salem¹ · Eman A. Tafweek² · Essam N. Abd El-Azziz¹

Received: 20 September 2023 / Accepted: 16 June 2024 / Published online: 11 July 2024
© The Author(s) 2024

Abstract

Breast cancer is the second most prevalent cause of cancer death and the most common malignancy among women, posing a life-threatening risk. Treatment for breast cancer can be highly effective, with a survival chance of 90% or higher, especially when the disease is detected early. This paper introduces a groundbreaking deep U-Net framework for mammography breast cancer images to perform automatic detection of abnormalities. The objective is to provide segmented images that show areas of tumors more accurately than other deep learning techniques. The proposed framework consists of three steps. The first step is image preprocessing using the Li algorithm to minimize the cross-entropy between the foreground and the background, contrast enhancement using contrast-limited adaptive histogram equalization (CLAHE), normalization, and median filtering. The second step involves data augmentation to mitigate overfitting and underfitting, and the final step is implementing a convolutional encoder-decoder network-based U-Net architecture, characterized by high precision in medical image analysis. The framework has been tested on two comprehensive public datasets, namely INbreast and CBIS-DDSM. Several metrics have been adopted for quantitative performance assessment, including the Dice score, sensitivity, Hausdorff distance, Jaccard coefficient, precision, and F1 score. Quantitative results on the INbreast dataset show an average Dice score of 85.61% and a sensitivity of 81.26%. On the CBIS-DDSM dataset, the average Dice score is 87.98%, and the sensitivity reaches 90.58%. The experimental results ensure earlier and more accurate abnormality detection. Furthermore, the success of the proposed deep learning framework in mammography shows promise for broader applications in medical imaging, potentially revolutionizing various radiological practices.

Keywords Mammography · Breast cancer detection · Deep learning · UNet

Introduction

Breast cancer is a leading cause of mortality for women globally. As per the World Health Organization (WHO) data, 2.3 million women were diagnosed with breast cancer, and 685,000 succumbed to it by 2020 [1]. When breast cancer is identified at an initial stage, before it has grown significantly or spread, the chances of effective treatment are higher. The American Cancer Society states that for breast cancer caught in its early and localized phase, the 5-year relative survival

rate is an impressive 99% [2]. The most reliable strategy to detect breast cancer early is to have frequent screening examinations. Screening consists of tests and exams aimed at identifying diseases in individuals without any manifesting symptoms. Having routine mammograms is pivotal for the early identification of breast cancer, maximizing the chances of successful treatment [2].

Mammograms are performed with a machine that specifically looks at breast tissue. The machine uses lower-dose x-rays than those used to examine other regions of the body, such as the lungs or bones. Two plates on the mammography machine compress or flatten the breast to spread the tissue apart. This produces a higher-quality image while using less radiation [2]. Mammogram of each breast are taken from two different views which are either mediolateral-oblique (MLO) and cranio-caudal (CC). The CC view is a top-down view of the breast. The MLO, on the other hand, is taken from an angled perspective. For radiologists, the MLO view is pre-

✉ Nourhan S. Salem
Norhan.selem1450@el-eng.menofia.edu.eg

¹ Department of Industrial Electronics and Control Engineering, Faculty of Electronic Engineering, Menoufia University, 32952 Menouf, Egypt

² Department of Clinical Oncology, Faculty of Medicine, Menoufia University, Menoufia, Egypt

ferred because it shows more of the breast features in the upper outer quadrant [3]. While mammograms are universally employed regardless of the severity of breast cancer, breast magnetic resonance imaging (MRI) is typically utilized for detecting higher levels of breast cancer. This can sometimes yield misleading results, indicating cancer even when none is present [3].

Diagnosing tumors is a time-consuming process and often poses challenges for radiologists interpreting medical images due to the interference of noise, artifacts, and intricate structures [4]. This challenge is intensified by a worldwide scarcity of radiologists and medical professionals skilled in analyzing screening data, especially in under-served areas and developing nations [5]. Swift patient care encompassing screening, diagnosis, and treatment is crucial, as time plays a pivotal role in saving lives when it comes to breast cancer. Beyond the time aspect, there are instances where mammography might misread certain tumors, leading to false negatives. This can delay the detection and subsequent treatment of cancer. To address this, Computer Aided Diagnosis (CAD) was developed to serve as a supplementary opinion to the radiologist's judgment.

The use of imaging technology is crucial for the early diagnosis of breast cancer, a process that encompasses phases from screening and early detection to subsequent diagnosis and treatment. Various modalities such as mammography, contrast-enhanced mammography, microwave imaging, optical imaging, ultrasonography, magnetic resonance imaging (MRI), and nuclear medicine are employed in the detection of abnormalities [6–9]. The integration of intelligent systems capable of promptly and accurately identifying and diagnosing abnormalities is increasingly vital. Deep learning, a subset of machine learning algorithms, operates directly on images, autonomously selecting optimal features without human intervention. These algorithms are rapidly being adopted in medical image processing, especially in mammography. They enhance traditional computer-aided diagnosis systems, aiming to maximize the performance levels of radiologists [10–12].

We used DICOM (Digital Imaging and Communications in Medicine) images in our study, which offer several advantages over PNG (Portable Network Graphics) images when used in deep learning applications. Firstly, DICOM images preserve the highest quality of pixel data, ensuring that no critical information is lost during image processing. This high-quality representation enhances the performance and accuracy of deep learning algorithms, leading to more reliable results [13]. Additionally, DICOM files contain extensive metadata relevant to the patient, study, and other contextual information. This rich set of metadata provides valuable additional information for analysis and interpretation, enabling more comprehensive and accurate deep learning models [14]. By leveraging these benefits, DICOM

images prove to be superior to PNG images in deep learning applications for radiological studies.

Before implementing an intelligent system, it is crucial to pre-process mammographic scans effectively to improve the quality of imaging outputs [5]. One of the key elements impacting the quality of imaging outputs is data noise, and Mitigating the detrimental impact of noise is a critical step in biomedical image processing, as any unwanted information in images is referred to as noise [15]. Image processing techniques, such as image enhancement, restoration, analysis, and compression, can be applied to the input data to improve performance and reduce noise and distortion during analysis [16].

Mammography has long been viewed as a potent method for screening breast cancer. Yet, its effectiveness is compromised in certain cases due to limitations like reduced sensitivity towards dense breasts and limited contrast features [17]. Breast density is determined by the ratio of fibrous and glandular tissue to fatty tissue in the breast. A higher proportion of fibrous and glandular tissue indicates denser breasts. It's common for about half of women to have dense breasts. Women with dense breasts face a marginally increased risk of breast cancer. The presence of dense tissue complicates tumor detection in mammograms, as both tumors and fibrous/glandular tissue appear white, potentially obscuring concerning findings. There's ongoing debate among experts regarding the additional tests needed for women with dense breasts who don't have elevated breast cancer risk due to factors like genetics or family history [2]. To improve tumor detection accuracy, image contrast enhancement becomes essential.

In our study, we employed automated mass segmentation in full-field mammograms to address the limitation of lacking significant mass boundary information in mass detection. This approach offers practical benefits for detecting and diagnosing breast cancer by eliminating the need for manual extraction of regions of interest (ROI) by radiologists. By automating this process, time-consuming and tedious tasks are minimized, making the detection and diagnosis more efficient. Based on the advantages of using artificial intelligence (AI), and deep learning techniques for automatic breast screening tasks, we propose a novel framework for automatic breast cancer segmentation to carry out lumps identification for an effective mass diagnosis in a mammography. The proposed framework capable to achieve a desirable results of segmentation with several pre-processing steps. The proposed study includes of removing artefacts, Li algorithm to minimize the cross-entropy between the foreground and the background, enhancing contrast using contrast-limited adaptive histogram equalization (CLAHE), normalizing, and median filtering. An effective augmentation approaches are used to address the issues of over- and under-fitting. Afterwards, a convolutional encoder-decoder network on the basis

of U-net is used to accurately identify suspicious masses. The contributions of this work are summarized as follows:

- Various pre-processing steps are performed to accurately segment the breast masses to achieve desired results.
- The proposed frame work is tested on two different datasets of INbreast and CBIS-DDSM images.
- Extensive comparisons are presented between the proposed model and the state-of-the-art techniques.

Related works

Over the past few years, deep learning methodologies have surged in prominence within the domain of medical imaging [18–20], particularly in breast cancer diagnosis [21]. Numerous research endeavors have showcased the prowess of deep learning algorithms in pinpointing and categorizing breast abnormalities across mammograms, breast thermograms [22], ultrasound captures [23], and MRI scans [24]. A notable strength of deep learning frameworks is their capacity to assimilate vast data volumes, enhancing their precision and adaptability. Further, the concept of transfer learning has gained traction; this technique involves priming a model using an expansive dataset before refining it with a more task-specific, smaller dataset. Predominantly, the encouraging outcomes from such research underscore the transformative potential of deep learning in refining the diagnostic accuracy for breast cancer, setting the stage for improved patient prognoses. As we delve into related works, the application of the deep learning models for breast cancer segmentation stands out as a critical area of exploration.

Sun et al. [25] presented the Attention-guided Dense Upsampling segmentation (AUNet). This technique aims to harness the most pertinent details from both high-level and low-level feature sets. In their approach, they employed Bilinear upsampling succeeded by a convolution layer. Subsequently, a Batch normalization layer was utilized as a substitute for the dense upsampling convolution. Notably, they refrained from applying any image augmentation or processing techniques. Their evaluations on the INbreast dataset, employing 5-fold cross-validation, yielded a Dice score of 79.1% for INbreast and 81.8% for CBIS-DDSM. However, cross-validation can sometimes result in overfitting on validation sets, especially when the fold count is minimal. In contrast, our proposed method achieved a Dice score of 87.98% for CBIS-DDSM and 85.61% for INbreast. Our training exclusively utilized the CBIS-DDSM dataset, and testing was conducted on 20% of CBIS-DDSM and all mass images from the INbreast dataset. This suggests that our strategy potentially offers better generalization across varied mammogram datasets.

Hou et al. [26] introduced MTLNet, an attentive multi-task learning network for breast cancer segmentation. By integrating group convolution and attention mechanisms, the model optimizes feature-learning, reducing redundancy. Their model's effectiveness was validated through a five-fold cross-validation. However, incorporating regularization techniques like dropout and weight decay could potentially enhance its performance on unseen data, suggesting avenues for future refinement.

Yan et al. [27] proposed an advanced deep detection system for automated mass localization that leverages a multi-scale fusion technique. For accurate mass delineation, they employed a convolutional encoder-decoder framework enriched with hierarchical and dense skip linkages. By benchmarking various architectures such as U-Net, CGAN, cascaded U-Net, and v19U-Net++, they observed that the collective average Dice score, when trained on both CBIS and INbreast, amounted to 80.44% for INbreast test visuals.

Li et al. [28] introduced a design where a densely connected convolutional network serves as the encoder and a U-Net with integrated attention gates functions as the decoder. The encoder's output is transformed into a gating signal vector to filter out irrelevant and noisy responses. Their method attained an F1 score of 82.24% on the DDSM dataset. In contrast, our model recorded an F1 score of 87.98% on the CBIS-DDSM dataset. While they relied solely on the DDSM dataset for both training and testing, there's a potential benefit in validating their approach across diverse datasets.

In their research, Zeiser et al. [29] utilized the U-Net framework to identify masses in mammograms, enhancing their approach with various data augmentation techniques like extraction from nine specific regions of interest, image zoom, and horizontal image inversion. Using the DDSM dataset for training and testing, they reported a Dice score of 79.39%, emphasizing the synergistic benefits of U-Net combined with data augmentation for mammogram analysis. Yet, their model encountered difficulties discerning masses that had similar densities to the surrounding breast tissue. Their exclusive dependence on the DDSM dataset for training (70%), validation (10%), and testing (20%) suggests the necessity for broader dataset validations. Interestingly, while they trained their model for 140 epochs, our proposed approach, trained for only 100 epochs, secured a Dice score of 87.98% on the CBIS-DDSM dataset.

Exploring different avenues, Vidal et al. [30] and Singh et al. [31] introduced unique lesion extraction and segmentation techniques. Vidal et al. combined multiple input-connected U-Net models in an ensemble approach, testing on the TCGA-BRCA database. Singh et al., on the other hand, employed max-mean and least-variance models, utilizing morphological operations and image gradient methods for tumor detection.

Further adding to the diversity of methodologies, Khan et al. [32] segmented breast cancer images using color thresholding and unsupervised segmentation, while another study [33] used a region-growing approach with dragonfly optimization for seed point determination in cancer detection. Wang et al. [34] presented a model combining principal component analysis with U-Net for hyperspectral image segmentation, achieving an impressive accuracy of 87.14%.

For MRI and ultrasound image analysis, Rahman et al. [35] and Byra et al. [36] contributed novel techniques. Rahman et al. enhanced MRI image quality using mean and Gaussian filters, while Byra et al. introduced a Selective Kernel U-Net for ultrasound image segmentation. Lastly, Ilesanmi et al. [37] improved breast ultrasound segmentation using contrast-constrained histogram equalization.

The integration of machine intelligence in supply chain management, as discussed by Myvizhi et al. [38], parallels the use of deep learning in medical imaging. Similarly, the challenges and opportunities in sustainable transportation outlined by Nabeeh et al. [39] reflect the balance of efficiency and accuracy in healthcare, akin to how advanced algorithms guide breast cancer segmentation. Additionally, Sallam's et al. [40] insights into IoT in supply chain management are comparable to the use of interconnected technologies in healthcare, highlighting shared concerns like data security and interoperability.

In summary, these studies collectively underscore the importance of AI and ML in specialized domains and their role in enhancing the precision and effectiveness of medical applications like breast cancer segmentation.

Materials and methodology

The suggested workflow for the segmentation process is illustrated in Fig. 1, with each stage represented as individual blocks.

Datasets

INbreast and CBIS-DDSM are two open datasets, used to validate the proposed model. A recent updated version of the DDSM database, CBIS-DDSM (Digital Database for Screening Mammography) [43, 44], includes images in the DICOM format. The CBIS-DDSM dataset is extensive, with a size surpassing 160GB and image resolutions roughly around 4000×6000 pixels. Additionally, it encompasses pixel-level annotations for masses and specific lesion pathology labels. Each view from the CBIS-DDSM, including the mediolateral oblique (MLO) and craniocaudal (CC) views, was used as a separate image. In our study, the binary segmentation masks from the CBIS-DDSM dataset were used to extract the ROIs. The dataset, comprising a total of 1588 images, was divided

into training and testing sets according to the BIRADS category. For this division, 20% of the cases, amounting to 358 images, were allocated for the testing set, with the remaining cases used for training. In Fig. 2, The top left subplot illustrates the distribution of BIRADS categories across the dataset, segmented into 'Test' and 'Training' groups. The top right subplot displays the distribution of cases by breast side. The bottom left subplot showcases the distribution of breast density categories, a critical factor in mammographic diagnosis. Finally, the bottom right subplot presents the count of different mammographic image views (CC and MLO), which are essential for a comprehensive evaluation of breast tissue. Together, these visualizations offer a multifaceted view of the dataset's composition, with a clear distinction between testing and training instances, facilitating an understanding of the data's structure and potential biases.

The INbreast dataset, referenced in [45, 46], comprises 410 mammographic images sourced from 115 individuals at St. John's Hospital's Breast Centre located in Porto, Portugal. 90 of these included ladies who had the cancer on both breasts. However, only 107 images with masses and precise delineations created by experts are used as testing set. The images in this database include two perspectives: MLO and (CC). The images in INbreast are officially as ".DICOM" format and have a size of 3328×4084 or 2560×3328 pixels. In Fig. 3 The top left panel displays the distribution of BIRADS assessment categories, indicating the range of breast cancer risk as interpreted by radiologists. The top right panel shows the distribution of the mammographic images by the breast side, denoting the prevalence of cases in either the left or right breast. The bottom left panel represents the distribution of breast tissue density, with categories ranging from A (almost entirely fatty) to D (extremely dense). The bottom right panel illustrates the distribution between the two types of mammographic views, CC and MLO, used in the screening process. These distributions provide insight into the dataset's composition and can help in understanding the variety of cases analyzed in the study.

Image pre-processing

In this section, we illustrate how to improve the quality of mammograms by removing artefact in background and enhance image contrast as illustrated in Fig. 4. To make any model learn features and patterns easily from the enhancement quality information. Achieving a great performance out of any model requires removing artefacts. We can see that mammograms contain the following flaws based on a visual inspection of the raw mammograms, as shown in Fig. 5:

1. Bright white borders or corners.
2. Artefacts floating around such as letters and labels.
3. The orientation in which the breasts face varies.

Table 1 Pros and limitations of some breast cancer segmentation models

Reference	Dataset	Segmentation method	Augmentation	Image Pre-processing	Limitations	Metrics
Shen et al. (2019) [41]	CBIS-DDSM INbreast	Pre-trained patch classifier(ResNet50,VGG16)			Downsized mammograms due to limited GPU memory, potentially affecting image quality and future model performance	AUC = 88%
Li et al. (2019) [28]	DDSM	U-Net with attention gates	✓	Gamma transform dot multiplication	Need to Validate the proposed method on different dataset	Dice = 82.24%
Sun et al. (2020) [25]	CBIS-DDSM INbreast	Attention-guided Dense Up-sampling segmentation	–	–	Evaluation with only 5 folds-cross validation which leads to overfitting	Dice = 81.8% Dice = 79.1%
Hou et al. (2021) [26]	CBIS-DDSM INbreast	Multitask learning network	✓	–	It's not a limitation; however, the model's performance on unseen data could be enhanced by adding regularization techniques, such as dropout and weight decay, to prevent overfitting	Dice = 86.3% Dice = 82.6%
Zeiser et al. (2020) [29]	DDSM	U-net	✓	CLAHE method	The model struggled to detect mass that its density being similar to the surrounding breast tissue	Accuracy = 85.95% Dice = 79.39%
Ramesh et al. (2021) [42]	CBIS-DDSM INbreast	DS U-Net		CLAHE method		Dice = 82.9% Dice = 79%
Our proposed study	CBIS-DDSM INbreast	U-net	✓	CLAHE method		Dice = 87.98% Dice = 85.61%

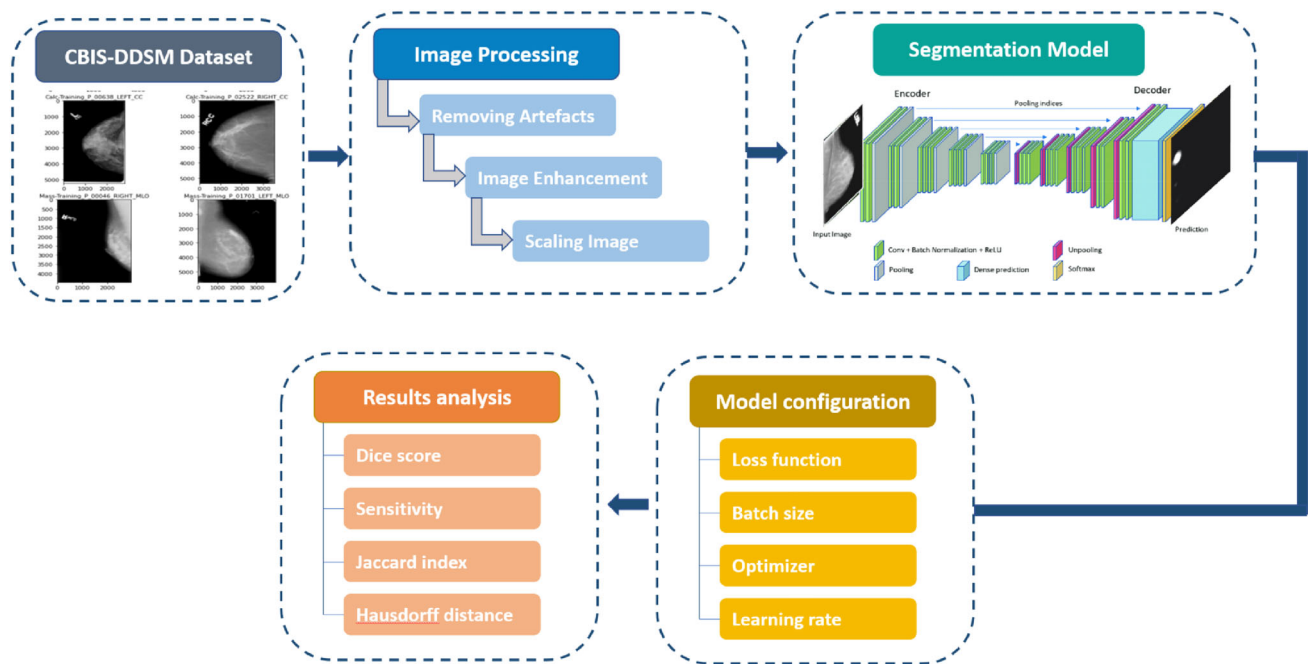


Fig. 1 Working flow of the entire segmentation process

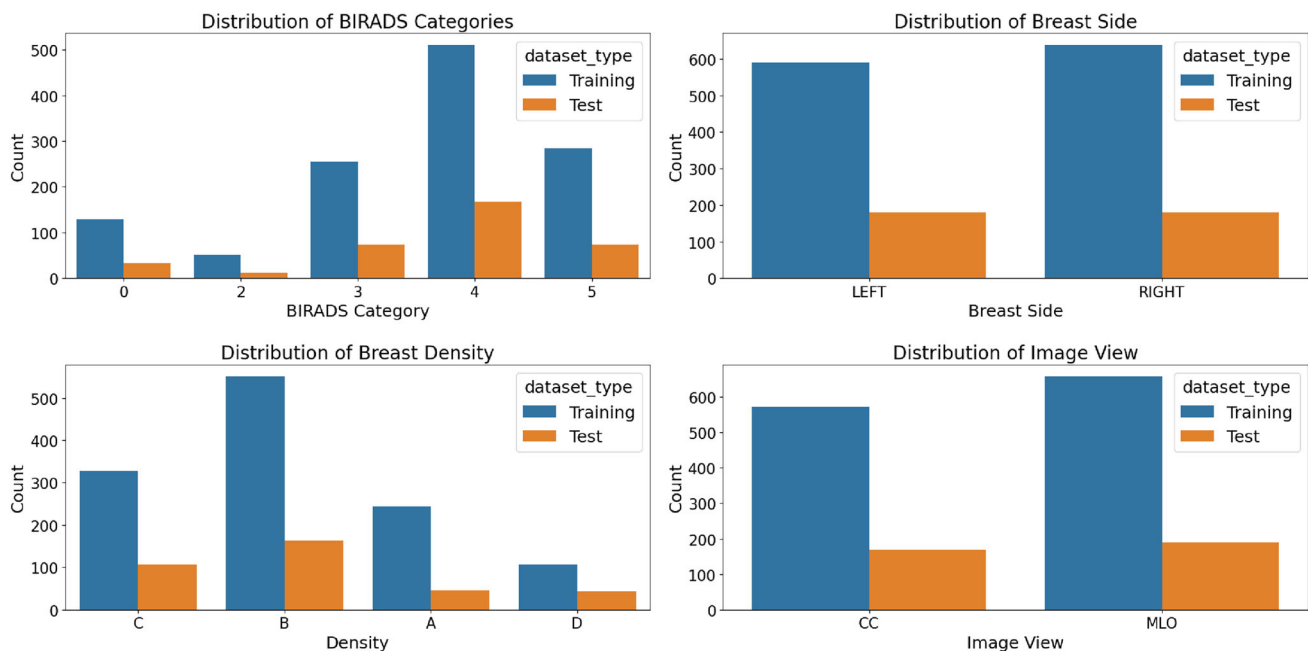


Fig. 2 Comparative Analysis of CBIS-DDSM Dataset Characteristics by Dataset Type

4. Low susceptibility to dense breasts.
5. The images have no fixed size.

Normalize images

Normalizing data is essential to ensure uniform distribution for every input parameter, in this context, each pixel. This step aids in expediting the model's training convergence. Normal-

ization is achieved by deducting the mean from each pixel and then dividing by its standard deviation, resulting in data distributed around a Gaussian curve centered at zero. However, for image inputs, positive pixel values are necessary, so the adjusted data is rescaled to fall within the [0, 1] range.

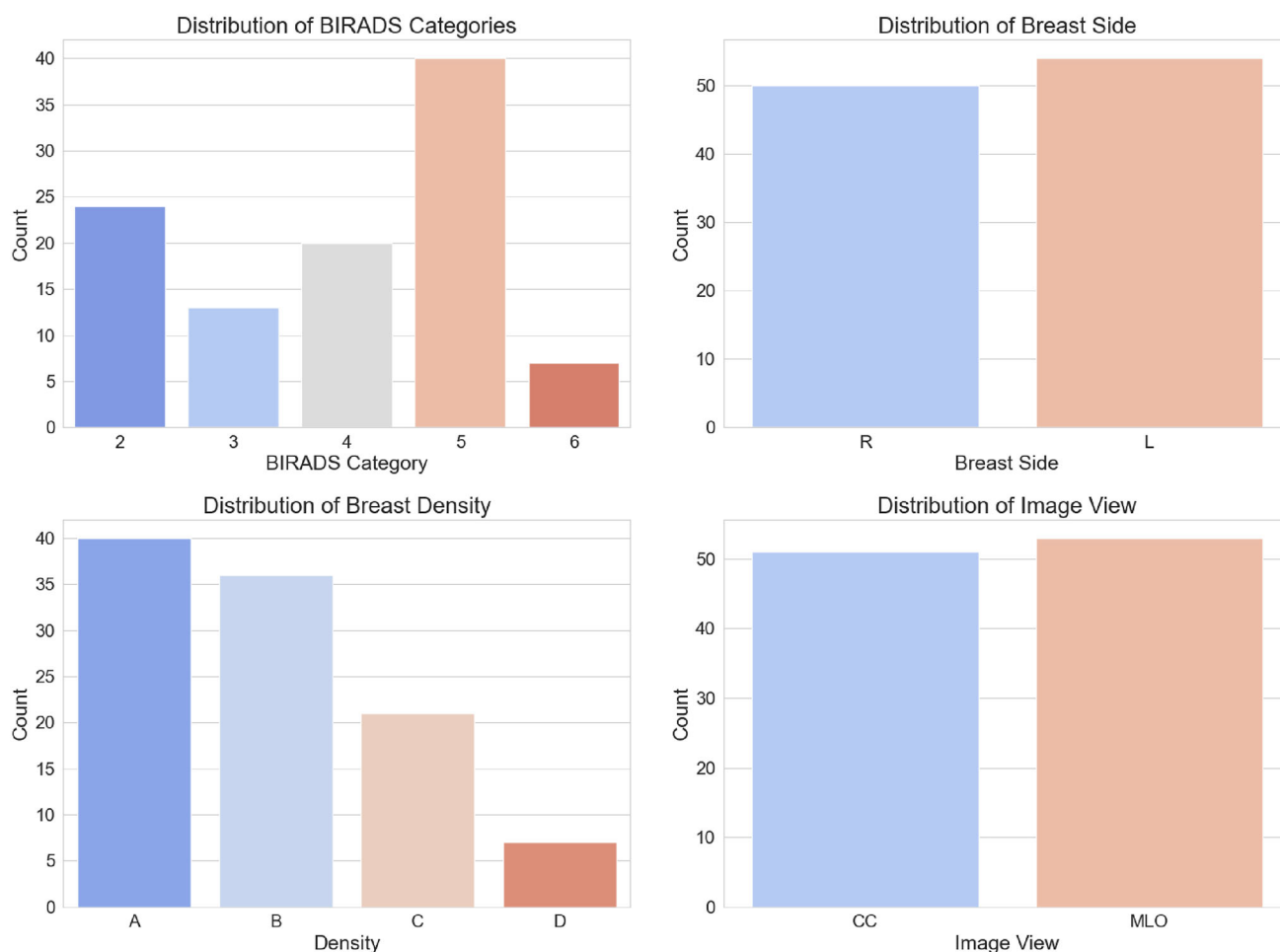


Fig. 3 Overview of INbreast dataset characteristics



Fig. 4 Pipeline diagram of image pre-processing

Right orientation mammogram

The orientation in which the breasts face varies. Some are looking to the left, while others are looking to the right. This is a problem since it may make image preprocessing more difficult. We must first right-orient the images in order for the algorithm to generalize across all mammograms as in Fig. 6. We just compare the number of nonzero pixels on both halves of the images to determine left-oriented breast images. This is a really basic method of detecting orientation, and it works since the background pixels are completely black, giving us sense of where the breasts are on either half of the image.

Removing artefacts

Unwanted items or areas can arise in images by accident, as demonstrated in Fig. 5. In the background, there are artefacts floating around. In practice, these artefacts are used by radiologists to distinguish between the left and right breasts, as well as to determine the scan's orientation. For example, "LCC" stands for "left breast, CC orientation". Also, there are some of the margins have bright white borders or corners. As a feature in the images that any model learns, these borders may generate an arbitrary edge. The removal of this border is critical since the tumor's intensity and that of this border

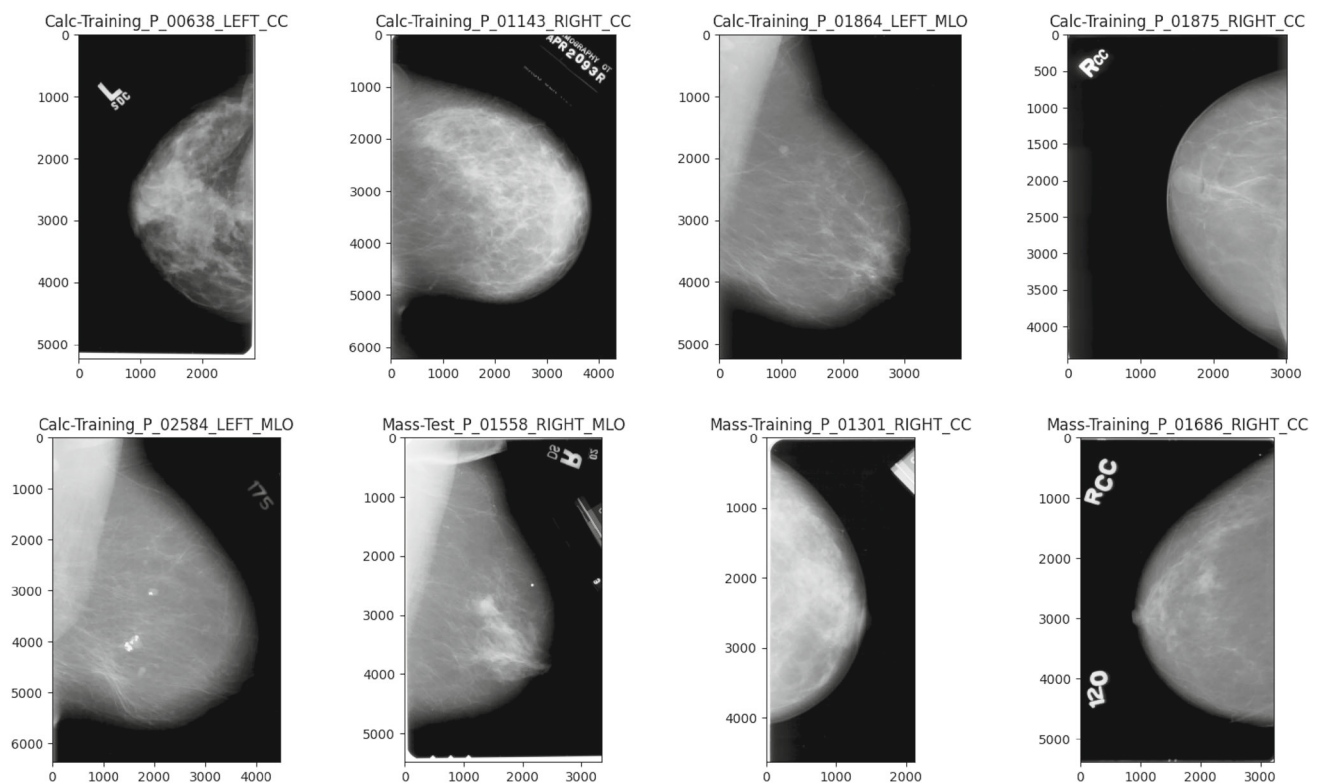


Fig. 5 Sample images of CBIS-DDSM dataset

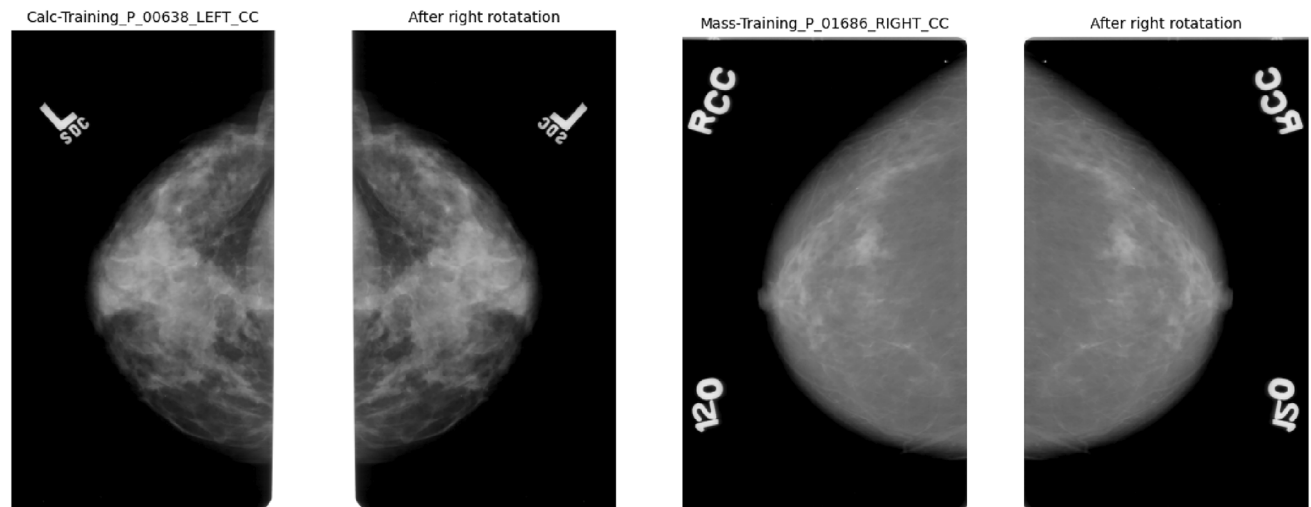


Fig. 6 Right orientation mammogram

are nearly identical, potentially affecting the performance of any model.

Thresholding algorithms

Masks are the most common way to remove or select certain areas of an image. Transforming an image to a binary image based on a pixel intensity threshold to assess whether or not it is of importance to us to isolate particular parts of an image. We use Image thresholding technique to construct

a mask. Masks are typically formed by performing one or more logical operations on an image.

The most basic thresholding method employs a manually defined image threshold. Using an automated threshold method on an image, on the other hand, calculates the numerical value of the image better than the human eye and can be easily repeated. We'll experiment with some Thresholding algorithms to see which thresholding strategies work best.

As shown in Fig. 7, Li and the Triangle technique appear to be doing well in our sample image. The other outcomes in this scenario are significantly worse. To transform our image to a binary image, we'll apply Li thresholding.

Li Thresholding algorithm

In 1993, Li and Lee introduced an innovative technique to ascertain the ideal threshold to differentiate the image's foreground from its background. They aimed to reduce the cross-entropy between the two means, which consistently provided optimal thresholding results. Before 1998, the method of determining the best threshold involved evaluating all potential thresholds and opting for the one yielding the least cross-entropy. However, Li and Tam later developed an iterative method, rooted in the gradient of the cross-entropy function, to expedite the identification of this threshold [47].

Mathematically, consider T as the threshold, and $F(T)$ and $B(T)$ as the respective means of the image's foreground and background. The iterative equation to deduce the optimal threshold can be depicted as:

$$T_{n+1} = \frac{F(T_n) + B(T_n)}{2} \quad (1)$$

Here, T_n stands for the threshold during the n th iteration. This equation progressively adjusts the threshold by computing the mean of the current foreground and background values. The iterations persist until the threshold stabilizes at a value that optimally reduces the cross-entropy function.

Applying mask

The label function of Scipy library returns an array in which each object in the input is assigned an integer index. Unless the output parameter is specified, it returns a tuple containing the array of object labels and the number of items detected, in which case only the number of objects is returned. A structural element determines how the objects are connected. Capable of providing two-way connections, the structural matrix must be centrosymmetric. A symmetric matrix about the center is known as a centrosymmetric matrix [48].

The input can be any value that is not zero and is considered to be a part of the object. A squared connectivity equal to one is used to construct the structuring element. By experimenting with different label values to select breast pixels, we select half of the 1st dimension and left quarter of the 2nd dimension of the mammogram image and any other pixel not connected to the label equals to zero. Figure 8 shows removing artefacts after applying the mask.

Enhancing contrast

To improve tumor detection accuracy, image contrast enhancement becomes essential. The Contrast Limited Adaptive Histogram Equalization (CLAHE) technique augments local

contrast in specific image areas. Compared to alternative Histogram Equalization techniques, CLAHE not only evens out the histogram but also maximizes entropy, making it particularly apt for medical imaging [17]. However, CLAHE's efficacy largely hinges on the clip limit, governing the amplification of noise in images. An improperly set clip limit might degrade image quality and introduce noise [49].

Data augmentation

To combat overfitting, we employed data augmentation methods to expand our training dataset. It is crucial to select appropriate augmentations that generate realistic images beneficial for our task.

In this study, we created augmented images by randomly flipping the horizontal and vertical axes by 50% and randomly adjusting brightness by 30%. The binary images used to represent the mass segmentation masks were cropped, resized, and enhanced to match their corresponding mammograms.

Network architecture

In 2015, Ronneberger et al. [50] introduced the U-Net structure tailored for semantic segmentation, specifically aimed at biomedical image processing. As depicted in Fig. 9, U-Net is primarily composed of an encoding and a decoding block. The encoder, through convolutional and max pooling layers, extracts image features. After feature acquisition in the encoder, the decoder upsamples these feature maps non-linearly, integrates them with skip pathways from the encoder, and processes them through a duo of 3X3 convolution layers, each followed by a ReLU activation function. A final convolution layer then allocates a probability to every pixel, and a pixel-wise sigmoid classifier refines the output. This blend allows U-Net to utilize features discerned at all depths, generating a final segmentation map with full resolution, as initial CNN layers often grasp basic features while deeper layers capture more advanced ones.

For our decoder, we employed transposed convolution layers, enlarging feature maps while halving the channel count. Post-convolution layers, we integrated a 0.5 dropout for enhanced network stability. The output from the decoder's corresponding segment is added to a transposed convolution's output at every stage. Padding is applied to the output map to ensure the output segmentation matches the input image's dimensions.

We employed the Adam optimizer [51] with an initial learning rate set at $1e-3$ and a momentum of 0.9 to reduce the Dice loss during training. Our training spanned 100 epochs, closely monitoring the dice score of the validation dataset. In each epoch, every image from the training set was utilized once, using a batch size of 24. The input mammogram images

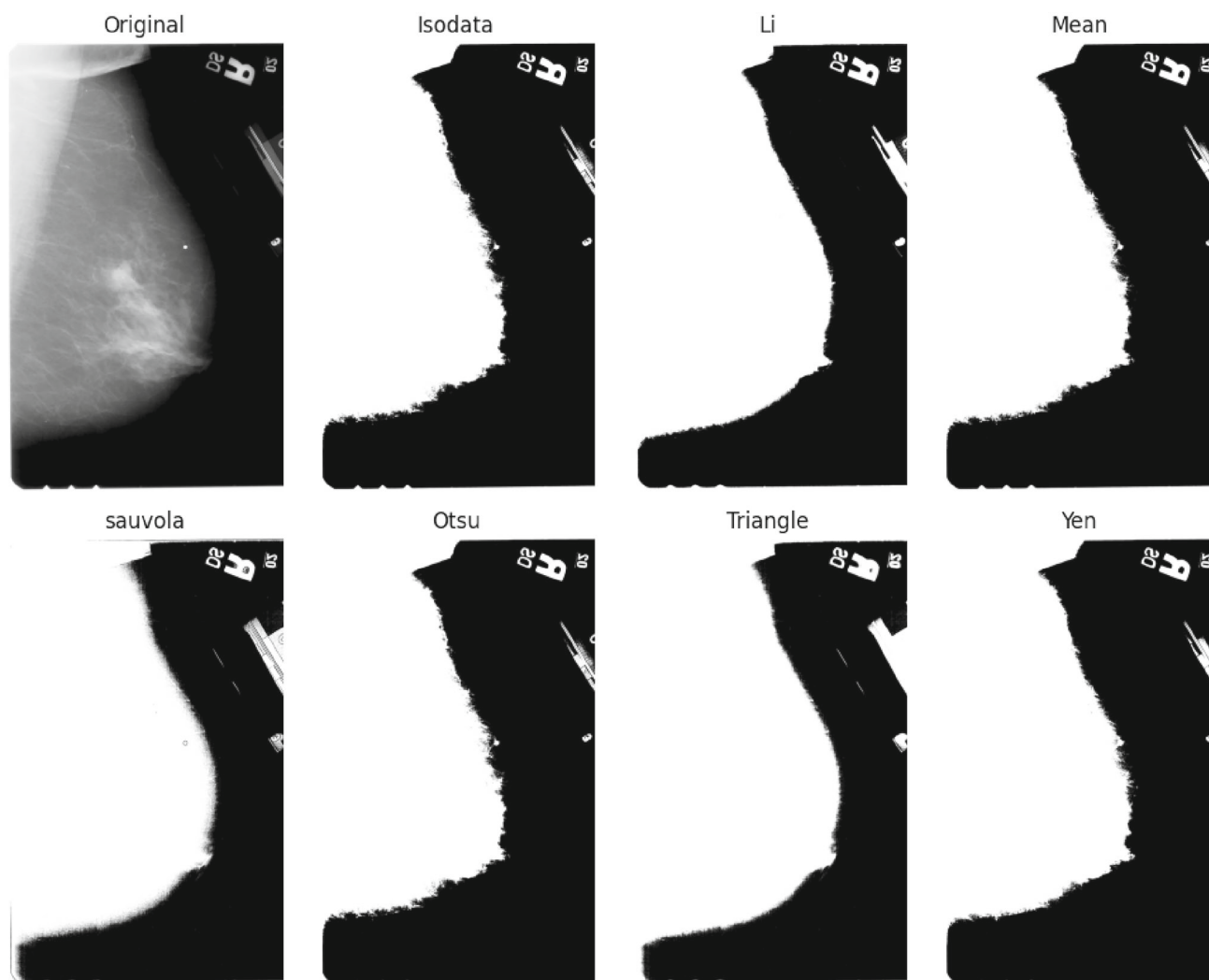


Fig. 7 Thresholding algorithms

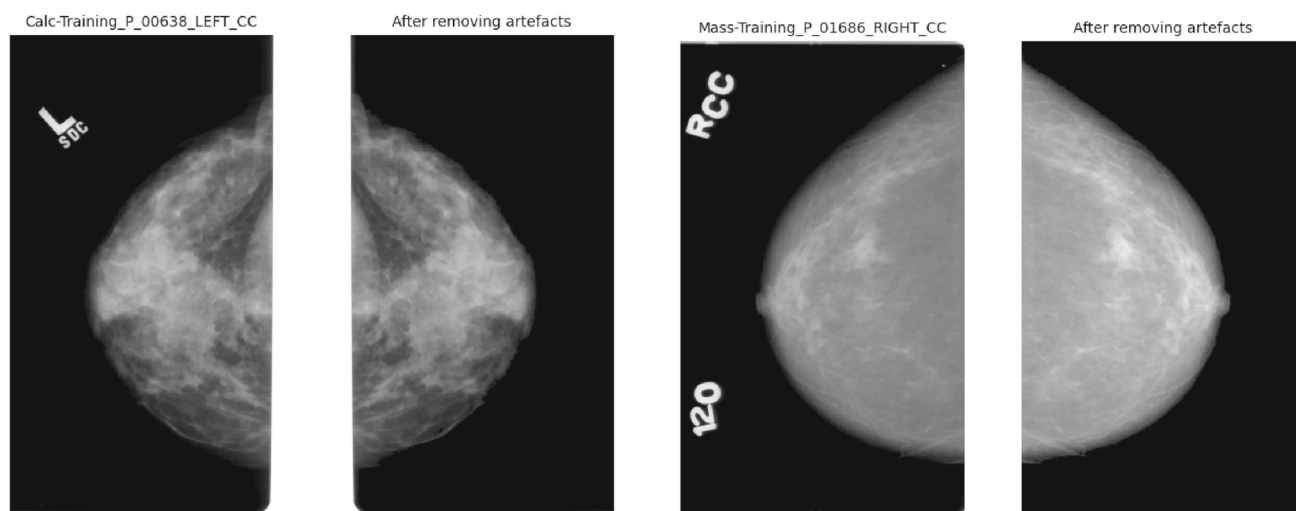


Fig. 8 After removing artefacts

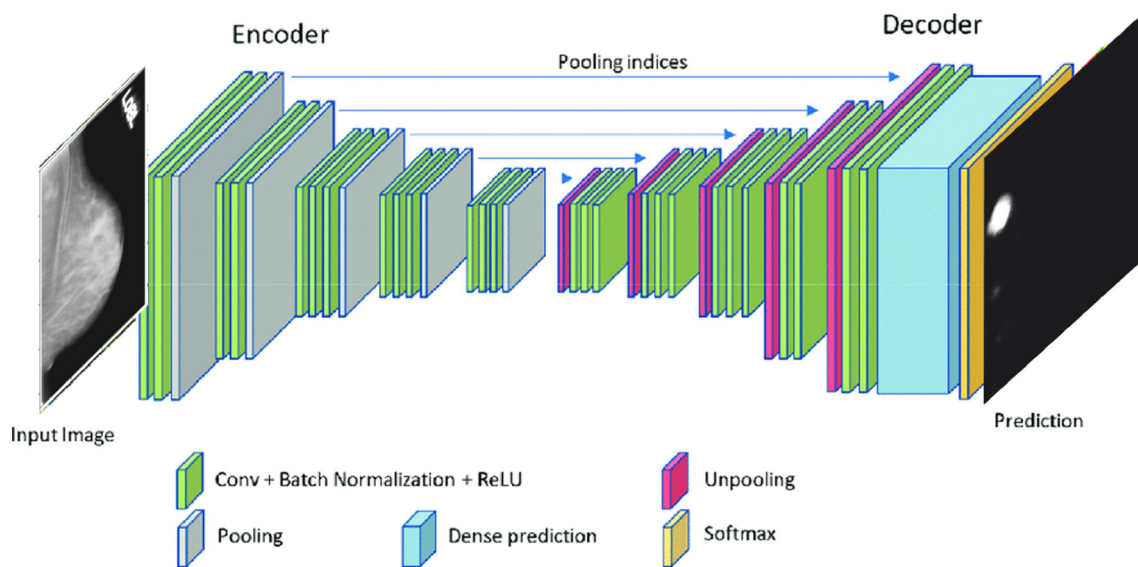


Fig. 9 U-net architecture

were resized to 224×224 dimensions. Our model's development, training, and testing were carried out using Python 3.9 and Tensorflow.

Evaluation metrics

In this study, a number of metrics: region level metrics and Pixel level metrics were utilized to evaluate how well the proposed approach was working. A few measurements at the pixel level are precision, sensitivity and f1 score. The precision evaluates the proportion of accurate predictions relating to the ground truth's mass region's pixels. The sensitivity (recall) evaluates the proportion of accurate predictions corresponding to the prediction map's normal region's pixels. The F1 score is the weighted average of precision and sensitivity. It factors in both the count of pixels correctly identified as mass and those mistakenly labeled as normal.

True positive (TP) denotes the count of pixels accurately classified as mass. True negative (TN) represents the number of pixels correctly identified as not being part of the mass. False positive (FP) indicates the count of pixels that are mistakenly classified as mass, even though they are normal. False negative (FN) signifies the number of mass pixels incorrectly labeled as normal.

Under-segmentation occurs when the mass is too small for the model to segment. Because there are many FN pixels, the sensitivity is low in this condition. over-segmentation occurs when the model segments normal region as a mass. as there are many FP pixels, the precision is affected.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (3)$$

$$\begin{aligned} F1 \text{ Score} &= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \\ &= \frac{2TP}{TP + \frac{1}{2}(FP + FN)} \end{aligned} \quad (4)$$

The area and boundary of corrective predictions are used by region level metrics for evaluation. The regions in the resulting prediction are evaluated using the Hausdorff distance, Jaccard coefficient, and dice score. Jaccard coefficient and dice score calculates how much the prediction mass and the ground truth overlap. There is no overlap between prediction and ground truth when the value is 0. A score of 1 denotes a perfect match between the prediction and ground truth. Hausdorff distance is the longest distance between any two points that are closest to one another and are located between the ground truth (GT) and the prediction (P) as cleared in Fig. 10.

$$\text{Dice score} = \frac{2|GT \cap P|}{|GT| + |P|} \quad (5)$$

$$\text{Jaccard coeff.} = \frac{|GT \cap P|}{|GT| + |P| - |GT \cap P|} \quad (6)$$

$$\text{Hausdorff distance} = \max \left\{ \max_{gt \in GT} \min_{p \in P} d(gt, p), \max_{p \in P} \min_{gt \in GT} d(p, gt) \right\} \quad (7)$$

Results

In this study, a large dataset of breast mammograms was used to be trained on an Unet model for mass detection.

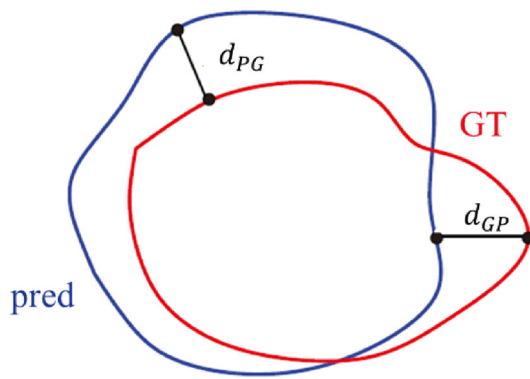


Fig. 10 Hausdorff distance

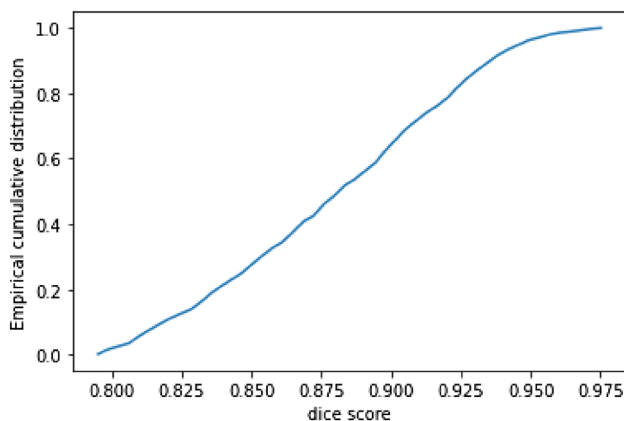


Fig. 11 Dice score empirical cumulative difference plot for the test images from the CBIS-DDSM dataset

It was demonstrated that the suggested model, which had been trained on a completely different dataset (CBIS-DDSM) than the tested dataset (INbreast), could be modified to accurately find masses in full mammogram. The suggested mass detection framework outperformed previous research in the literature in terms of better pixel-scale sensitivity and region-scale dice score.

Dice score evaluation

A more in-depth examination of the model's dice score metrics is depicted in the Empirical cumulative difference plot, as illustrated in Fig. 11. By mapping the empirical cumulative difference of the dice score, we gain a clearer insight into the alignment of predictions and actual values across the entirety of the CBIS-DDSM dataset. Given that the plot starts from a dice score of 0.8, it's evident that the U-Net effectively identifies lesions, as there are no predictions falling below this score. The steeper ascent in the U-Net plot further indicates the precision of most of its predictions.

Case studies

This section will feature specific cases illustrated to demonstrate the application of the methodology to CBIS-DDSM and INbreast images. The model's prediction outcome for sample images from both datasets is shown in Figs. 12, 13. The boundary of the projected regions is outlined using red lines, and the boundary of the ground truth is delineated using blue lines.

Based on both quantitative metrics and visual evaluations, our U-Net model demonstrates exceptional prowess in segmenting tumors within mammographic images. Notably, it is proficient not only in detecting minuscule tumors with a Dice score of 84.71% and a perfect sensitivity rate of 100% as in Fig. 12a, but also in diagnosing multiple tumors within a single image in Fig. 12b. The model's capability shines especially in challenging scenarios—accurately detecting tumors, regardless of their proximity to high-density regions or their diminutive size, as depicted in Fig. 12c and d. This is particularly noteworthy considering that radiologists might find these areas demanding and time-intensive for accurate diagnosis.

Despite being exclusively trained on the CBIS dataset, with no exposure to INbreast during training, our model showcased a commendable generalization capability by accurately predicting tumors in the INbreast dataset, achieving a Dice score of 85.61%. This accomplishment is particularly significant given the inherent differences in the nature of the two datasets, as illustrated in Fig. 13. The model's proficiency extends to both large and small tumors. Notably, even for diminutive tumors, as depicted in Fig. 13a, it achieved an impressive Dice score of 89.05%.

Comparison with state-of-the-art methods

In our evaluation, we selected state-of-the-art methods for a comparative analysis, adhering rigorously to the evaluation protocols of the CBIS-DDSM and INbreast datasets. Several of these methods are detailed in the literature review section.

The results in Tables 2 and 3 demonstrate that our proposed method significantly surpasses the performance of leading techniques. Evidently, our model showcases marked enhancements across crucial metrics compared to other strategies. For the CBIS-DDSM dataset, our approach achieved the top sensitivity of 90.58%, which is approximately 5% superior to previously documented methods. Additionally, there's a noticeable decrease in the Hausdorff distance metric by 0.78. For both CBIS-DDSM and INbreast datasets, our method stands out, registering the highest dice scores of 87.98% and 85.61%, respectively.

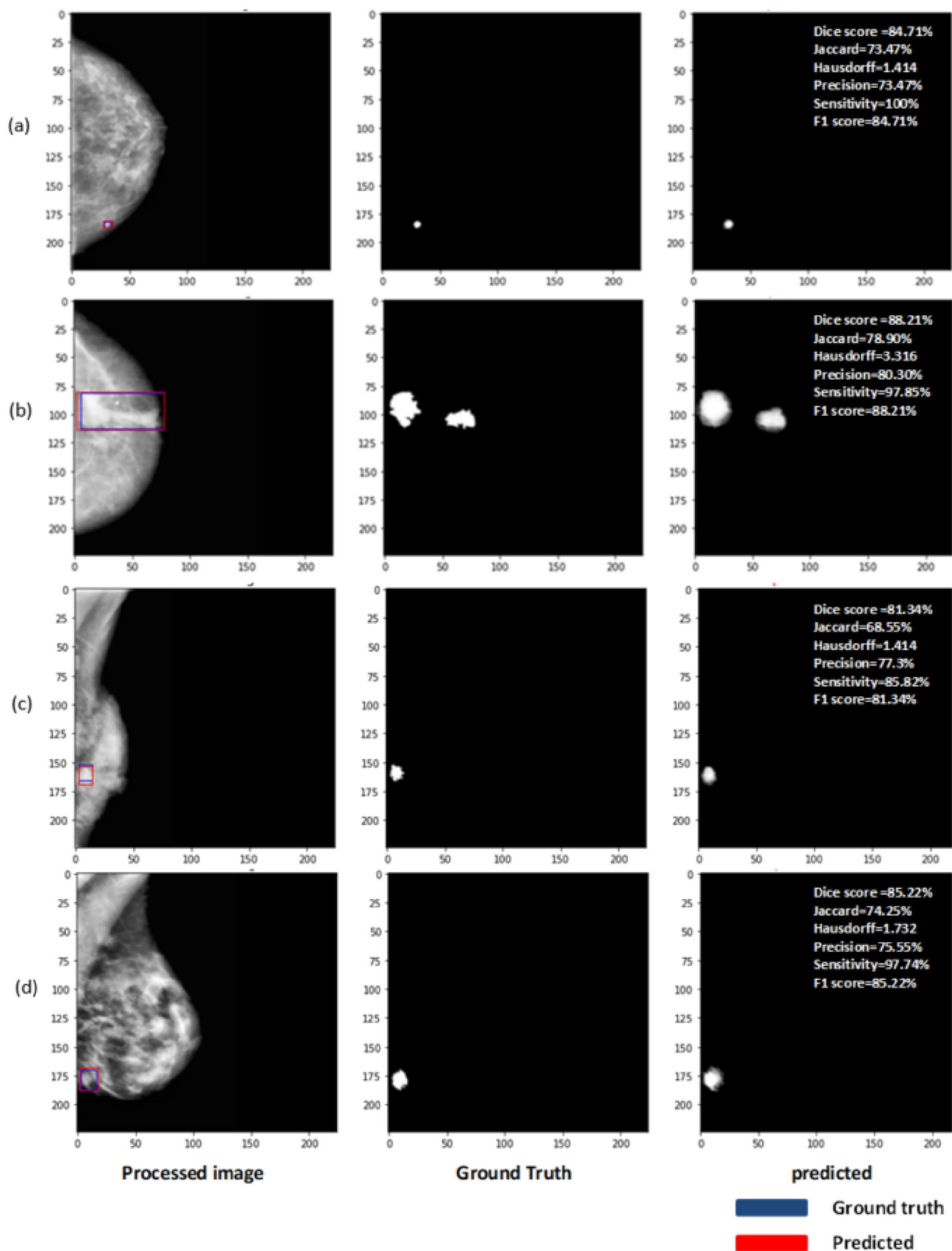


Fig. 12 Results of prediction for sample of test images from the CBIS-DDSM dataset, their IDs **a** Mass-Test-P00124-RIGHT-CC, **b** Mass-Test-P00116-RIGHT-CC, **c** Mass-Test-P00016-Left-MLO, and **d** Mass-Test-P01378-RIGHT-MLO

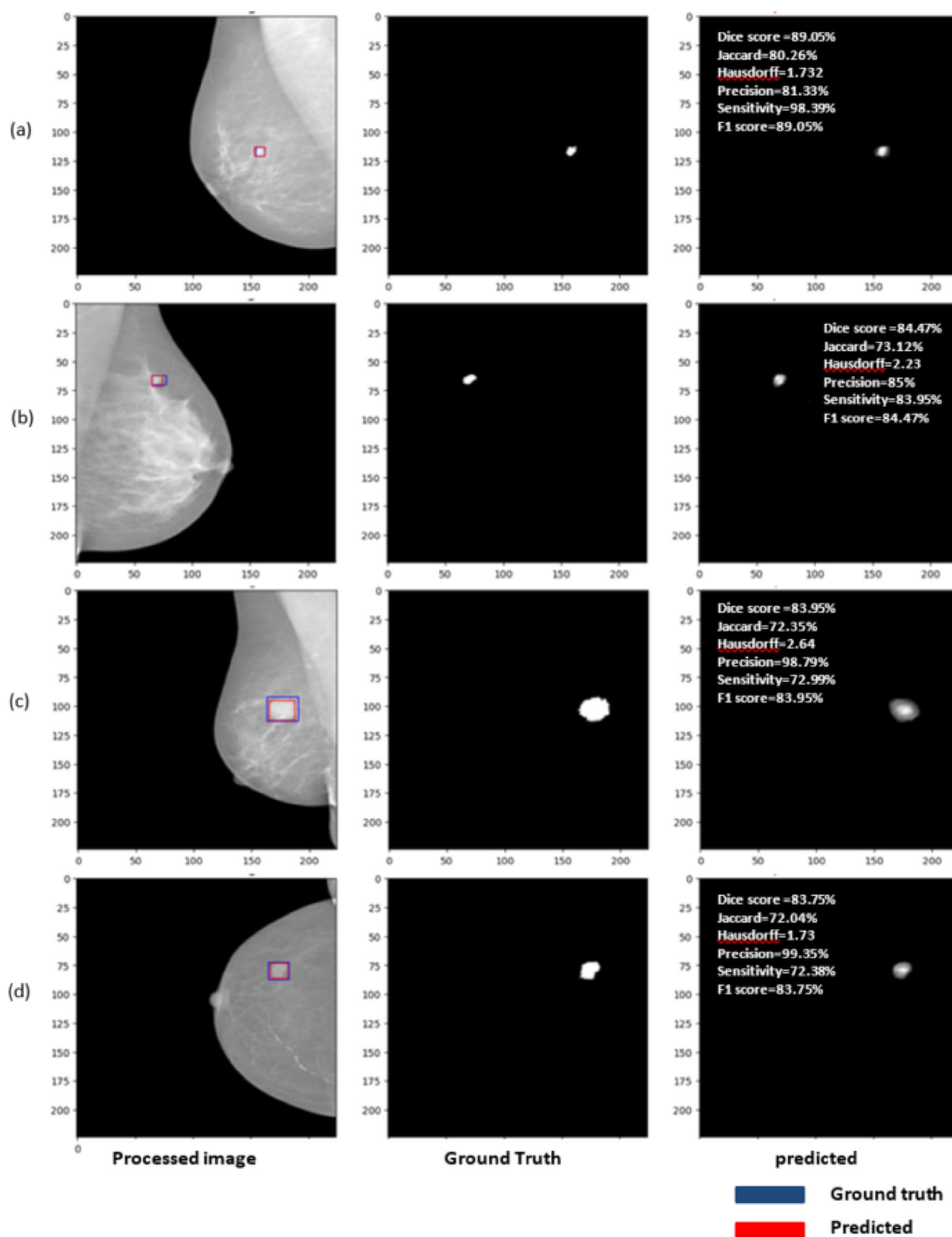


Fig. 13 Results of prediction for sample of test images from the INbreast dataset, their IDs **a** case-53581406, **b** case-20587810, **c** case-50996352, and **d** case-22670278

Table 2 Comparison of the proposed and state-of-the-art methods on CBIS-DDSM

Authors	Model	Dataset	Dice score	Sensitivity	Hausdorff distance
Sun et al. (2019) [25]	AUNet	CBIS-DDSM	81.8%	84.9%	2.96
Hou et al. (2021) [26]	MTLNet	CBIS-DDSM	86.3%	85.6%	2.64
Ramesh et al. (2021) [42]	DS U-Net	CBIS-DDSM	79%		2.43
Li et al. (2019) [28]	AD UNet	DDSM	82.24%	77.89%	–
Our proposed model		CBIS-DDSM	87.98%	90.58%	1.65

Table 3 Comparison of the proposed and state-of-the-art methods on INbreast

Authors	Model	Dataset	Dice score	Sensitivity	Hausdorff distance
Sun et al. (2019) [25]	AUNet	INbreast	79.1%	80.8%	6.85
Hou et al. (2021) [26]	MTLNet	INbreast	82.6%	79.2%	3.87
Ramesh et al. (2021) [42]	DS U-Net	INbreast	82.9%	81%	1.86
Singh et al. (2020) [52]	cGAN	INbreast	68.69%	–	–
Our proposed model		INbreast	85.61%	81.26 %	6.92

Conclusion

This research introduces a holistic approach for segmenting masses in digital mammography. The process encompasses data augmentation and image pre-processing, where the contrast of the mammograms is improved using CLAHE. Lesions are then identified using a deep supervised U-Net, eliminating the need for manual feature extraction or the selection of specific parameters like machine learning algorithms.

A number of metrics at the pixel and region levels are employed to assess the model's performance. When compared to other state-of-the-art approaches, the proposed architecture is determined to perform better segmentation outcomes on masses of completely different sizes and shapes with an overall average Dice of dice score 87.98% for CBIS-DDSM and 85.61% for INbreast. We used 80% of CBIS-DDSM for training and test on 20% of CBIS-DDSM and all mass images of INbreast dataset.

Building upon the findings presented, our research has yielded a model that not only excels in accuracy but also demonstrates the practical viability of deep learning in medical diagnostics. The utilization of advanced image pre-processing techniques, such as CLAHE, has resulted in improved visibility of mammographic features, thus enabling the deep supervised U-Net to perform precise lesion segmentation. This negates the need for manual feature extraction, streamlining the diagnostic process.

The employment of various evaluation metrics at both pixel and region levels has provided a comprehensive understanding of the model's performance, affirming its superiority over other state-of-the-art methods. The ability of our model to adeptly handle masses of varying sizes and shapes underscores its potential as a valuable tool in clinical settings.

Moreover, the training and testing on distinct subsets of the CBIS-DDSM and INbreast datasets have demonstrated the robustness and adaptability of our approach.

An additional advantage of our proposed deep U-Net framework in mammography image analysis is its significant contribution to the early detection of breast cancer. Key clinical benefits include:

1. **Enhanced Diagnostic Accuracy:** The framework's capability to discern subtle nuances in mammograms elevates diagnostic precision. This is crucial for identifying minor tissue changes that may be indicative of early-stage breast cancer.
2. **Expedited Analysis Process:** Our approach streamlines the image analysis procedure, thereby reducing response times. This acceleration is vital for the early detection of breast cancer, as it allows for quicker intervention.
3. **Reduction in Human Error:** The deep U-Net model's automated and precise analysis minimizes the likelihood of human error, which is a critical factor in medical diagnostics. By relying on advanced algorithmic processing, the system ensures a more consistent and reliable interpretation of mammographic images.

Future work

Future efforts will concentrate on further advancing validating and integrating our breast cancer segmentation models into clinical practice. This includes testing with varied datasets for broader applicability and implementing advanced deep learning methods for enhanced performance. Key to our approach is conducting clinical trials with healthcare professionals to ensure the model's efficacy and compliance with healthcare standards, ultimately aiming to

embed AI-driven diagnostics into routine clinical workflows for better breast cancer management.

Acknowledgements On behalf of all authors, the corresponding author states that there is no conflict of interest.

Funding Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

Data availability In the study we used two datasets which is available in a public repository: firstly, **CBIS-DDSM** which is available in a public repository of The Cancer Imaging Archive (**TCIA**) which hosts and identifies a sizable collection of openly accessible medical images of cancer. TCIA is run by the Frederick National Laboratory for Cancer Research and supported by the National Cancer Institute of the United States' Cancer Imaging Program. Secondly, **INBreast** contains images that were taken at the Breast Center in the Hospital de So Joo in Porto, Portugal.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. World health organization, a. URL <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>. Accessed 26 Mar 2022
2. American cancer society, b. URL <https://www.cancer.org/cancer/breast-cancer.html>. Accessed 01 May 2022
3. Hamed G, Marey M, Amin S, Tolba M (2021) Comparative study and analysis of recent computer aided diagnosis systems for masses detection in mammograms. *Int J Intell Comput Inf Sci* 21:33–48. <https://doi.org/10.21608/ijicis.2021.56425.1050>
4. Bozek J, Mustra M, Delac K, Grgic M (2009) A Survey of image processing algorithms in digital mammography, pages 631–657. https://doi.org/10.1007/978-3-642-02900-4_24
5. Reddy BA, Sami A, Mirjam J, Bharanidharan S, Krishnan K, Adnan A (2021) Preprocessing of breast cancer images to create datasets for deep-cnn. *IEEE Access* 9:33438–33463. <https://doi.org/10.1109/ACCESS.2021.3058773>
6. Yong LS, Ik JS, Seulhee J, Jae CI, Cheol-Hee A (2014) Targeted multimodal imaging modalities. *Adv Drug Deliv Rev* 76:60–78. <https://doi.org/10.1016/j.addr.2014.07.009>
7. El-Hag Noha A, Ahmed S, El-Banby Ghada M, Walid El-Shafai, Khalaf Ashraf AM, Waleed Al-Nuaimy, El-Samie Fathi Abd E, El-Hoseny Heba M (2021) Utilization of image interpolation and fusion in brain tumor segmentation. *Int J Numer Methods Biomed Eng* 37(8):e3449. <https://doi.org/10.1002/cnm.3449>
8. Michael H, Subrata C, Biswajeet P, Manoranjan P, Douglas G, Anwaar U-H, Abdullah A (2021) Deep mining generation of lung cancer malignancy models from chest x-ray images. *Sensors*. <https://doi.org/10.3390/s21196655>
9. Horry Michael J, Subrata C, Biswajeet P, Manoranjan P, Jing Z, Wen LH, Datta BP, Rajendra AU (2023) Development of debiasing technique for lung nodule chest x-ray datasets to generalize deep learning models. *Sensors*. <https://doi.org/10.3390/s23146585>
10. El-Shafai W, El-Hag Noha A, El-Banby Ghada M, Khalaf Ashraf AM, Soliman Naglaa F, Algarni Abeer D, El-Samie Fathi E. Abd (2021) An efficient cnn-based automated diagnosis framework from covid-19 ct images. *Comput Mater Continua* 69(1), 1323–1341. <https://doi.org/10.32604/cmc.2021.017385>
11. El-Hag Noha A, Ahmed S, Walid E-S, El-Hoseny Heba M, Khalaf Ashraf AM, El-Fishawy Adel S, Waleed A-N, Abd El-Samie Fathi E, El-Banby Ghada M (2021) Classification of retinal images based on convolutional neural network. *Microsc Res Techn* 84(3):394–414. <https://doi.org/10.1002/jemt.23596>
12. Khalil Hager, El-Hag Noha A, Sedik Ahmed, El-Shafai Walid, Mohamed Abd, Khalaf Ashraf AM, El-Fishawy Adel, El Banby Ghada, El-Samie Fathi Abd (2019) Classification of diabetic retinopathy types based on convolution neural network (cnn). 12
13. Varma Dandu R (2012) Managing dicom images: tips and tricks for the radiologist. *Indian J Radiol Imaging* 22:4–13. <https://doi.org/10.4103/0971-3026.95396>
14. Novaes Magdala de A (2020) Telecare within different specialties, pages 185–254. Elsevier, <https://doi.org/10.1016/B978-0-12-814309-4.00010-0>
15. Wang S, Emre CM, Zhang Y-D, Yu X, Lu S, Yao X, Zhou Q, Miguel M-G, Tian Y, Gorriz Juan M, Tyukin I (2021) Advances in data preprocessing for biomedical data fusion: an overview of the methods, challenges, and prospects. *Inf Fusion* 76:376–421. <https://doi.org/10.1016/j.inffus.2021.07.001>
16. da Silva E, Mendonca G (2005) Digital image processing, pages 891–910. 12. ISBN 9780121709600. <https://doi.org/10.1016/B978-012170960-0/50064-5>
17. Susama B, Kim Gaik T, Audrey H, Sanjoy Kumar D. Image processing and machine learning techniques used in computer-aided detection system for mammogram screening - a review. *Int J Electr Comput Eng (IJECE)*, 10:2336, 2020. <https://doi.org/10.11591/ijece.v10i3.pp2336-2348>
18. Maryam F, Subrata C, Biswajeet P, Oliver F, Datta BP, Hossein C, Rajendra A (2024) Deep learning techniques in pet/ct imaging: a comprehensive review from sinogram to image space. *Comput Methods Progr Biomed* 243:107880. <https://doi.org/10.1016/j.cmpb.2023.107880>
19. Shehzadi S, Hassan Muhammad A, Rizwan M, Kryvinska N, Vincent K (2022) Diagnosis of chronic ischemic heart disease using machine learning techniques. *Comput Intell Neurosci* 3823350:2022. <https://doi.org/10.1155/2022/3823350>
20. Deshpande N M, Gite S, Pradhan B, Kotecha K, Alamri A (2022) Improved otsu and kapur approach for white blood cells segmentation based on lebltbo optimization for the detection of leukemia. *Math Biosci Eng*.
21. Liu J, Lei J, Ou Y, Zhao Y, Tuo X, Zhang B, Shen M (2022) Mammography diagnosis of breast cancer screening through machine learning: a systematic review and meta-analysis. *Clin Exp Med* 23(3):1–16. <https://doi.org/10.1007/s10238-022-00895-0>
22. Roslidar R, Mohd S, Khairun S, Biswajeet P, Fitri A, Maimun S, Khairul M (2022) Brecnet: a high-accuracy breast thermogram classifier based on mobile convolutional neural network. *Math Biosci Eng* 19:1304–1331. <https://doi.org/10.3934/mbe.2022060>

23. Webb Jeremy M, Adusei Shaheeda A, Yinong W, Naziya S, Kalie A, Meixner Duane D, Fazzio Robert T, Mostafa F, Azra A (2021) Comparing deep learning-based automatic segmentation of breast masses to expert interobserver variability in ultrasound imaging. *Comput Biol Med* 139:12. <https://doi.org/10.1016/j.compbiomed.2021.104966>
24. Hirsch L, Huang Y, Luo S, Rossi Saccarelli C, Lo Gullo R, Daimiel Naranjo I, Bitencourt AG, Onishi N, Ko ES, Leithner D, Avendano D, Eskreis-Winkler S, Hughes M, Martinez DF, Pinker K, Juluru K, AE El-Rowmeim, Elnajjar P, Morris EA, LC Parra, Sutton EJ (2021) Radiologist-Level Performance Using Deep Learning for Segmentation of Breast Cancers on MRI. *Radiol: Artif Intell* 4(1). <https://doi.org/10.1148/ryai.200231>
25. Sun H, Li C, Liu B, Liu Z, Wang M, Zheng H, Feng DD, Wang S (2020) Aunet: attention-guided dense-upsampling networks for breast mass segmentation in whole mammograms. *Phys Med Biol* 65(5):055005. <https://doi.org/10.1088/1361-6560/ab5745>
26. Xuan H, Yunpeng B, Yefan X, Ying L (2021) Mass segmentation for whole mammograms via attentive multi-task learning framework. *Phys Med Biol* 66(10):105015. <https://doi.org/10.1088/1361-6560/abfa35>
27. Yan Y, Conze P-H, Quellec G, Lamard M, Cochenier B, Coatrieux G (2021) Two-stage multi-scale breast mass segmentation for full mammogram analysis without user intervention. *Biocybern Biomed Eng* 41(2):746–757
28. Li Shuyi, Dong Min, Guangming Du, Xiaomin Mu (2019) Attention dense-u-net for automatic breast mass segmentation in digital mammogram. *IEEE Access* 7:59037–59047
29. Zeiser Felipe A, da Costa CA, Zonta T, Marques NMC, Roehe AV, Moreno M, da Rosa RR (2020) Segmentation of masses on mammograms using data augmentation and deep learning. *J Digit Imaging* 33:858–868
30. Joel V, Vilanova Joan C, Robert M et al (2022) A u-net ensemble for breast lesion segmentation in dee mri. *Comput Biol Med* 140:105093
31. Anuj Kumar S, Bhupendra G (2015) A novel approach for breast cancer detection and segmentation in a mammogram. *Procedia Comput Sci* 54:676–682
32. Khan AM, El-Daly H, Simmons E, Rajpoot NM (2013) Hymap: a hybrid magnitude-phase approach to unsupervised segmentation of tumor areas in breast cancer histology images. *J Pathol Inform* 4(2):1
33. Punitha S, Amuthan A, Joseph SK (2018) Benign and malignant breast cancer segmentation using optimized region growing technique. *Futur Comput Inform J* 3(2):348–358
34. Wang J, Wang Y, Tao X, Li Q, Sun L, Chen J, Zhou M, Menghan H, Zhou X (2021) Pca-u-net based breast cancer nest segmentation from microarray hyperspectral images. *Fund Res* 1(5):631–640
35. Rahman M, Hussain M G, Hasan M R, Babe S, Akter S (2020) Detection and segmentation of breast tumor from mri images using image processing techniques. In 2020 fourth international conference on computing methodologies and communication (ICCMC), pages 20–724. IEEE,
36. Byra M, Jarosik P, Szubert A, Galperin M, Ojeda-Fournier H, Olson L, O'Boyle M, Comstock C, Andre M (2020) Breast mass segmentation in ultrasound with selective kernel u-net convolutional neural network. *Biomed Signal Process Control* 61:102027
37. Enitan IA, Utairat C, Makhanov Stanislav S (2021) A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning. *Biocybern Biomed Eng* 41(2):802–818
38. Myvizhi M, Ali Ahmed (2023) Sustainable supply chain management in the age of machine intelligence: Addressing challenges, capitalizing on opportunities, and shaping the future landscape. *Sustain Mach Intell J* <https://doi.org/10.61185/SMIJ.2023.33103>
39. Nabeeh N (2023) Assessment and contrast the sustainable growth of various road transport systems using intelligent neutrosophic multi-criteria decision-making model
40. Sallam K, Mohamed M, Mohamed A W (2023) Internet of things (iot) in supply chain management: challenges, opportunities, and best practices
41. Li S, Margolies Laurie R, Rothstein Joseph H, Eugene F, Russell M, Weiva S (2019) Deep learning to improve breast cancer detection on screening mammography. *Sci Rep* 9:12495. <https://doi.org/10.1038/s41598-019-48995-4>
42. Ravitha Rajalakshmi N, Vidhyapriya R, Elango N, Nikhil R (2021) Deeply supervised u-net for mass segmentation in digital mammograms. *Int J Imaging Syst Technol* 31(1):59–71. <https://doi.org/10.1002/ima.22516>
43. Sawyer LR, Francisco G, Assaf H, Kawai MK, Mia G, Rubin Daniel L (2017) Data descriptor: a curated mammography data set for use in computer-aided detection and diagnosis research. *Sci Data*. <https://doi.org/10.1038/sdata.2017.177>
44. Kenneth C, Bruce V, Kirk S, John F, Justin K, Paul Koppel, Stephen M, Stanley P, David M, Michael P, Lawrence T, Fred P (2013) The cancer imaging archive (tcia): maintaining and operating a public information repository. *J Digit Imaging* 26:1045–1057. <https://doi.org/10.1007/s10278-013-9622-7>
45. Mei-Ling H, Ting-Yu L (2020) Dataset of breast mammography images with masses. *Data Brief* 31:105928. <https://doi.org/10.1016/j.dib.2020.105928>
46. Moreira Inês C, Inês DIAL, António C, João CM, Cardoso Jaime S (2012) Inbreast: toward a full-field digital mammographic database. *Acad Radiol* 19(2):236–248. <https://doi.org/10.1016/j.acra.2011.09.014>
47. Li CH, Tam PKS (1998) An iterative algorithm for minimum cross entropy thresholding. *Pattern Recog Lett* 19(8):771–776. [https://doi.org/10.1016/S0167-8655\(98\)00057-9](https://doi.org/10.1016/S0167-8655(98)00057-9)
48. Weaver JR (1985) Centrosymmetric (cross-symmetric) matrices, their basic properties, eigenvalues, and eigenvectors. *Am Math Monthly* 92(10):711–717
49. Kuran U, Kuran EC (2021) Parameter selection for clahe using multi-objective cuckoo search algorithm for image contrast enhancement. *Intell Syst Appl* 12:200051
50. Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18, pages 234–241. Springer,
51. Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. *dec. arxiv.org*
52. Kumar SV, Rashwan Hatem A, Santiago R, Farhan A, Nidhi P, Mostafa SMD, Kamal SA, Meritxell A, Miguel A, Domenec P et al (2020) Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network. *Expert Syst Appl* 139:112855

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