Breast Mass Detection and Segmentation Using Multi-scale Morphological Sifting and K-Means Clustering

Abdelrahman HABIB, Muhammad Zain Amin, Md Imran Hossain, Andrew Dwi Permana, Anita Zhudenova

June 18, 2023

Report submitted for **Image Processing and Analysis 2023** at the University of Cassino and Southern Lazio, Cassino, Italy

1 Introduction

Breast cancer not only holds the highest incidence rate amongst all types of cancer affecting women globally but also ranks as the leading cause of cancer-related mortality [9]. However, early detection and diagnosis of breast cancer can substantially reduce the death rate and provide more treatment options for patients at risk [3]. Several studies provide evidence of a significant reduction (between 20 and 30%) in breast cancer mortality due to early screening programs. Mammography is considered one of the most efficient primary tools for breast screening and plays a crucial role in decreasing death from breast cancer. There are a significant number of features in mammograms, such as mass and microcalcification, can be assessed in order to detect the existence of breast cancer. The presence of breast mass in the mammograms indicates the likelihood of breast cancer. The detection of breast mass is considered the most difficult challenge due to irregular size and shape, and lack of contrast of masses. To address this challenge, the researchers proposed several computer-aided diagnoses (CAD) systems using advanced image processing and machine learning to interpret mammogram images, which assist in identifying masses accurately even though the size is very small. The CAD systems for the detection of breast cancer using mammograms have already displayed a promising outcome for radiologists.

2 Materials and Methods

Our proposed approach for mass detection and segmentation consists of five main steps for region segmentation, which are pre-processing, region candidate generation using multi-scale morphological sifting, mean shift filtering, k-means clustering, and finally post-processing. All of those steps are explained in detail in the following subsections. As we used ucasML tool for obtaining a classification score for each region candidate, we also included the feature extraction process and the evaluation methods used in this section. The complete code for this project is publicly available at the following repository *Breast Mass Detection and Segmentation Using Multi-scale Morphological Sifting and K-Means Clustering GitHub repository*: click here.

2.1 Mammographic Dataset

The dataset used in this project includes various examples of normal mammograms, mammograms with masses, mammograms with calcifications, architectural distortions, asymmetries, and images with multiple

findings. A total of 410 images, depending on the compression plate used during the image acquisition, each image has a matrix size of either 3328 x 4084 or 2560 x 3328 pixels [7]. Calcification is present in 301 images out of 410, the tag "cluster" was only used in 27 sets of calcifications, in 21 images, which is around 1.3 clusters per image. Out of these 21 images, only 2 images did not have a single calcification annotation. In 299 images a total number of 6880 calcifications were individually detected [7]. Around 86% of the cases in this dataset are malignant. In 107 images, there are 116 masses with the average mass size being 479 mm2 and standard deviation of 619 mm2, the area of the smallest and largest mass is 15 mm^2 and 3689 mm^2 respectively.

2.2 Pre-processing

As the mammograms are large in size, the raw images for both positive and negative labels were pre-processed and re-scaled to a lower resolution to speed up the process. The raw positive images with breast mass regions were segmented by thresholding and the redundant black background has been cropped away as in [6, 5] by extracting the largest connected region in the binary image. A contrast limited adaptive histogram equalization (CLAHE) was then applied with a tile grid size of (4, 4) and a clip limit of 8. These parameters were determined experimentally along the entire process. The breast region is then sub-sampled to 1/4 of its original size using bi-cubic interpolation. Finally, the breast region is normalized to 16-bit. This pre-processing was also applied on negative candidate images only for the purpose of using them for segmentation evaluation. The raw images (both positive and negative candidates) were also cropped and sub-sampled to match the pre-processed images dimensions for using them in feature extraction. Fig. 1 displays the intermediate results of the pre-processing stage.

2.3 Region candidate generation using a Multi-scale Morphological Shifting (MMS)

Following a similar implementation of multi-scale grayscale morphological filters by using multiple pairs of oriented linear structuring elements (OLSE) [5]. The multi-scale morphological sifting (MMS) utilize morphological filters with oriented linear structuring elements to extract mass-like patterns and regions from mammographic images. The sifting process is applied on different scales to accommodate different breast mass sizes.

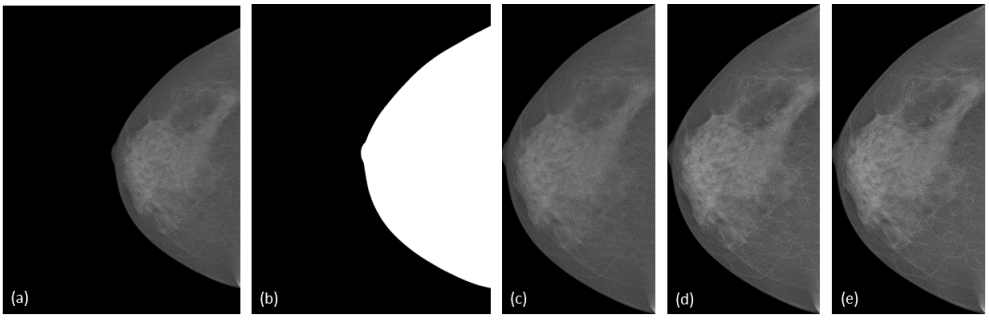

Figure 1: A visual representation for the pre-processing steps. (a) Original mammogram image. (b) Thesholded mask for the original image. (c) Segmented image without redundant black background. (d) CLAHE applied on segmented region. (d) Final pre-processed image with a subsampled size 1/4 of its original size.

Table 1: MMS parameters for the proposed work.

$[Area_{\mathbf{min}}, Area_{\mathbf{max}}]$	P	M	N
[15,3689]	70 µm	10	18

We implemented a similar approach to the multi-scale morphological by [5], where it enables extracting masses of different sizes that fall within a size range of masses $[Area_{\min}, Area_{\max}]$, pixel size P, number of scales M, N elements in each set of structuring elements, and orientation range is $[0°, 180°]$, where the orientation difference was set to $\Delta\theta = 180/N$. Table 1 summarizes the parameter values used in our implementation.

One important parameter is the number of scales that were modified on different iterations to obtain the best evaluation results. As 10 scales were used on every image for extracting the mass region, this increased the processing time compared to using 4 scales as proposed in the literature.

As this is a multi-scale approach with multiple structuring elements used on different orientations, we used the generated kernels and applied top-hat transform and morphological opening operations on the input processed. The output scale result is then scaled back to the full range of a 16-bit image. Thus, this stage generates 10 scales for every single image, each scale is the result of a morphological operation with a different kernel size and orientation to handle the variety of mass sizes. Figure 2 shows the output of MMS from five different scales after the pre-processing step was executed.

2.4 Mean Shift Filtering

In this stage, mean shift filtering technique was applied on every scale of the MMS result. Mean shift is a non-parametric technique [4], where the filtering stage of mean shift segmentation can be used to smooth images while preserving important structures and edges. The advantage of using the mean shift algorithm for filtering is that edges and other continuities are maintained; pixels across the discontinuity are excluded when they are smoothed with a weighted average of the neighbors in both space and color range [1]. We used *spatial window radius* = 15 and *color window radius* = 60 in our implementation after converting the input scale image to 3-channel image. The filtered image is converted back to a 1-channel grayscale 16-bit image. The resulting image enhances the region intensity and makes it more distinct for further processing. Figure 2 shows the output of mean shift filtering approach taking an input image from five different scales after MMS step was executed.

2.5 K-Means Clustering

K-mean clustering algorithm is a simple and fast unsupervised algorithm that can be used to segment the region of interest from the surrounding background, which is done by classifying a given set of data into k number of disjoint cluster [2]. In our implementation, K-means clustering algorithm with $k = 2$ parameter was used to segment the mean shift filtered images into two clusters. We defined $k = 2$ as in at least one of the 10 scales, the mass will have brighter intensity after the filtering making the segmentation result to be mass region in white (foreground) and every other region in black (background), thus, it behaves like a binarization process. As a result of the k-means, we obtain the labels and the array of centers of clusters, which are used to replace pixel values with their center value. The final segmented image is then converted to 16-bit image as in the previous steps. Figure 2 shows the output of k-means approach taking an the mean shift filtered images as an input.

2.6 Post-processing

Post-processing plays an important role in refining the results after image segmentation. In this work, a set of morphological operators were applied to refine the resulting mask from the k-means clustering step. Among the 10 scales, it was observed that the first few scales had the highest number of mask regions after the segmentation, which varied in shape and size that could exceed or go below the target mass sizes and shape. Thus, in this final stage, we applied adaptive thresholding as the

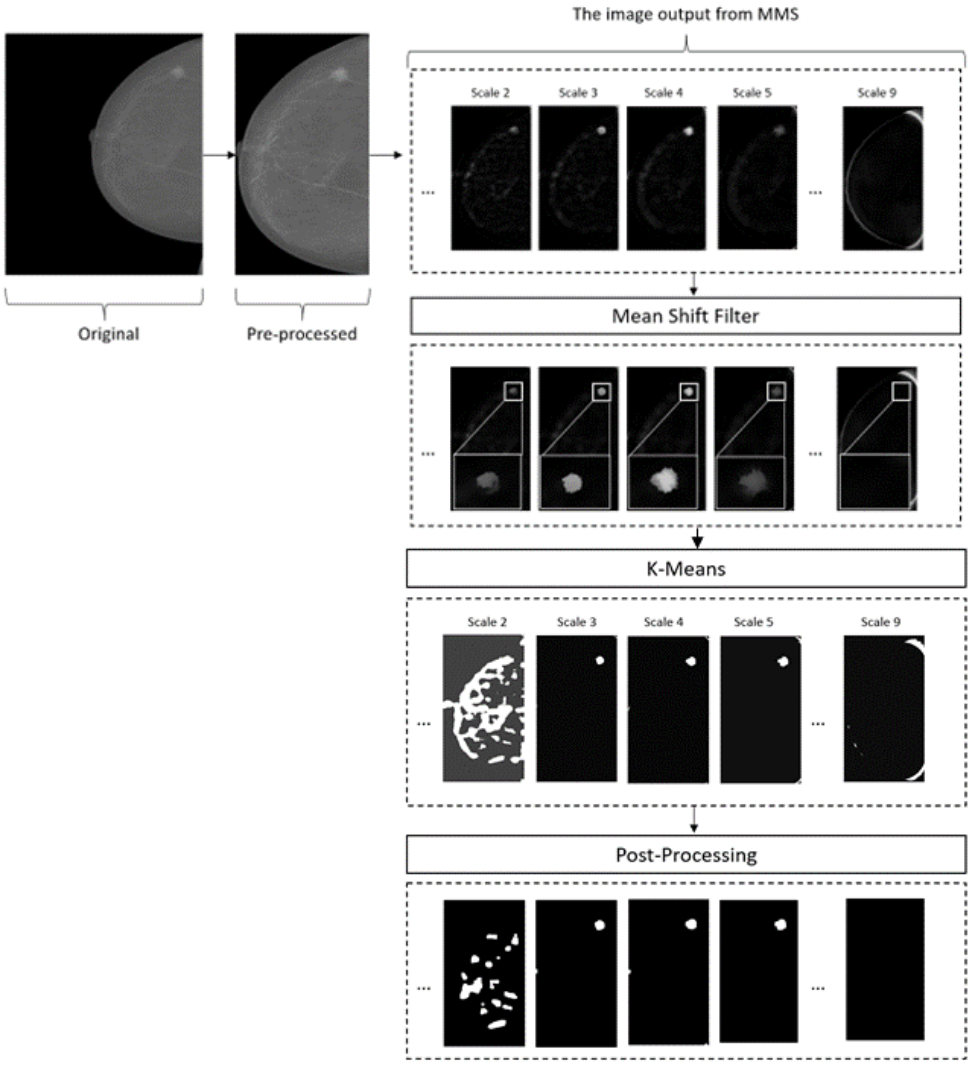

Figure 2: A visual representation of all of the methods with their outputs.

threshold value changes dynamically over the image [8], then the contours of the thresholded image were obtained and contours with areas outside this area range $[110, 551132]$ were eliminated. Finally, thin lines were eliminated using a square kernel $15x15$ kernel was used to perform opening operation on the binary image. This is to make sure that any breast borders were eliminated from the final mask. Figure 2 shows 5 different output scales for a single image after the post-processing steps.

2.7 Feature Extraction

After generating the region candidates, texture features such as grey-level co-occurrence matrix (GLCM) as well as Local Binary Patterns (LBP) features were extracted from the region bounding box, shape features and intensity features were extracted from the candidate region, shape features. A full list of all extracted features can be found in table 2. All of the features has been extracted from the cropped raw mammograms within each region candidate. Pixel intensity mean and standard deviation have been extracted as intensity features. Shape features has been extracted from the shape properties of the region candidates such as area, perimeter, aspect ratio, and compactness.

The GLCM features has been calculated in one distance and four degree angles (0, 45, 90, 135) where the contrast, dissimilarity, homogeneity, energy, correlation, and ASM has been calculated on each angle. Shape features such as area, perimeter, aspect ratio and compactness has been extracted from the region candidate. LBP uniform features were extracted in this work using two different parameter values for the both circularly symmetric neighbour set points and the radius of the neighborhood around each pixel. To calculate both GLCM and LBP features, the bounding rectangle of the region candidate was calculated and then was used to extract a bounding box that surrounds both the the region candidate and its surrounding region.

Table 2: Extracted features in the proposed work.

Intensity features	**Texture features**	**Shape features**
Mean	GLCM related features (6)	Area
Standard deviation	LBP patterns at 8:1 and 16:2	Perimeter
		Aspect ratio
		Compactness

From the feature extraction, 92 features were extracted from both positive and negative candidates on the cropped raw images. All of those features were separated into two files, and normalised using Min-Max normalisation approach where each feature column is normalised individually.

As features were extracted from both positive labelled mammograms that has a ground truth, and negative labelled mammograms. The extraction process which is region-wise and not pixel-wise, for the positive features starts by iterating on all of the ground truth images, up-sampling and binarizing them, then extracting their contours for every image iteration. Once the contours are extracted, an inner iteration starts for

iterating on all of the ground truth contours, where every contour candidate is separated in this step in a separate mask. Each of these ground truth candidates were compared to each and every candidate from every scale. Thus, a third inner loop is made for all of the extracted 10 scales results from the region candidate generation step. For scale, we upsample, binarize, then extract the contours, and finally iterate on those region candidate contours separately. Thus in this stage, we are able to reach each and every ground truth candidate and compare it to each and every segmented contour candidate.

The intersection and union were calculated on every segmented mask candidate and ground truth candidate, where the non-zero pixels were counted for both intersection and union, then the dice similarity index was computed using Eq.1.

$$DSI(I, U) = \frac{2I}{I + U} \tag{1}$$

As we iterate on all of the scales for every image, the maximum dice score of all scales is tracked for a later stage of sensitivity evaluation. The condition for extracting features that indicates a mass is based on the dice score for every scale is shown in Eq.2.

$$Extract\ if\ DSI(I, U) = \frac{2I}{I + U} >= 0.6 \tag{2}$$

The extraction procedure is applied on the negative images, however, as they don't contain any ground truth, we iterate directly on the images, then on all of the scales masks, then on all of the contours candidates, and extract the features with no condition as in the previous two equations. This extraction procedure is also explained in the algorithm (see algorithm 1).

Algorithm 1 Sensitivity Evaluation and Positive Feature Extraction

$DICE\_DATASET\_VECTOR \leftarrow \{\}$
for Every ground truth image **do**
    Read the ground truth image, up-sample, and binarize
    Get the contours of the ground truth
    Read the cropped raw positive image and convert to 8-bit
    $DICE\_IMAGE\_VECTOR \leftarrow \{\}$
    for All the ground truth contours **do**
        $gt\_candidate \leftarrow current\ contour$
        $MAX\_DICE \leftarrow 0$
        for Every Scale of the 10 Scales **do**
            Read the segmentation for the current scale for the current image
            Upsample and binarize the segmentation image
            Get the contours of the segmentation results
            for All the segmentation contours on current scale and image **do**
                $mask\_candidate \leftarrow current\ contour$
                $intersection \leftarrow$ cv2.bitwise_and($mask\_candidate, gt\_candidate$)
                $union \leftarrow$ cv2.bitwise_or($mask\_candidate, gt\_candidate$)
                $int\_pixels \leftarrow$ cv2.countNonZero($intersection$)
                $union\_pixels \leftarrow$ cv2.countNonZero($union$)
                Calculate dice score
                if $Current\ dice\ score >$ MAX_DICE **then**
                    $MAX\_DICE \leftarrow$ current dice
                end if
                Extract all features from the cropped raw and the $mask\_candidate$
                if $Current\ dice\ score >= 0.6$ **then**
                    Save current candidate feature row
                end if
            end for
        end for
        $DICE\_IMAGE\_VECTOR \leftarrow$ append(MAX_DICE)
    end for
    $DICE\_DATA\_VECTOR \leftarrow$ append($DICE\_IMAGE\_VECTOR$)
end for

2.8 Model training using ucasML tool

After the feature extraction, the ratio between positive features to negatives is approximately 1:47 generated for training and evaluation, mak-

ing a total of 240 positive feature candidates and 11,419 negative feature candidates. ucasML tool was used with 10-fold-cross validation as it is robust to such class imbalance. The tool was provided the two features files, where the positives file holds the features of the extracted candidates lesions that match with the true lesion condition only from the positive mammograms, and negatives file that holds the negatives features for the extracted candidates from the negative mammograms with no mass. This tool automatically trains and provides the classification score associated to all positive and negative samples provided. Multiple training iterations were made to find the optimal classification scores that were used for evaluation in the next sections.

3 Experimental and Evaluation Methods

Our system was evaluated on the INBreast dataset. The evaluation was divided into two parts, first, the segmentation sensitivity was determined by evaluating the ground truth masks with the segmentation masks; second, the final model prediction scores were used to plot the Free-Response Receiver Operating Characteristic (FROC) curve.

The segmentation sensitivity was evaluated on the positive images as described in the algorithm (see algorithm 1). The final threshold for the dice scores obtained to indicate if the score for every image counts as $TruePositive$ or $FalseNegative$ was set to be the same as the threshold for extracting the feature as a positive candidate feature, greater than or equal to 0.6.

4 Results and Discussion

In this work, a mammogram mass segmentation method using multi-scale morphological sifting combined with mean shift filtering and k-means clustering was proposed to segment mass lesions with different sizes and irregular shapes from the background. Figure 3 demonstrates a sample result of mass segmentation on 7 different images evaluated from different scales, with different mass sizes. As shown in the figure, some mammogram images contain more than one mass, with a wide variety of sizes and shapes. The red contours represent the candidate segmentation contour from our approach, while the black refers to the ground truth.

By evaluating our implementation with the given ground truth, we obtained 76.92% sensitivity with a dice score of 0.6 as the threshold. This sensitivity reflects the sufficient ability of the segmentation algorithm to

correctly identify true positive instances of masses. This value could be significantly improved by lowering the threshold value below 0.6.

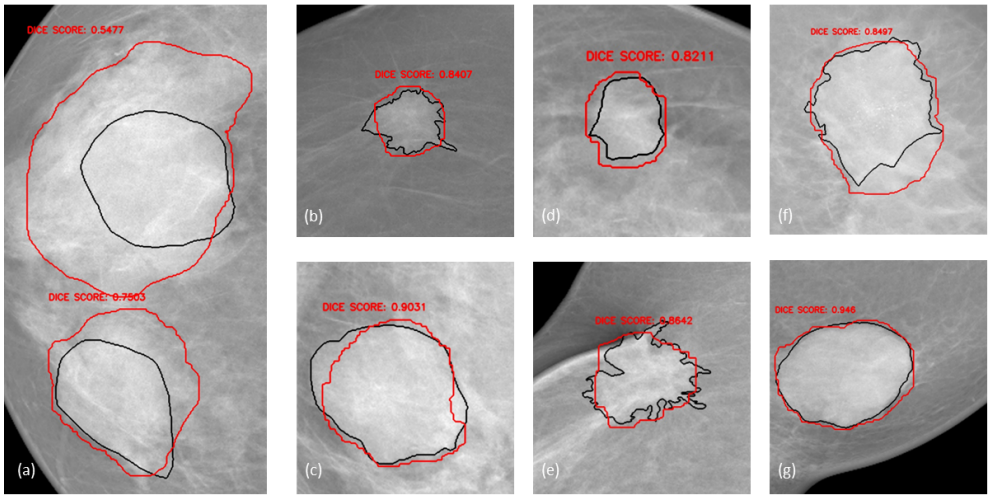

Figure 3: Segmentation results from different scale examples compared to ground truth for various mass shapes, positions, and sizes. The black lines represent the ground truth, and the red lines represent our segmentation results and the corresponding candidate dice score. (a, g) Result from scale 7. (b,d) Result from scale 3. (c) Result from scale 6. (e, f) Results from scale 5.

The results for the mass detection using the ucasML tool score results were evaluated using the FROC curve as in figure 4. By altering the detection threshold to be below 0.6 as mentioned in the previous step, it is guaranteed to improve the curve as the number of positive candidates with slightly increase.

One possible improvement that could be made is to apply k-mean clustering on the image superpixels instead of the pixels by integrating Simple Linear Iterative Clustering (SLIC) for superpixel generation. This approach was tested on a separate iteration on the project but was not finalised, however, it could result in a better segmentation.

5 Conclusion

In this work, we have presented a mammographic mass detection and segmentation approach using a multi-scale morphological sifting approach integrated with a mean shift filter, k-means, and post-processing that detects and segments breast masses. This approach was implemented and evaluated on the InBreast mammographic dataset, which was able

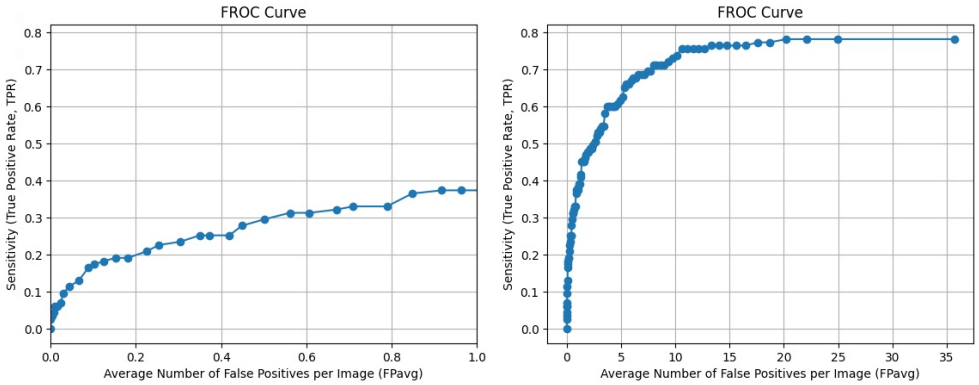

Figure 4: FROC curve obtained with the proposed work

to segment mass lesions from the background accurately with a sensitivity of 76.92%. Various features were extracted from raw images, and the ucasML tool with 10-fold-cross validation was used for a binary classification task as it is robust to class imbalance. ucasML tool was able to provide satisfactory scores for both labels based on the features trained with while facing a high class imbalance. In general, the proposed system achieves promising results in both detection and segmentation of mammogram masses.

References

[1] Damir Demirović. An implementation of the mean shift algorithm. *Image Processing On Line*, 9:251–268, 2019.

[2] Nameirakpam Dhanachandra, Khumanthem Manglem, and Yambem Jina Chanu. Image segmentation using k -means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, 54:764–771, 2015. Eleventh International Conference on Communication Networks, ICCN 2015, August 21-23, 2015, Bangalore, India Eleventh International Conference on Data Mining and Warehousing, ICDMW 2015, August 21-23, 2015, Bangalore, India Eleventh International Conference on Image and Signal Processing, ICISP 2015, August 21-23, 2015, Bangalore, India.

[3] K Ganesan, UR Acharya, CK Chua, LC Min, KT Abraham, and KH Ng. Computer-aided breast cancer detection using mammograms: a review. *IEEE Rev Biomed Eng*, 6:77–98, 2013.

[4] Thomas Melzer. Non-parametric segmentation of als point clouds using mean shift. 1(3):159–170, 2007.

[5] Hang Min, Shekhar S Chandra, Stuart Crozier, and Andrew P Bradley. Multi-scale sifting for mammographic mass detection and segmentation. *Biomedical Physics & Engineering Express*, 5(2):025022, 2019.

[6] Hang Min, Devin Wilson, Yinhuang Huang, Siyu Liu, Stuart Crozier, Andrew P Bradley, and Shekhar S. Chandra. Fully Automatic Computer-aided Mass Detection and Segmentation via Pseudo-color Mammograms and Mask R-CNN. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 1111–1115, 2020.

[7] IC Moreira, I Amaral, I Domingues, A Cardoso, MJ Cardoso, and JS Cardoso. INbreast: toward a full-field digital mammographic database. *Acad Radiol*, 19:236–48, Feb 2012.

[8] Payel Roy, Saurab Dutta, Nilanjan Dey, Goutami Dey, Sayan Chakraborty, and Ruben Ray. Adaptive thresholding: A comparative study. In *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, pages 1182–1186, 2014.

[9] Hyuna Sung, Jacques Ferlay, Rebecca L. Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global

Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3):209–249, feb 2021.