



## A Comparative Study of Breast Mass Detection Using YOLOv8 Deep Learning Model in Various Data Scenarios on Multi-View Digital Mammograms

Muhammet Üsâme ÖZİÇ<sup>1\*</sup>, Ayşe Sidenur YILMAZ<sup>1</sup>, Halil İbrahim SANDIRAZ<sup>1</sup>  
Baihaqi Hilmi ESTANTO<sup>1</sup>



<sup>1</sup>Department of Biomedical Engineering, Faculty of Technology, Denizli, Türkiye

(ORCID: [0000-0002-3037-2687](https://orcid.org/0000-0002-3037-2687)) (ORCID: [0009-0005-9014-1038](https://orcid.org/0009-0005-9014-1038)) (ORCID: [0009-0002-0904-5806](https://orcid.org/0009-0002-0904-5806))

(ORCID: [0009-0005-8468-2970](https://orcid.org/0009-0005-8468-2970))

**Keywords:** Breast Cancer, Deep Learning, Mass, VinDr-Mammo, YOLOv8

### Abstract

Breast cancer is one of the most common lethal cancer types in the female population globally. It typically begins with abnormal cell growth in the breast glands or milk ducts and can spread to other tissues. Many breast cancer cases start with the presence of a mass and should be carefully examined. Masses can be monitored using X-ray-based digital mammography images, including left mediolateral oblique, right craniocaudal, left craniocaudal, and right mediolateral oblique views. In this study, automatic mass detection and localization were performed on mammography images taken from the VinDr-Mammo full-field digital mammography dataset using the YOLOv8 deep learning model. Three different scenarios were tested: raw data, data with pre-processing to crop breast regions, and data with only mass regions cropped to a 1.2x ratio. The data were divided into 80% for training and 10% each for validation and testing. The performance results were calculated using metrics such as precision, recall, F1-score, mAP, and training graphs. At the end of the study, it is demonstrated that the YOLOv8 deep learning model provides successful results in mass detection and localization, indicating its potential use as a computer-based decision support system.

### 1. Introduction

Today, breast cancer is considered one of the most common cancers that threaten the lives of women. Breast cancer occurs as a result of uncontrolled growth of breast tissue. Approximately 685,000 women died due to breast cancer in 2020. These statistics account for 16% or 1 in every 6 cancer deaths in women. Therefore, research and development to detect breast cancer early with correct diagnosis and treatment is extremely important to enhance the survival rate. Many cases of breast cancer begin with the presence of a mass in the breast, and when a patient presents with suspicion of cancer, various methods are carefully employed to examine whether a mass exists. Masses can be benign, non-cancerous, or malignant, indicating a potential for cancer. Usually, breast cancer manifests itself with

the formation of a mass or masses and can be detected in the early stages using different imaging methods [1-3]. Mammography is a commonly used X-ray-based imaging technique for screening and diagnosing breast masses. In general, images are taken from four different angles and views: Left Mediolateral Oblique (LMLO), Right Mediolateral Oblique (RMLO), Right Craniocaudal (RCC), and Left Craniocaudal (LCC) [4, 5]. Radiologists analyze these four views to analyze cancerous and suspicious masses. In recent years, the use of deep learning methods in breast cancer screening and diagnosis using mammography images has made great progress. Many studies have shown that artificial intelligence systems can reduce the workload of screening mammograms. In the literature, artificial intelligence-based mass detection and breast cancer classification studies are still hot topics on which applications are

\*Corresponding author: [muozic@pau.edu.tr](mailto:muozic@pau.edu.tr)

Received: 21.09.2023, Accepted: 01.12.2023

being carried out [6-19]. The data used in the studies are generally collected from international databases. They are extremely variable regarding image size, resolution, number of data, image views, labels, and breast density differences [20-23]. Therefore, comparisons across studies are entirely relative, but the studies performed show promising results for early-stage diagnosis of breast cancer. VinDr-Mammo dataset is a project with richer content than the current databases, with four-view mammography of each recently released case, high-resolution radiology images, different findings labels, different breast density values, data observed by several radiologists, and other features [24]. You Only Look Once (YOLO) deep learning model among object detection methods has been popularly used in medical and other fields [25, 26]. The algorithm, which is updated periodically with different versions, emerges with its powerful features compared to other object detection algorithms [27-30]. The n, s, m, l, and x models with different parameters and network sizes of the YOLOv8 model, which is one of the latest versions, are available to researchers [31]. In this study, the performance of the YOLOv8 algorithm on the automatic detection and localization of mass regions was investigated by creating datasets in three different scenarios from the mass-labeled RCC, LCC, RMLO, and LMLO views using the VinDr-Mammo [24]. Although the RCC, LCC, RMLO, and LMLO views have different directions and pixel density values, how the algorithm models perform has been investigated. Three different scenarios were prepared for this research. Four views were taken as raw data in the first scenario, and applications were carried out. In the second scenario, a dataset was created by removing the black pixel regions in the background as much as possible with a pre-processing method developed for this study. In the third scenario, a dataset was created by cropping the mass-labeled regions to a 1.2x ratio, resulting in only the mass regions and some surrounding tissues. These scenarios aim to investigate how algorithm performance will perform on raw and Region of Interest (ROI) data. Data were labeled and annotated according to the mass coordinate regions given from the VinDr-Mammo dataset. Performance criteria were calculated by training the data separated as 80% training, 10% validation, and 10% testing with the n model of YOLOv8. The most important contribution of this study, compared to other studies in the literature, is that it examines the breast mass detection performance in classes created in different scenarios in the rich data set mentioned above, without applying complex pre-processing methods to the images (filtering, adaptive histogram equalization, denoising

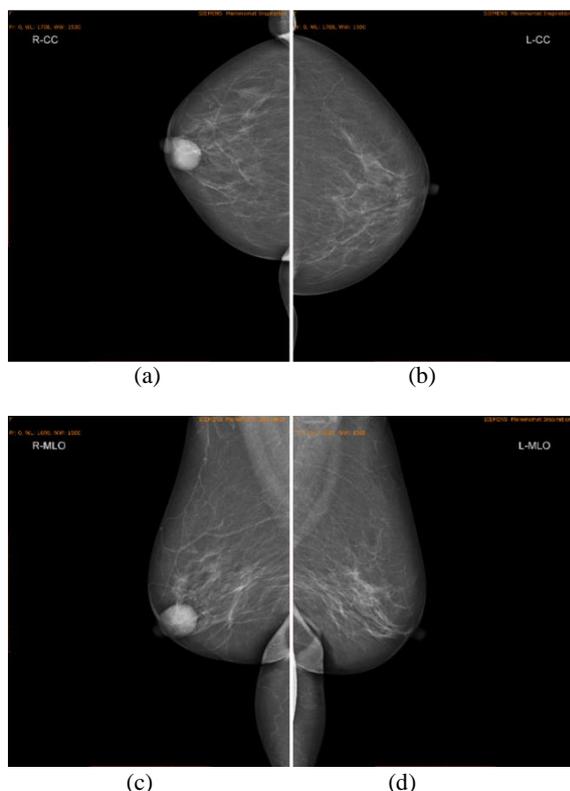
flipping images right or left, histogram equalization, etc.). At the same time, it has been observed how the algorithm performance is affected as the examined region becomes smaller. At the end of the study, the advantages and disadvantages of the experiments of the proposed systems were discussed. The results demonstrate that the YOLOv8 model is suitable for computer-based clinical decision support processes as it can detect masses on mammography images regardless of views and densities.

## 2. Material and Method

### 2.1. Data

The VinDr-Mammo dataset used in this study is an open-source project developed by Vietnam National University and VinBrain AI [24]. Using this data, researchers have developed artificial intelligence-based computer-aided diagnostic systems [32-36]. This dataset contains 20,000 images, 5,000 of which are four-view (full-field) radiographs (LMLO, LCC, RMLO, RCC). Evaluations, findings, and annotation information of the images are also given. Images are labeled as laterality, view information, image size, Breast Imaging Reporting and Data System (BI-RADS) category, composition information and breast density (A, B, C, D), category of findings (skin retraction, suspicious lymph node, skin thickening, asymmetry, focal asymmetry, suspicious calcification, nipple retraction, architectural distortion, global asymmetry, mass). BI-RADS is a classification system developed by the American College of Radiology and used to standardize breast imaging results. BI-RADS 0 to 6 categorizes the severity of the findings. Generally classified as BI-RADS-0 additional imaging or further evaluation, BI-RADS-1 normal, BI-RADS-2 benign, BI-RADS-3 possibly benign, BI-RADS-4 suspicious malignant lesion (between 2% and 95% probability of malignancy), BI-RADS-5 high-grade suspicious malignant lesion, BI-RADS-6 lesion known to be malignant by biopsy [37, 38]. Breast density information is divided into four main categories, A, B, C, and D, which represent the ratio of fat, fibroglandular tissue, and dense tissue in the breast. The breast density, which is a relatively easy-to-detect anomaly and consists of mostly adipose tissue with low density, is classified as category A, and the breast with a balanced structure between adipose tissue and fibroglandular tissue, with medium density, is classified as category B. The increase in density, in which the detection of anomaly becomes relatively more difficult with the increase in the ratio of fibroglandular tissue, is classified as category C, and

the composition consisting of mostly fibroglandular tissue with high density and difficult-to-detect anomaly is classified as category D [39]. In this study, a dataset with different BI-RADS categories and breast density levels labeled with mass was created from the VinDr-Mammo dataset. The mass represents a lesion or abnormality. 1226 findings are labeled as "mass" and the coordinates (Xmin, Ymin, Xmax, Ymax) where the mass is located are given in the \*.CSV file. These findings can be one or more in an image. Therefore, there are 1113 mammograms with 1226 mass findings. Mammograms are recorded in Digital Imaging and Communications in Medicine (DICOM) format. Therefore, the radiographs were converted to high-contrast Portable Network Graphics (PNG) format using a DICOM viewer. The coordinate information given by VinDr-Mammo was converted to YOLO format and checked in the browser-based "makesense" labeling program. These images have dimensions of 2800x3518 and a depth of 8 bits. Mass labeled dataset has different BI-RADS categories, different views, and breast density. In Figure 1, a patient's RCC, RMLO, LCC, and LMLO image is given. While the mass is seen in RCC and RMLO views, there is no finding in LCC and LMLO images. Table 1 shows the distribution label of the data used in the study regarding BI-RADS category, density, and view.



**Figure 1.** (a) RCC (b) LCC (c) RMLO (d) LMLO views of a case

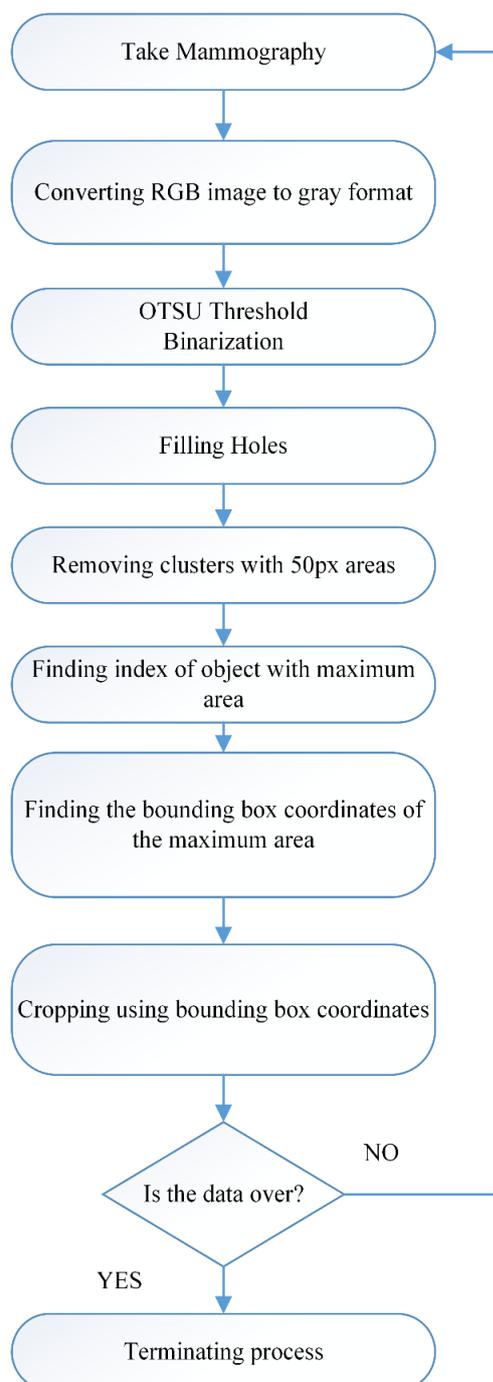
**Table 1.** The distribution of the data used in the study in terms of BI-RADS category, density, and view (According to the distribution of findings )

<b>BI-RADS</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>-</b>
	530	481	215	-
<b>Density</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>
	5	153	998	70
<b>View</b>	<b>LCC</b>	<b>RCC</b>	<b>LMLO</b>	<b>RMLO</b>
	310	267	343	306

## 2.2. Image Pre-processing and ROI Extraction

All mammography images were pre-processed for the second and third scenarios. In the second scenario, a pre-processing method was presented to remove as much of the breast area as possible. Thus, the breast regions in the entire image were automatically cropped. In the third scenario, ROIs were cropped to sizes with dimensions that changed by 1.2 times, according to the bounding box coordinates provided in the VinDrMammo dataset. First of all, high-resolution mammography images with DICOM extension were converted to PNG format for all scenarios. For the second scenario (ROI1), an automated pre-processing pipeline was developed that only takes the breast area, as the mammography images have the text and a lot of black pixel backgrounds. Images were recorded in 3 channels. Therefore, conversion from color image to gray level was performed. Binarization was applied by finding the adaptive threshold value specific to the image with the OTSU thresholding method [40]. The holes in the object were filled and the binary regions with 50-pixel clusters were removed from the image. By finding the pixel areas of independent objects, the index of the object with the highest pixel area was obtained. Since the largest object with this index was the breast region, the smallest bounding box coordinates surrounding this region were saved. The largest object corresponding to these coordinates in the original image was cropped and saved as a new image. In order not to lose the mass region coordinates given in VinDr-Mammo, resize was not performed. This developed pipeline was looped and all images were cropped automatically. The developed pipeline is given in Figure 2 as a flow chart. Figure 3 pre-processing steps are given visually on mammography images. For the third scenario (ROI2), considering the bounding boxes in the VinDrMammo dataset, the relevant region of each image was cropped by a 1.2x ratio (This expression is not the zoom of the image, but how much the relevant coordinate points are expanded when cropping the ROI boundaries). As a

result, a new dataset was created with the truncated ROIs by enlarging the bounding box 1.2 times. The new coordinate values were calculated according to the original image and saved in YOLO format. The goal of the approach here is to calculate how the performance of the algorithm may change as the ROI region becomes smaller. Figure 4 shows the cropping of a mass ROI.



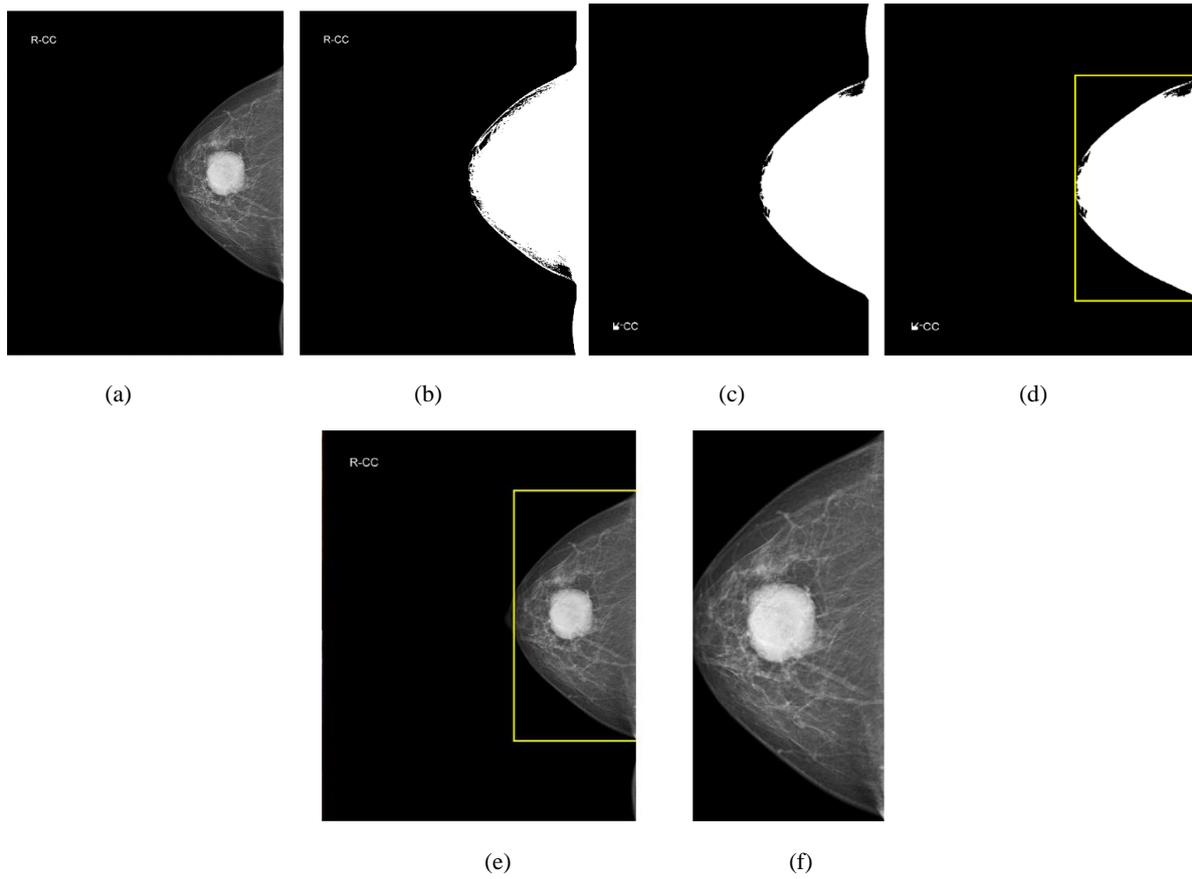
**Figure 2.** Pipeline developed to crop the breast area

### 2.3. YOLOv8 Network Settings

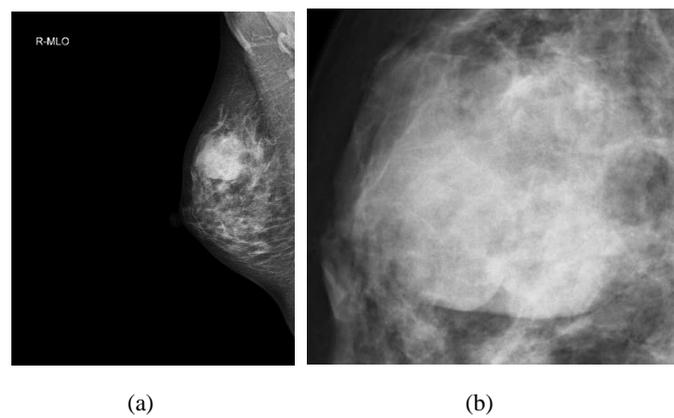
The YOLOv8 deep learning model was preferred for breast mass detection because it is one of the newest versions of the YOLO family. During the training stage, a high-performance computing environment was obtained by purchasing the Google Colab Pro version. The training process was carried out using the YOLOv8 model with the GPU and Tesla T4 graphics processing unit provided by Google Colab Pro. Model hyperparameters were set to epoch 600, optimization algorithm Stochastic Gradient Descent, mini-batch size 16. Other hyperparameters were left as default. The n model of YOLOv8 was downloaded from the Ultralytics repository, and transfer learning and fine-tuning operations were carried out. This model was a pre-trained network with 80 classes of output, normally trained with the COCO dataset. The classes in the "coco.yaml" file were deleted and replaced with the name "mass" to give a single class output. The data for the three scenarios were divided into 80% training, 10% validation and 10% testing. The bounding box in YOLO format and data were uploaded to the COLAB platform and the training process was started.

### 2.4. Performance Evaluation Metrics

Performance evaluation metrics were examined in two stages: training and testing. IoU (Intersection over Union) value is taken into account in the performance criteria during the training stage. This metric is a measure of how much the annotated region overlaps with the bounding box obtained by the algorithm. For IoU, 50% overlap is the critical limit. If the overlap is largely achieved, True Positive (TP), False Positive (FP), and False Negative (FN) values are calculated. Using these values, precision (Equation 1), recall (Equation 2), and F1-score (Equation 3) values of the training phase are obtained. Another metric for the training process is mAP (mean Average Precision). This value is calculated after determining the average precision (AP) value for each class. The AP value is the area under the precision-recall graph of the classes. mAP is the average of the AP values of all classes. The mAP .5 value represents the performance when the IoU is 50%. mAP .5: .95 is the average of the IoU threshold from 50% to 95% by calculating one by one with a step size of 0.05. The higher the AP and mAP values, the more successful the training. By using the weight file obtained as a result of the training, the detection, and localization of the masses were performed in test images. By evaluating these data, TP, FP, and FN values were calculated, and precision (Equation 1), recall



**Figure 3.** (a) Original mammography converted to gray level (b) OTSU Threshold and Binarization (c) Filling holes and removing 50-pixel areas (d) Finding maximum area, indexing and finding bounding box coordinates (e) Cropping using bounding box coordinates (f) Cropped mammography



**Figure 4.** (a) Original Mammography (b) Cropped Mass Region

(Equation 2), F1-score (Equation 3) values and test performance were evaluated.

$$precision = \frac{TP}{TP + FP} \quad (1)$$

$$recall = \frac{TP}{TP + FN} \quad (2)$$

$$F1-score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3)$$

### 3. Results and Discussion

#### 3.1. Training and Testing Results

Training results for three scenarios are given in Table 2, and test results are given in Table 3. Figure 5 shows training graphs for different metrics by epoch progression for the three scenarios. When Table 2 was examined, it was observed that the values of training performance evaluation metrics increased as the evaluated region became smaller. Although 600 epochs of hyperparameter input were given for each scenario, the training was stopped by performing an early stopping operation on the system at the 88th epoch for raw data, the 71st epoch for ROI1, and the 102nd epoch for ROI2. This process was performed automatically by the system when there was no significant change in the last epochs. Thus, it reduced the training time. The pre-processing performed for ROI2 generally increased the performance values.

Since mAP values were above the determined threshold values, it showed that the detection process was successful. It was observed that the graphical progress given in Figure 5 approaches zero for loss graphs and approaches one value for other metrics, indicating that the training process was progressing at appropriate values. The test results given in Table 3 were calculated by comparing the detection operations performed on mammography images with the ground truth labels. The precision value was obtained as 95.7% for raw data, 97.9% for ROI1, and 100% for ROI2. It has been determined that the mass regions detected by the model are largely accurate and that models trained by pre-processing increase the precision results. Recall value was obtained as 78.8% for raw data, 82.4% for ROI1, and 100% for ROI3. This metric shows how much of the mass regions the algorithm needs to detect. Thus, the results proved that as the ROI size decreased, the amount of FN decreased and therefore the recall value increased.

#### 3.2. Visual Results

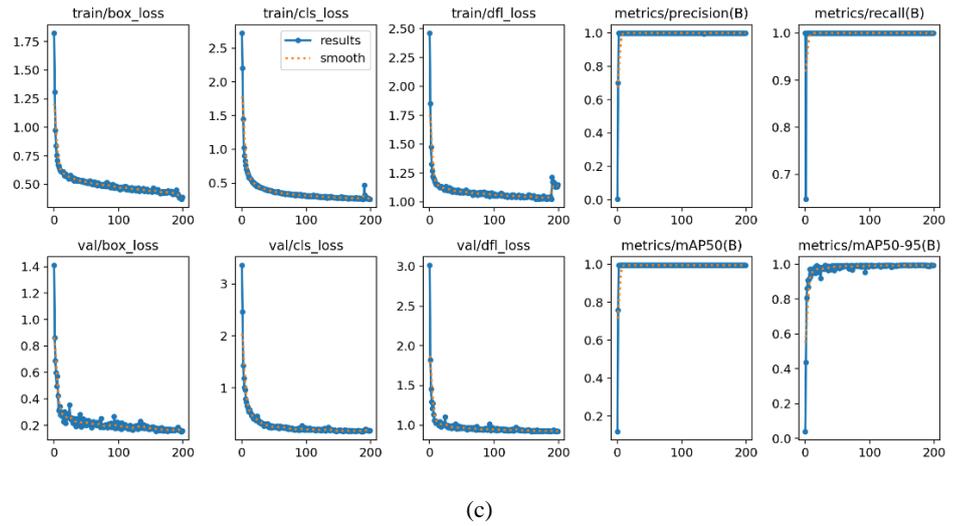
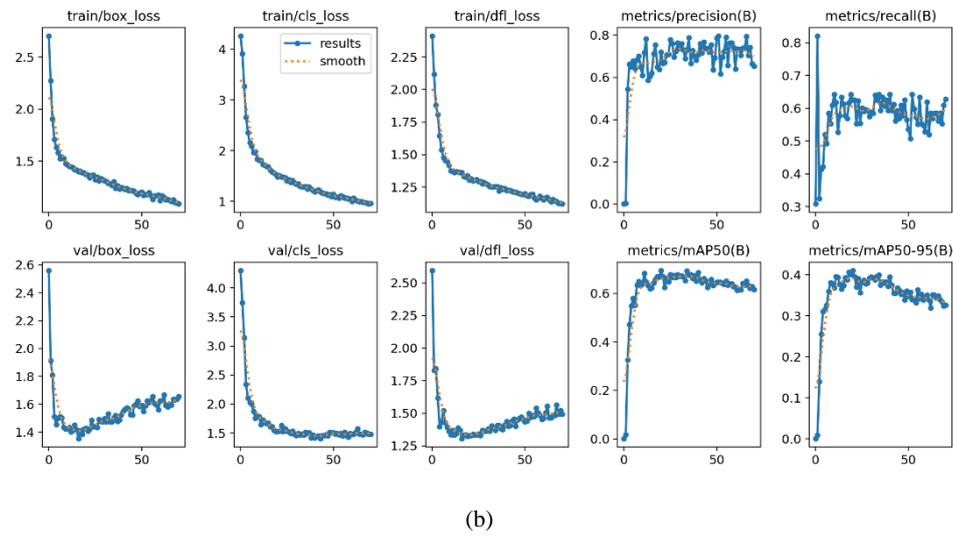
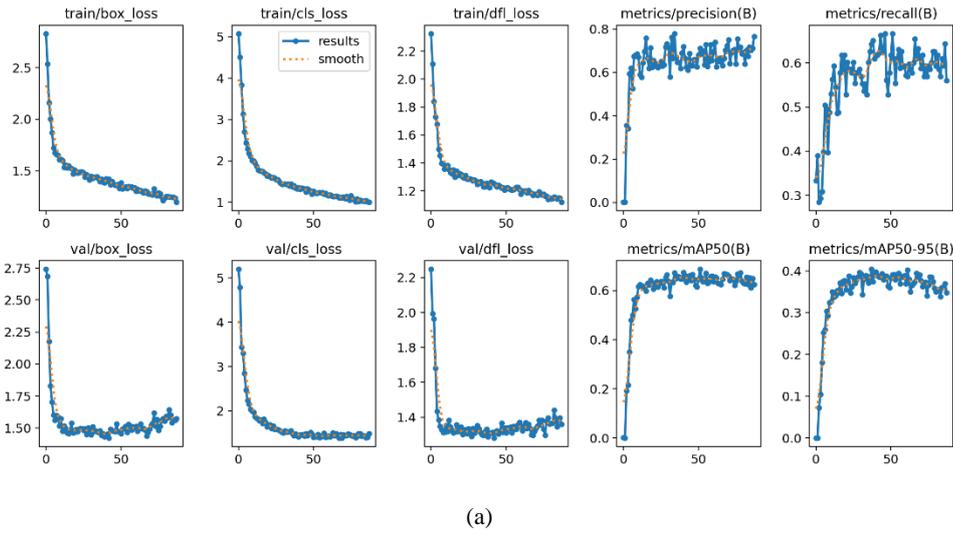
Detection and localization of mass on six test raw mammography images at different views in Figure 6, detection and localization of mass on six test ROI1 (Breast Region) mammography images at different views in Figure 7, detection of masses on eight ROI2 (Mass Region) mammography patches in Figure 8 are given as sample images from the visual results.

**Table 2.** Training Results

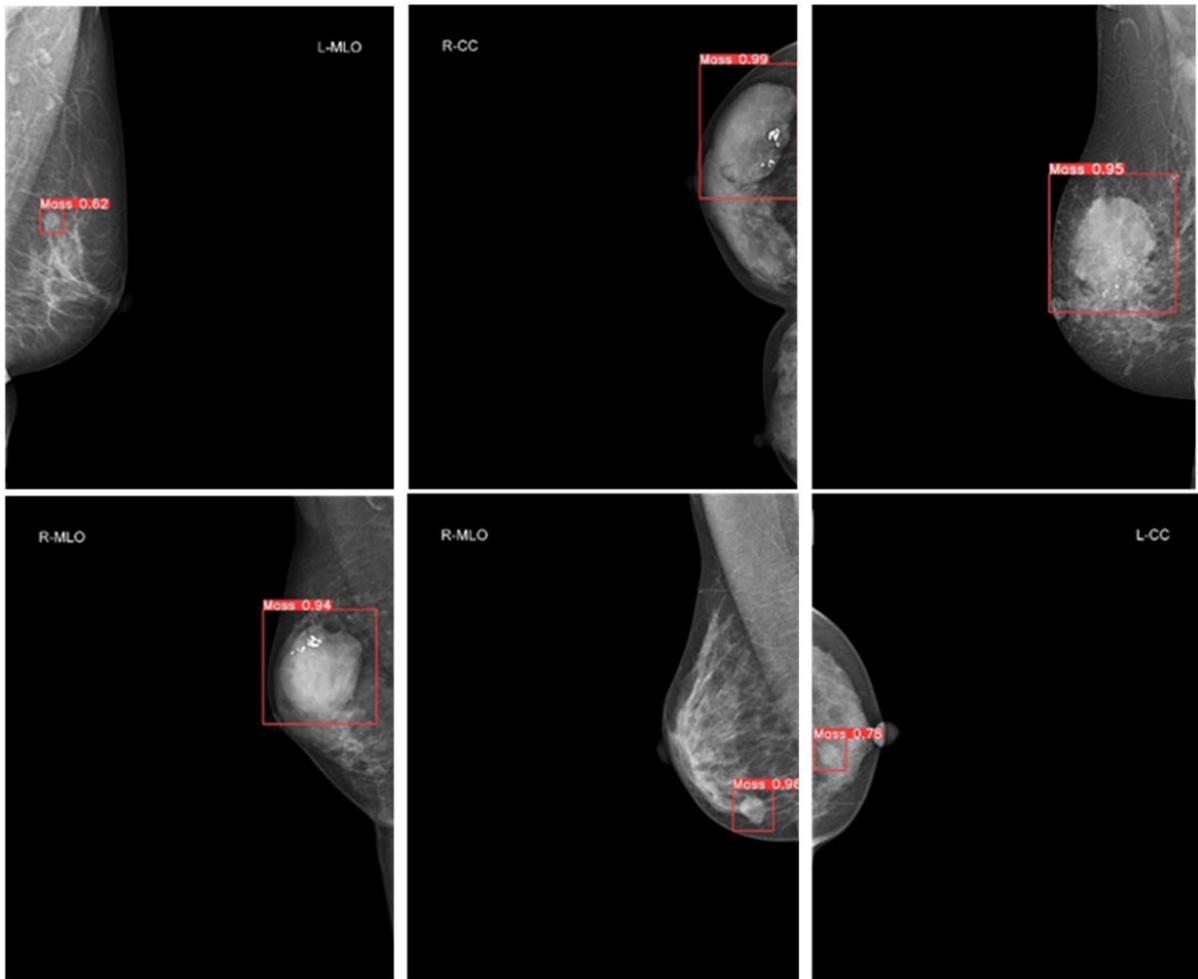
Scenario	Precision	Recall	F1-Score	mAP.5	mAP.5:95	Training Time	Epoch
Raw Data	64.5	63.4	63.9	65.6	40.4	1.72 h	88/600
ROI1(Breast Region)	73.0	62.6	67.4	69.4	41.0	0.74 h	71/600
ROI2(Mass Region)	100	100	100	99.5	99.5	0.62 h	102/600

**Table 3.** Testing Results

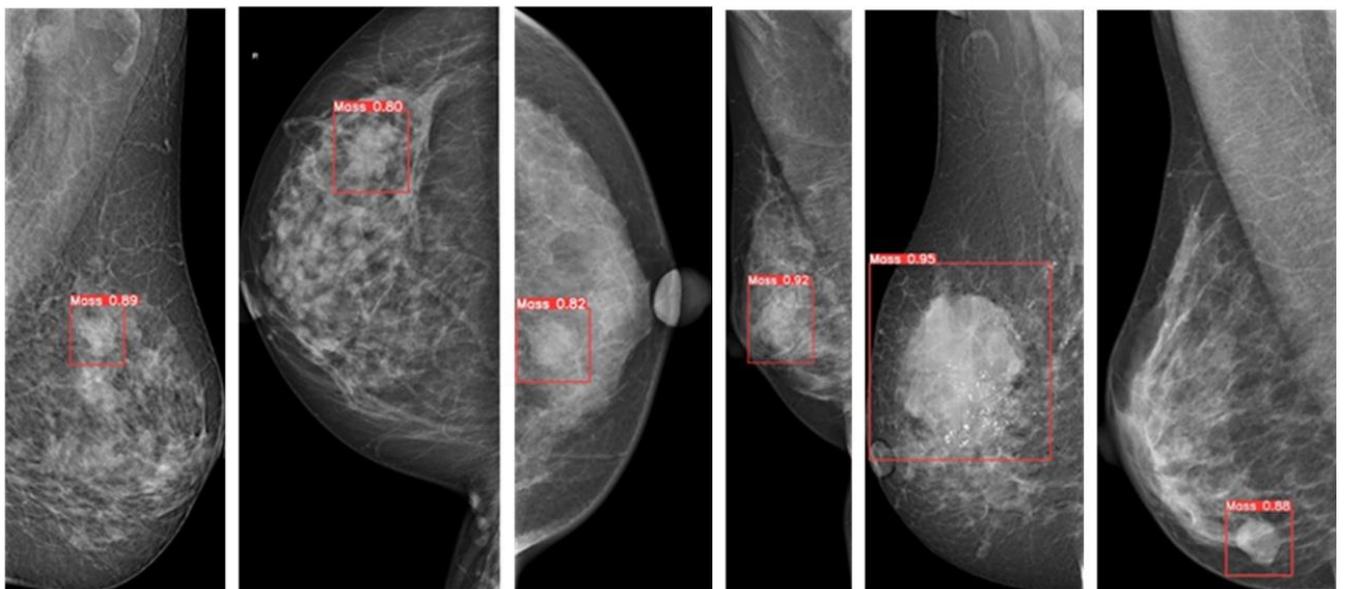
Scenario	TP	FP	FN	Precision	Recall	F1-Score
Raw Data	90	4	24	95.7	78.8	86.4
ROI1(Breast Region)	94	2	20	97.9	82.4	89.4
ROI2(Mass Region)	114	0	0	100.0	100.0	100.0



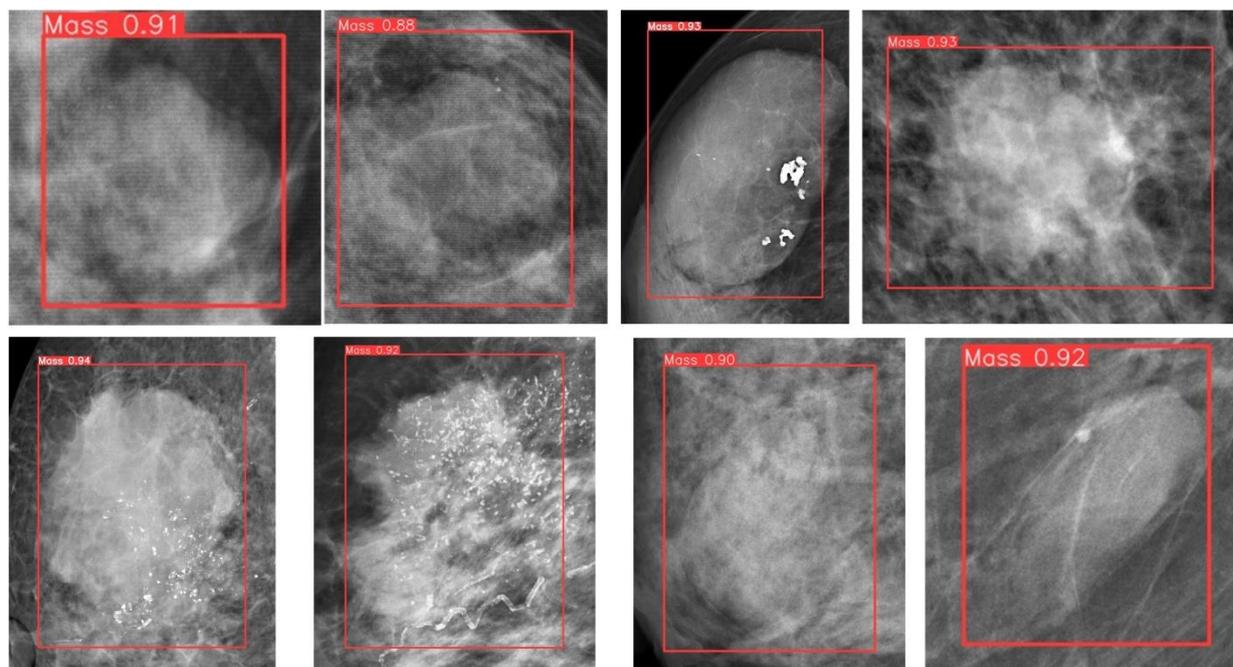
**Figure 5.** Training and validation graphs according to epoch progress (a) Raw data (b) ROI1 data (c) ROI2 data



**Figure 6.** Visual results of mass detection on six raw test images



**Figure 7.** Visual results of mass detection on six ROI1 test images



**Figure 8.** Visual results of mass detection on eight ROI2 test images

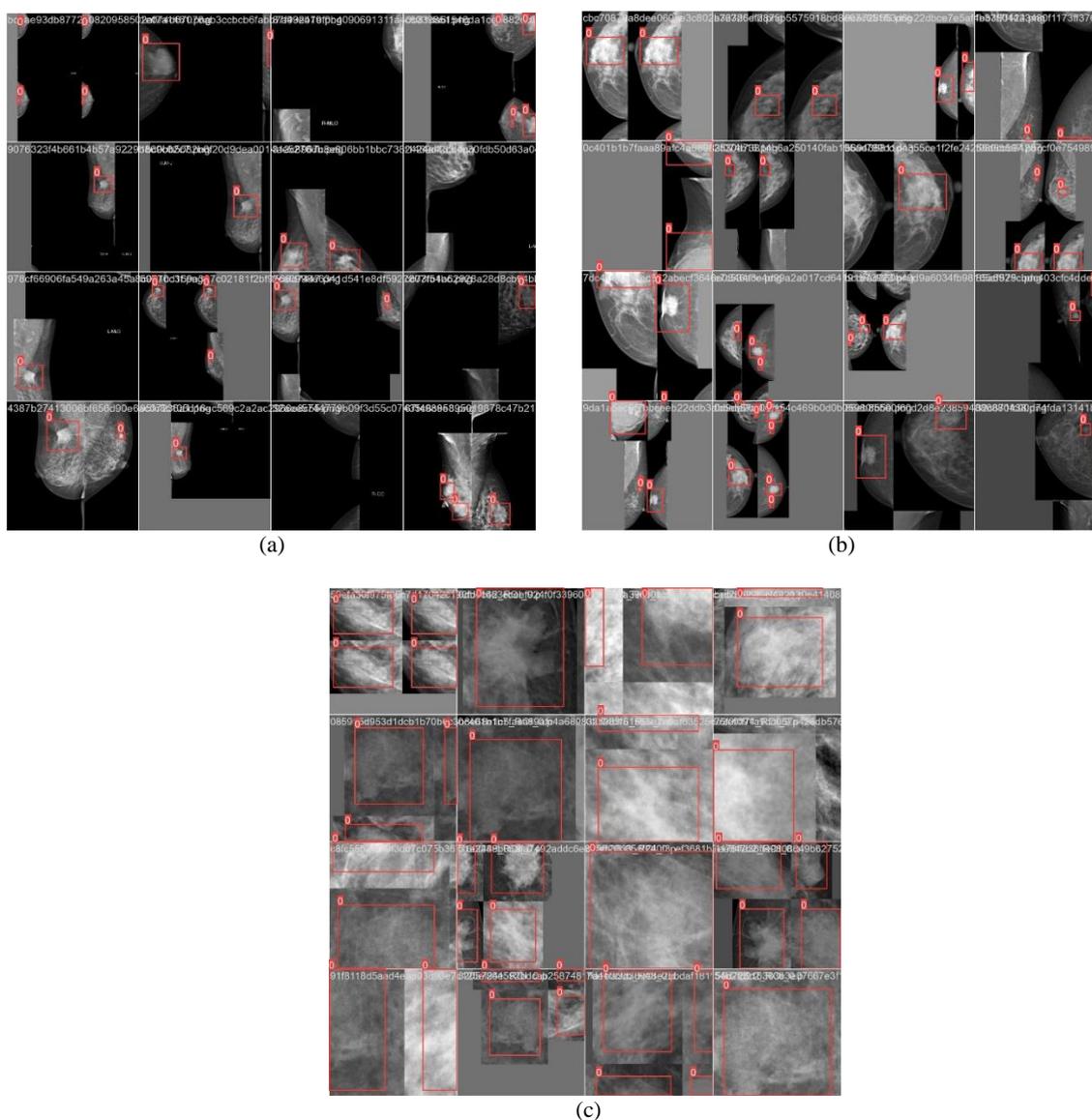
### 3.3. Discussion

Automated mass detection and classification of mammography images, usually using open-source datasets and sometimes privately collected datasets, is a hot topic studied in the literature [6-19]. The Digital Database for Screening Mammography (DDSM) dataset, the Mammographic Image Analysis Society Digital Mammogram Database (MIAS), and the INbreast dataset have been popularly used for years. It is stated that the low-resolution images in MIAS, one of the oldest databases, cause inaccurate training and performance. DDSM, which has low mammography quality and pixel-level contour labels, is not sensitive enough for segmentation and classification processes. Although the INBreast dataset releases high-resolution full-field digital mammography, the number of data is low [20-23]. The VinDr-Mammo dataset is a relatively new study and consists of 5000 cases and contains 20000 high-resolution images, each with four views. Laterality, size, view status, BI-RADS category, and breast density category information of the images are shared with researchers as finding annotations. The data were independently reviewed twice, and in case of discrepancy, they were evaluated by a third radiologist. The VinDrMammo dataset contains the symptoms and label information for skin retraction, suspicious lymph nodes, skin thickening, asymmetry, focal asymmetry, suspicious calcification, nipple retraction, architectural distortion, global asymmetry, and mass. Therefore, it has a richer structure than other databases, as it includes the number of data, the

number of symptoms, finding annotation, and high-resolution images [24]. While many applications are using other databases, the number of studies conducted with VinDrMammo is relatively low [34, 35, 41-44]. Some studies using the VinDr-Mammo dataset applied the YOLO algorithm as a pre-processing tool to crop only the breast region from the image [35, 44]. There is one study that automatically detects the mass area from VinDr mammograms using YOLO [34]. Mahoro and Akhloufi determined the mass regions with YOLOv7 and YOLOv8 with different pre-processing, resize processes, and data augmentation, using 1029 labeled data from the VinDr Mammo dataset. Different pre-processing techniques were performed on mammograms with bilateral filter, median filter, and contrast-limited adaptive histogram equalization (CLAHE). Images resized to 640x640 pixels and converted to JPG format in DICOM were divided into 75% train, 22% validation, and 3% test. Data augmentation was performed during training with different transformations. In experiments conducted with CLAHE, median filter, bilateral filter, and raw data, the mAP values obtained for YOLOv7 were 0.51, 0.53, 0.42, 0.44, and for YOLOv8 were 0.61, 0.65, 0.61, 0.64, respectively (IoU=0.5). In the study, in which no pre-processing was performed to determine the area of interest of the breast region, the results were evaluated only on mAP, and the precision, recall, and F1-score values of the test results were not given. It has been stated that YOLOv8, which is a state-of-the-art model, gives better results in terms of

model architecture, speed, accuracy, and detection of small objects compared to YOLOv7. YOLOv8, one of the latest versions of the YOLO family, has a lot of new features provided to researchers by Ultralytics and offers a user-friendly GitHub repo. The Pytorch-based algorithm uses a dynamic head network strategy to increase its speed and accuracy compared to its competitors. Many studies have shown that YOLOv8 gives better results in terms of performance evaluation metrics than YOLOv7 [31, 45, 46]. In this study, mass areas were automatically detected with the YOLOv8 algorithm in three different scenarios using high-resolution mammography images. DICOM images were saved in PNG format to achieve high contrast. No filters or resizing were applied to

mammograms, thus reducing the computational cost. Using the same set of hyperparameters, the experiments produced the following precision, recall, and F1-score values for the test data: 95.7%, 78.8%, and 86.4% for raw data, 97.9%, 82.4%, and 89.4% for ROI1, and 100%, 100%, and 100% for ROI3. One of the most significant advantages of this study was the ability to capture high-contrast images from mammography without applying any complex filters. The second advantage was the examination of the impact of the detection process when experiments were conducted with ROIs. As seen in Table 2 and Table 3, the success rates of performance metrics increased as processing was performed with ROIs compared to raw data.



**Figure 9.** (a) Mosaic data augmentation for raw data (b) Mosaic data augmentation for ROI1 (c) Mosaic data augmentation for ROI2

In Figure 5, the training conducted with ROIs was shown to be more successful. While performance criteria may vary, as demonstrated in Figure 6, Figure 7, and Figure 8, the detection and localization of mass regions can be accurately determined. It has been determined that the performance evaluation results have increased with the developed automatic background subtraction pre-processing method. 100% success was achieved in the detection processes with ROI2 regions that were cropped by 1.2x. The third advantage was that, as seen in the data distribution provided in Table 1, high detection results were achieved despite the complex distribution of factors such as different BI-RADS, breast density, and views. One of the notable findings was that as the background, which did not contain meaningful information in the image, was removed, the accuracy results increased, demonstrating that higher results can be achieved with patch-based approaches. This may be due to the automatic mosaic data augmentation process implemented during the training of the YOLOv8 algorithm. This process was automatically carried out at the algorithm's input. During the training process, YOLOv8 was provided with images in mini-batches of 16. The training batch was designed as a mosaic through various scaling, resizing, and cropping operations. When examining the mosaic form in Figure 9a, it was evident that there were many black regions in the raw image, which negatively affected the algorithm's performance. As shown in Figure 9b, even with pre-processing, there were relatively black areas left due to the structure of the breast, which did not entirely cover the bounding box of the breast as a whole. In Figure 9c, the mosaic data augmentation form of the patch-based approach with 16 batches is demonstrated. Here, having only breast tissue and no black pixel background regions has positively influenced the performance metrics. As a result of such training, a precision and recall value of 100% was achieved on test data, and the algorithm detected mass in all patches. Another significant finding highlighted by this study was how patch-based approaches significantly improve results and the necessity of breaking down the image into as small patches as possible without background for detection processes. The limitation of this study is the irregularity in the distributions in the BI-RADS, density, and view categories in the data set. Having an equal number of examples in these categories can increase the results of performance evaluation

## References

- [1] M. Arnold *et al.*, "Current and future burden of breast cancer: Global statistics for 2020 and 2040," *The Breast*, vol. 66, pp. 15-23, 2022.

metrics. However, since the VinDr-Mammo dataset was used as a benchmark, this study was performed according to the procedures shared in the database.

## 4. Conclusion and Suggestions

In this study, breast mass detection was performed in different scenarios using the VinDr-Mammo full-field digital mammography dataset and a deep-learning model. The results demonstrate that the deep learning model provides successful results in the detection and localization of masses in breast cancer. In future studies, applications for the detection of other anomalies in mammograms will be explored. AI methods have gained popularity in recent years for the detection, localization, grading, segmentation, and classification of anomalies in medical images. Therefore, the automated analysis of anomalies in medical images, presenting preliminary findings to doctors, is expected to be advanced software used in the future. The applications conducted in this study, which automatically identify and localize breast masses related to breast cancer, have the potential to be used as a computer-based decision support system, aiding doctors in the diagnostic process.

## Acknowledgment

This study was supported by the Scientific Research Coordination Unit of Pamukkale University under project number 2023LÖKAP007.

## Contributions of the authors

ASY and MÜÖ contributed to the organization of the data and the pre-processing. MÜÖ carried out the study and algorithm design. HİS and HBE wrote the code and calculated the performance criteria. All authors contributed equally to article writing and literature review.

## Conflict of Interest Statement

There is no conflict of interest between the authors.

## Statement of Research and Publication Ethics

The study complies with research and publication ethics

- [2] L. Wilkinson and T. Gathani, "Understanding breast cancer as a global health concern," *The British Journal of Radiology*, vol. 95, no. 1130, p. 20211033, 2022.
- [3] R. Ali, A. Sultan, R. Ishrat, S. Haque, N. J. Khan, and M. A. Prieto, "Identification of New Key Genes and Their Association with Breast Cancer Occurrence and Poor Survival Using In Silico and In Vitro Methods," *Biomedicines*, vol. 11, no. 5, p. 1271, 2023.
- [4] M. Buda *et al.*, "Detection of masses and architectural distortions in digital breast tomosynthesis: a publicly available dataset of 5,060 patients and a deep learning model," *arXiv preprint arXiv:2011.07995*, 2020.
- [5] V. Barros *et al.*, "Virtual biopsy by using artificial intelligence-based multimodal modeling of binational mammography data," *Radiology*, vol. 306, no. 3, p. e220027, 2022.
- [6] Y. Yan, P.-H. Conze, G. Quellec, M. Lamard, B. Cochener, and G. Coatrieux, "Two-stage multi-scale breast mass segmentation for full mammogram analysis without user intervention," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 2, pp. 746-757, 2021.
- [7] I. Domingues and J. Cardoso, "Mass detection on mammogram images: a first assessment of deep learning techniques," in *19th Portuguese Conference on Pattern Recognition (RECPAD)*, 2013.
- [8] R. Gayathri and K. Kasirajan, "Yolo Algorithm Based Breast Masses Detection And Classification Technique For Digital Mammograms," *Latin American Journal of Pharmacy*, vol. 42, no. 3, pp. 374-381, 2023.
- [9] A. Mohiyuddin *et al.*, "Breast tumor detection and classification in mammogram images using modified YOLOv5 network," *Computational and mathematical methods in medicine*, vol. 2022, pp. 1-16, 2022.
- [10] Z. Cao *et al.*, "Deep learning based mass detection in mammograms," in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2019: IEEE, pp. 1-5.
- [11] M. A. Al-masni *et al.*, "Detection and classification of the breast abnormalities in digital mammograms via regional convolutional neural network," in *2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, 2017: IEEE, pp. 1230-1233.
- [12] G. Ayana *et al.*, "Vision-Transformer-Based Transfer Learning for Mammogram Classification," *Diagnostics*, vol. 13, no. 2, p. 178, 2023.
- [13] G. H. Aly, M. Marey, S. A. El-Sayed, and M. F. Tolba, "YOLO based breast masses detection and classification in full-field digital mammograms," *Computer methods and programs in biomedicine*, vol. 200, p. 105823, 2021.
- [14] Y. Cui, Y. Li, D. Xing, T. Bai, J. Dong, and J. Zhu, "Improving the prediction of benign or malignant breast masses using a combination of image biomarkers and clinical parameters," *Frontiers in Oncology*, vol. 11, p. 629321, 2021.
- [15] H. Zhang, Z. Xu, D. Yao, S. Zhang, J. Chen, and T. Lukasiewicz, "Multi-Head Feature Pyramid Networks for Breast Mass Detection," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023: IEEE, pp. 1-5.
- [16] Y. Yan, P.-H. Conze, M. Lamard, G. Quellec, B. Cochener, and G. Coatrieux, "Multi-tasking siamese networks for breast mass detection using dual-view mammogram matching," in *Machine Learning in Medical Imaging: 11th International Workshop, MLMI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Proceedings 11*, 2020: Springer, pp. 312-321.
- [17] M. A. Al-Masni *et al.*, "Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system," *Computer methods and programs in biomedicine*, vol. 157, pp. 85-94, 2018.
- [18] Y. Su, Q. Liu, W. Xie, and P. Hu, "YOLO-LOGO: A transformer-based YOLO segmentation model for breast mass detection and segmentation in digital mammograms," *Computer Methods and Programs in Biomedicine*, vol. 221, p. 106903, 2022.
- [19] B. Abhisheka, S. K. Biswas, and B. Purkayastha, "A comprehensive review on breast cancer detection, classification and segmentation using deep learning," *Archives of Computational Methods in Engineering*, pp. 1-30, 2023.
- [20] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, "Inbreast: toward a full-field digital mammographic database," *Academic radiology*, vol. 19, no. 2, pp. 236-248, 2012.

- [21] S. Singh and K. Bovis, "An evaluation of contrast enhancement techniques for mammographic breast masses," *IEEE Transactions on Information Technology in Biomedicine*, vol. 9, no. 1, pp. 109-119, 2005.
- [22] M. P. Sampat, A. C. Bovik, G. J. Whitman, and M. K. Markey, "A model-based framework for the detection of spiculated masses on mammography a," *Medical physics*, vol. 35, no. 5, pp. 2110-2123, 2008.
- [23] E. Song *et al.*, "Hybrid segmentation of mass in mammograms using template matching and dynamic programming," *Academic radiology*, vol. 17, no. 11, pp. 1414-1424, 2010.
- [24] H. T. Nguyen *et al.*, "VinDr-Mammo: A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography," *Scientific Data*, vol. 10, no. 1, p. 277, 2023.
- [25] M. Ü. Öziç, M. Barstuğan, and A. Özdamar, "An autonomous system design for mold loading on press brake machines using a camera platform, deep learning, and image processing," *Journal of Mechanical Science and Technology*, pp. 1-9, 2023.
- [26] F. Yuçe, M. Ü. Öziç, and M. Tassoker, "Detection of pulpal calcifications on bite-wing radiographs using deep learning," *Clinical Oral Investigations*, vol. 27, no. 6, pp. 2679-2689, 2023.
- [27] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263-7271.
- [28] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
- [29] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.
- [30] G. Jocher *et al.*, "ultralytics/yolov5: v3. 0," *Zenodo*, 2020.
- [31] J. Terven and D. Cordova-Esparza, "A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond," *arXiv preprint arXiv:2304.00501*, 2023.
- [32] D. G. Petrini, C. Shimizu, R. A. Roela, G. V. Valente, M. A. A. K. Folgueira, and H. Y. Kim, "Breast cancer diagnosis in two-view mammography using end-to-end trained efficientnet-based convolutional network," *Ieee access*, vol. 10, pp. 77723-77731, 2022.
- [33] D. Anyfantis, A. Koutras, G. Apostolopoulos, and I. Christoyianni, "Breast Density Transformations Using CycleGANs for Revealing Undetected Findings in Mammograms," *Signals*, vol. 4, no. 2, pp. 421-438, 2023.
- [34] E. Mahoro and M. A. Akhloufi, "Breast masses detection on mammograms using recent one-shot deep object detectors," in *2023 5th International Conference on Bio-engineering for Smart Technologies (BioSMART)*, 2023: IEEE, pp. 1-4.
- [35] H. T. Nguyen, S. B. Tran, D. B. Nguyen, H. H. Pham, and H. Q. Nguyen, "A novel multi-view deep learning approach for BI-RADS and density assessment of mammograms," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2022: IEEE, pp. 2144-2148.
- [36] S. B. Tran, H. T. Nguyen, C. Phan, H. H. Pham, and H. Q. Nguyen, "A Novel Transparency Strategy-based Data Augmentation Approach for BI-RADS Classification of Mammograms," *arXiv preprint arXiv:2203.10609*, 2022.
- [37] S. Magny, R. Shikhman, and A. Keppke, "Breast Imaging Reporting and Data System. 2022 Aug 29," *StatPearls [Internet]. StatPearls Publishing (Treasure Island, FL)*, 2022.
- [38] Y. Liu *et al.*, "High-temporal resolution DCE-MRI improves assessment of intra-and peri-breast lesions categorized as BI-RADS 4," *BMC Medical Imaging*, vol. 23, no. 1, p. 58, 2023.
- [39] A. Vourtsis and W. A. Berg, "Breast density implications and supplemental screening," *European radiology*, vol. 29, pp. 1762-1777, 2019.
- [40] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [41] R. Walsh and M. Tardy, "A Comparison of Techniques for Class Imbalance in Deep Learning Classification of Breast Cancer," *Diagnostics*, vol. 13, no. 1, p. 67, 2022.
- [42] S. B. Tran, H. T. Nguyen, C. Phan, H. Q. Nguyen, and H. H. Pham, "A Novel Transparency Strategy-based Data Augmentation Approach for BI-RADS Classification of Mammograms," in *2023 IEEE Statistical Signal Processing Workshop (SSP)*, 2023: IEEE, pp. 681-685.

- [43] V. Sampaio and F. R. Cordeiro, "A Study on Class Activation Map Methods to Detect Masses in Mammography Images using Weakly Supervised Learning," in *Anais do XIX Encontro Nacional de Inteligência Artificial e Computacional*, 2022: SBC, pp. 437-448.
- [44] H. N. Huynh, A. T. Tran, and T. N. Tran, "Region-of-Interest Optimization for Deep-Learning-Based Breast Cancer Detection in Mammograms," *Applied Sciences*, vol. 13, no. 12, p. 6894, 2023.
- [45] B. Gašparović, G. Mauša, J. Rukavina, and J. Lerga, "Evaluating YOLOV5, YOLOV6, YOLOV7, and YOLOV8 in Underwater Environment: Is There Real Improvement?," in *2023 8th International Conference on Smart and Sustainable Technologies (SpliTech)*, 2023: IEEE, pp. 1-4.
- [46] F. N. Ortataş and M. Kaya, "Performance Evaluation of YOLOv5, YOLOv7, and YOLOv8 Models in Traffic Sign Detection," in *2023 8th International Conference on Computer Science and Engineering (UBMK)*, 2023: IEEE, pp. 151-156.