

Breast detection in X-ray mastography with machine learning techniques

Luis Fernando Parra-Sánchez¹, Uriel Isaac Álvarez-Cárdenas¹, Arturo Daniel Jiménez Salazar¹, Blanca Olivia Murillo Ortiz², Luis Carlos Padierna García¹

¹División de Ciencia e Ingenierías, Universidad de Guanajuato, Campus León

²Unidad de Investigación en Epidemiología Clínica, Unidad Médica de Alta Especialidad No. 1 Bajío, Instituto Mexicano del Seguro Social, León, Guanajuato, México

lf.parrasanchez@ugto.mx, ui.alvarezcardenas@ugto.mx, ad.jimenezsalazar@ugto.mx, bomo907@hotmail.com, lc.padierna@ugto.mx

Abstract

Breast cancer is one of the leading causes of death in women in the world. An early detection is crucial to be able to perform an appropriate treatment and thus reduce mortality. An important tool for early detection is mastography, which is a technique to obtain images through X-rays. Through this technique it is possible to detect abnormalities in the breast tissue, thus allowing to detect the disease at an early stage. On the other hand, machine learning algorithms have provided promising results in automatic detection. In this work we combine both X-ray mastography and machine learning algorithms to perform breast detection. The YOLO neural network was used, which is a convolutional network that performs the detection in a single stage. For the training of the network, 216 mastographies from the dataset of breast mammography images with masses were used. Experimental results in position-independent breast detection showed an average confidence of 0.983, and a standard deviation of 0.024 in a sample of 200 images. This result allows to propose YOLO as an accurate preprocessing method for more sophisticated neural networks designed to solve breast cancer related tasks.

Key words: breast cancer, mastography, machine learning, breast detection.

Introduction

Cancer is the main cause of death around the world. Breast Cancer is the most common type of Cancer and with around 685,000 deaths is in the fifth place among different types of cancer. The use of screening tools as X-ray mammography are important to help practitioners with early detection of the disease. However, the breast cancer detection based on this technique requires the expertise of radiologists, and thus, the reliability of the result depends on the ability of human experts. Recently, machine learning models have been developed to help radiologists to analyze and draw conclusions based on automatic image processing. In this work we test a recent neural network model known YOLO (You only look once) to verify if this algorithm is capable to detect breast located in different position from a given mastography.

Materials and Methods

The INbreast database

The dataset presented in this study comprises breast mammography images with masses, obtained from the INbreast database. The original INbreast database was collected from the Breast Center at Centro Hospitalar de S. Joao, Porto, and includes data from August 2008 to July 2010. Among the 410 mammograms in the INbreast database, a total of 106 images with breast masses were selected for this dataset.

To enhance the dataset's size and diversity, data augmentation techniques were applied to the original images, resulting in a total of 7632 breast mammography images. The augmentation process involved multi-angle rotation ($\theta = 30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330^\circ$) and horizontal and vertical flipping, contributing to better model generalization and preventing overfitting. The dataset encompasses four

categories of breast density, categorized as Density 1, Density 2, Density 3, and Density 4, according to the BI-RADS standards. Additionally, each image is labeled with either a benign or malignant status, creating eight classification groups in total.

Before applying data augmentation, the images underwent contrast limited adaptive histogram equalization (CLAHE) preprocessing to enhance the visibility of breast masses. This dataset serves as a valuable resource for researchers and practitioners in the field of breast cancer screening, providing diverse and annotated mammography images for early detection and accurate classification of breast masses. It is made publicly accessible as an open-access dataset through a provided DOI (Digital Object Identifier) link, enabling further advancements in breast cancer diagnosis and treatment. [5]

YOLO Neural Network

Object detection is a computer vision task that works to identify and locate of specific objects in images or videos. In other words, the main goal consists in detect and delimit the regions where the objects of interest are located. There are several architectures used for object detection such as: R-CNN, Fast R-CNN, SSD, YOLO, among others. Each architecture, due to its structure and operation, presents different results in terms of performance, speed, and accuracy.

YOLO (You Only Look Once) is a popular neural network architecture used in computer vision tasks for object detection, the YOLO architecture consists of a convolutional neural network based on a Darknet architecture that performs the object detection task. Detection is normally carried out in several stages to obtain the coordinates of each object per image; however, the YOLO architecture simultaneously obtains the location and class probability for each object, which makes it more efficient than other architectures.[1] There are several versions of YOLO, we will use YOLOv3 illustrated in Figure 1.

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1x	Convolutional	32	1 × 1	128 × 128
	Convolutional	64	3 × 3	
	Residual			
	Convolutional	128	3 × 3 / 2	64 × 64
2x	Convolutional	64	1 × 1	64 × 64
	Convolutional	128	3 × 3	
	Residual			
	Convolutional	256	3 × 3 / 2	32 × 32
8x	Convolutional	128	1 × 1	32 × 32
	Convolutional	256	3 × 3	
	Residual			
	Convolutional	512	3 × 3 / 2	16 × 16
8x	Convolutional	256	1 × 1	16 × 16
	Convolutional	512	3 × 3	
	Residual			
	Convolutional	1024	3 × 3 / 2	8 × 8
4x	Convolutional	512	1 × 1	8 × 8
	Convolutional	1024	3 × 3	
	Residual			
	Avgpool		Global	
	Connected		1000	
	Softmax			

Figure 1: Darknet-53 architecture. Image obtained from [1]

YOLOv3 proposes a single convolutional neural network called Darknet-53 as its backbone, Darknet-53 consists of 53 convolutional layers, and is responsible for the extraction of the input image features.[2].

The network takes an input image that it divides into an N x N grid. An example is provided in Figure 2. Subsequently, localization and identification are applied on each of the cells to predict the bounding box and the confidence that the object belongs to a certain class that is sought to be identified in the image.[1].

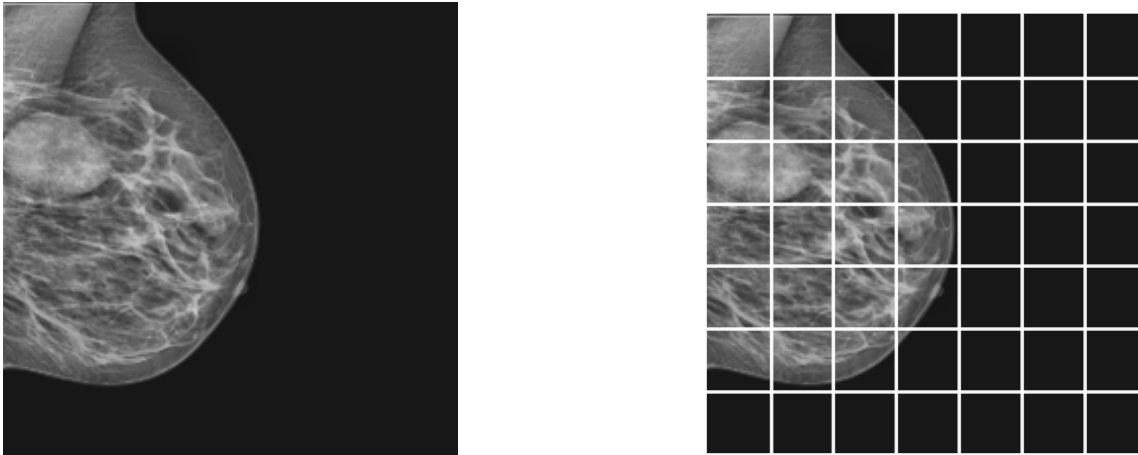


Figure 2: Initial processing of YOLO. Example image of the dataset described in [4].

Each bounding box is labeled with a vector with the following values.

Table 1. Bounding box predictions

Parameter	Representation
Confidence value	pc
Bounding box coordinate (x axis)	bx
Bounding box coordinate (y axis)	by
Height of the bounding box	bh
Bounding box width	bw
Class	c

pc is the “confidence value” or the probability that the object is correctly delimited and belongs to the class, c represents the class to which the object belongs, bx , by , bh and bw specify the bounding box (x, y, height and width coordinates). So, with the results it is possible to obtain a rectangular or a circular detection.

The intersection over the union (IOU) between the actual and predicted bounding box is used to calculate the confidence value in the bounding box.[1]

Let’s look at the following scenario (Figure 3) where the purple box is the real delimitation and the red box is the one obtained by the network.

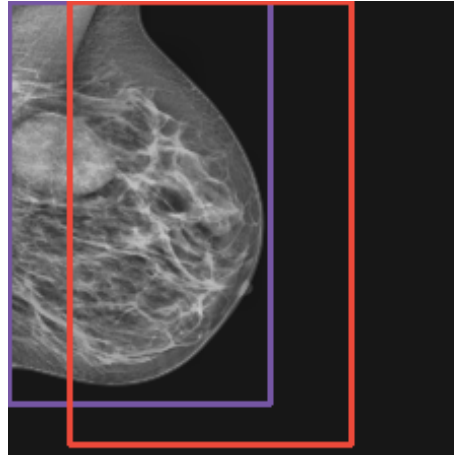


Figure 3: Representative figure of the detection.
Example image of the dataset described in [4].

In this case, the confidence value can be calculated by the following equation.

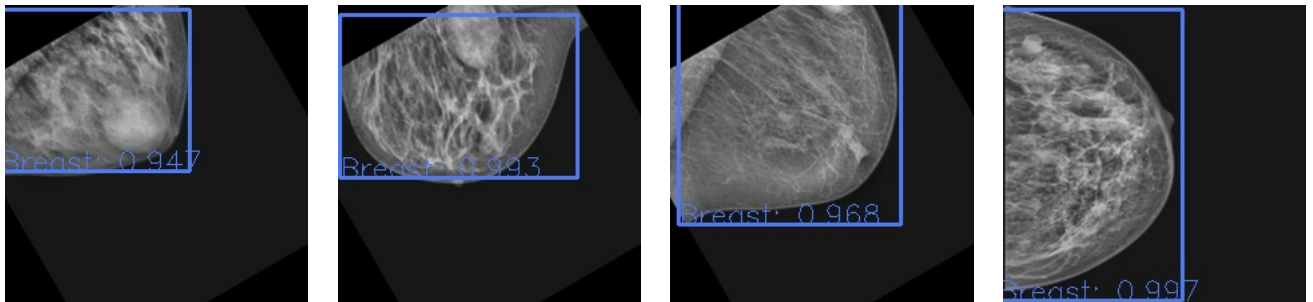
$$IOU = \frac{\text{Intersection area}}{\text{Union area}}$$

We can categorize the confidence values using an arbitrary threshold, e.g., $IOU > 0.5$, where the confidence must be greater than the threshold to be considered a good detection.[3]

Detection may encounter the problem of obtaining multiple overlapping detections for the same object, to solve this issue, a technique called Non-Maximum Suppression is used to select the most accurate detections by eliminating the others.[3]

Results

For the network training, 212 images from the set of images [4] were used. Each image was previously labeled in the same way as presented in Table 1 except for the confidence parameter. Some of the results obtained are presented below.



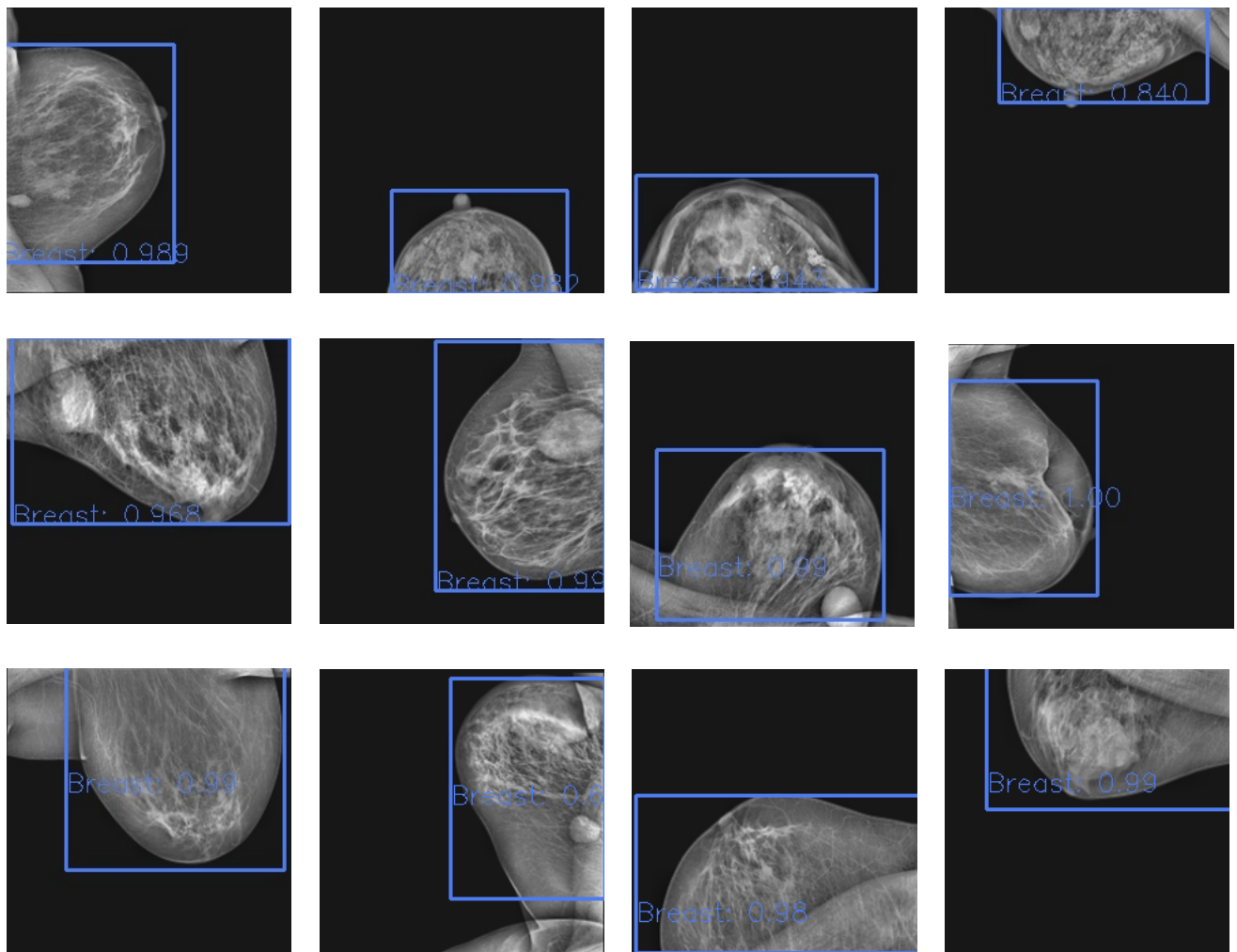


Figure 4.: Detection results

Subsequently, a total of 200 images were used to test the network. The results of the tests are shown in the following statistics.

Table 2. Statistics obtained in the screening test

Average accuracy	Variance (σ^2)	Standard deviation (σ)
0.983	0.001	0.024

To test the effectiveness of the network, it was subjected to unusual situations using images with multiple objects to cause stress on the network, in this case the following result was obtained (Figure 5)

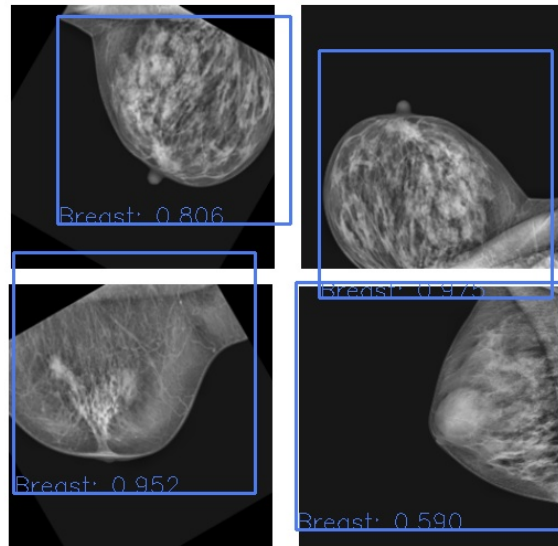


Figure 5: Multiple detection

We observed a decrease in the accuracy of some detections, which is to be expected given that the training considered only one breast. However, this is not a problem since in the detection phase only one breast per image is expected. In fact, having little decrease in accuracy, given these completely unforeseen conditions for the network indicates an excellent performance. Thus, the trained model can be applied in non-optimal conditions in the medical context.

Finally, to ensure the reliability of the network, another stress test was performed, where the network was subjected to scenarios radically different to the morphology of the breast to evaluate its behavior under unexpected situations.

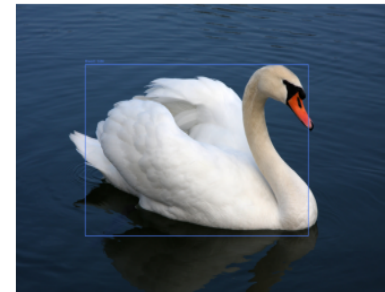
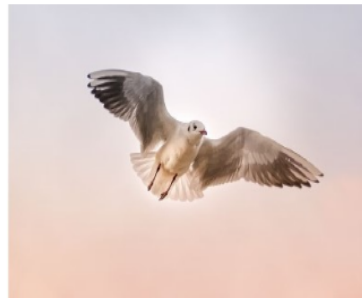


Figure 6: Stress test results

Where it can be observed that in the first two images we obtain a correct detection, but in the image of the goose a breast was detected with a confidence of 0.842, this may be due to the fact that in the case of that image there are similar shape characteristics, so it could be said that one of the vulnerabilities of the network occurs when images with a morphology similar to that of a breast are introduced, which makes it difficult for the neural network to distinguish them.

Conclusion

A YOLO type neural network architecture was implemented for breast detection in X-ray mastography. Good results were obtained in terms of breast detection, this can be implemented as a part of image preprocessing, which is the first step in the process of detecting tumors or abnormal bodies in breasts. Having a correct segmentation of the breast allows us to give an optimal image processing before resorting to deep learning models for tumor classification and avoid as many errors as possible. Having already obtained a good performance of YOLO type networks for whole breasts, indicates a good way forward in order to detect more specific issues within the breast.

Bibliography/References

- [1] Redmon, J., Divvala, S. K., Girshick, R., \& Farhadi, A. (2016). You only look once: Unified, Real-Time Object Detection. <https://doi.org/10.1109/cvpr.2016.91>
- [2] Redmon, J., \& Farhadi, A. (2018). YOLOV3: an incremental improvement. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.1804.02767>
- [3] Sharma, P. (2021). A practical guide to object detection using the popular YOLO framework – Part III (with Python codes). Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2018/12/practical-guide-object-detection-yolo-framewor-python/>
- [4] Huang, Mei-Ling; Lin, Ting-Yu (2020), "Dataset of Breast mammography images with Masses", Mendeley Data, V1, doi: 10.17632/x7bvzv6cvr.1
- [5] Huang, M., & Lin, T. (2020). Dataset of breast mammography images with masses. Data in Brief, 31, 105928. <https://doi.org/10.1016/j.dib.2020.105928>