Artificial Intelligence Research and Development
M. Villaret et al. (Eds.)
© 2021 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/FAIA210150

# Lesion Detection in Breast Tomosynthesis Using Efficient Deep Learning and Data Augmentation Techniques

# Loay HASSAN<sup>a,1</sup>, Mohamed ABDEL-NASSER<sup>a</sup>, Adel SALEH<sup>b</sup> and Domenec PUIG<sup>a</sup>

<sup>a</sup>Department of Computer Engineering and Mathematics, Universitat Rovira i Virgili, 43007 Tarragona, Spain <sup>b</sup>Gaist Solutions Ltd., Skipton BD23 2TZ, UK

> Abstract. Digital breast tomosynthesis (DBT) is one of the powerful breast cancer screening technologies. DBT can improve the ability of radiologists to detect breast cancer, especially in the case of dense breasts, where it beats mammography. Although many automated methods were proposed to detect breast lesions in mammographic images, very few methods were proposed for DBT due to the unavailability of enough annotated DBT images for training object detectors. In this paper, we present fully automated deep-learning breast lesion detection methods. Specifically, we study the effectiveness of two data augmentation techniques (channel replication and channel-concatenation) with five state-of-the-art deep learning detection models. Our preliminary results on a challenging publically available DBT dataset showed that the channel-concatenation data augmentation technique can significantly improve the breast lesion detection results for deep learning-based breast lesion detectors.

> Keywords. Digital breast tomosynthesis, Lesion detection, CAD systems, Deep learning

## 1. Introduction

Worldwide, breast cancer is one of the most common cancers in women and a leading cause of death [1]. The earlier breast cancer gets detected, the better of getting successful treatment. Mammography (X-ray images of the breast) and breast ultrasonography (BUS) are the most common diagnostic tools to detect breast cancer. However, the sensitivity and specificity of breast lesion detection in mammographic and BUS images are often limited by the presence of high breast density (dense fibro glandular tissue in the breast). The main reason for this limitation is the 2-D imaging modality of these images [2].

Digital breast tomosynthesis (DBT)—a new 3-D breast cancer screening technique has been introduced to reduce the effects of the superposition and overlapping problem of dense tissues in mammographic images by providing depth information, which yields

<sup>&</sup>lt;sup>1</sup>Corresponding Author; E-mail: loay.abdelrahimosmanhassan@urv.cat.

enhanced breast lesion detection rate [3]. In DBT, a sequence of mammographic images is acquired by moving the X-ray tube in an arc over a stationary compressed breast at multiple angles [4]. These individual mammographic images are then reconstructed into a series of 3-D and high-resolution slices by computer software. These 3-D images significantly reduce the effect of dense tissues that can hide breast lesions or make it difficult to be detected [5].

Although DBT has the ability to address the limitation of tissue superimposition in mammography [4] by providing superior tissue visualization, radiologists are required to examine a greater number of slices per breast volume, which creates clinical workflow challenges. Furthermore, as the number of slices to evaluate grows, physicians' oversight of findings increases. As a result, computer-aided detection (CAD) is necessary for clinical DBT and offers a larger clinical role in increasing work performance than traditional digital mammography. It is also likely that CAD will perform better with DBT images than with mammographic images due to the higher mass margin visibility [6].

Traditional CAD systems can lead to drawbacks as it relies on hand-engineered features. By contrast, the development of deep learning-based CAD systems, which relies instead on learning the features and classification decisions end-to-end, has been demonstrated to outperform traditional computer aid software in many computer vision problems. Deep learning has gained popularity in biomedical image analysis with a promising results in cancer classification [7] and lesions detection [8]. Practical experiments show that deep learning algorithms are robust and more accurate across datasets.

Recently, deep learning has been employed in DBT images to improve Lesion or cancer detection accuracy. Many researchers have employed several deep learning architectures to identify and classify breast cancer [9]. Fan et. al [10] introduced a deep learning framework based on a 3-D version of the Mask-RCNN model for mass detection and segmentation in DBT images. The main idea of this study is to use ResNet-Feature Pyramid Network (ResNet-FPN) as a backbone for their model. The DBT slices of the same image were passed into the mask-RCNN model one by one to generate bounding boxes for each slice with a confidence score of detected mass. These bounding boxes of a breast mass from the same set of DBT images are combined, if they have overlapping ratios (i.e. intersection over union ratios) greater than 0.5.

In [11], Lai et. al. presented a DBT mass automatic segmentation algorithm using a U-Net architecture. Their method consists of six stages: data pre-processing, patch extraction, data augmentation, U-Net segmentation, voting stage, and post-processing. In data pre-processing, a top-hat transform was applied to enhance the contrast between tumor location regions and background tissues. Images were split into patches and these patches were rotated by 90 degrees to increase the number of data for patch extraction and data augmentation steps. U-Net model with 23 layers was used to perform an end-toend pixel-wise segmentation. For the voting stage, a probabilistic prediction was made using the U-Net segmentation model for each slice. These predictions are then fused into the final image label using one of multiple voting schemes. Finally, they imposed volumetric constraints by removing clusters in the segmentation obtained by the U-Net to deal with small clusters that were mistakenly classified as breast masses.

Lotter et. al. [12] proposed a three-stage annotation-efficient deep learning model for breast cancer detection in mammographic and DBT images. In the first stage, a pretrained ResNet model was trained for lesion classification. This trained ResNet then is used as a backbone model for RetinaNet which is trained for lesion detection as a second stage. In the third stage, the ROI extracted from the second stage is used to classify optimized 2-D images condensed from DBT.

Buda et. al [13] developed a single-stage deep learning model to detect masses and architectural distortions in DBT images. The proposed model employs YOLO for 2-D object detection with DenseNet. In this work, different loss functions have been tested including binary cross-entropy, weighted binary cross-entropy, focal loss, and reduced focal loss. The results showed that their model performed best using focal loss.

Breast cancer detection in mammographic and DBT images is still a challenging task due to variability in breast lesion shapes and sizes and breast density. However, many automated lesion detection approaches were proposed in the literature to accurately detect breast cancer in mammographic images, very few approaches for DBT have been presented due to a lack of enough annotated DBT images for training object detectors. In this paper, we develop five automated deep learning-based breast lesion detection models for DBT images based on robust object detection models like YOLO [14] and Faster R-CNN [15]. Additionally, we investigate the impact of two data augmentation techniques called channel-replication and chancel-concatenation in improving the breast lesion detection results of deep learning models.

The remainder of this paper is designed as follows. Section 2 describes the proposed detection system. The experimental results and discussion are presented in section 3. Section 4 concludes the paper and provided some lines for future work.

## 2. Proposed Lesion Detection System

Figure 1 shows the proposed lesion detection method. The key elements of the proposed method are data augmentation, deep learning-based detector, and non-maximum suppression (NMS). Below, we describe the proposed method in detail.



Figure 1. Lesion detection in breast tomosynthesis images based on deep learning models and robust data augmentation techniques.

### 2.1. Data augmentation

In data augmentation techniques, the number of training images is increased with different image manipulation algorithms. We use two different data augmentation techniques in this study. • Augmentation technique 1 (Aug1): channel replication. We flipped all images in the training set horizontally, then for each original and flipped image  $I_s$  we apply gamma correction to adjust the overall brightness of an image using Equation 1.

$$I_{\gamma} = 255 \times \left(\frac{I_s}{255}\right)^{\gamma} \tag{1}$$

Where  $I_{\gamma}$  is the output image for gamma correction and  $\gamma$  is the gamma correction factor.

Also, we apply the contrast limited adaptive histogram equalization (CLAHE) to enhance the image Local Contrast [16]. To calculate the clip limit for the CLAHE algorithm, we used Equation 2.

$$cliplimit = \frac{W \times H}{L} \left( 1 + \frac{\alpha}{100} \left( S_{\max} - 1 \right) \right)$$
(2)

Where  $W \times H$  is the number of pixels in each histogram calculated region, L is the number of gray-scales,  $\alpha$  is a clip factor, and  $S_{max}$  is the maximum allowable slope. However,  $S_{max}$  should set to four for still X-ray images.

• Augmentation technique 2 (Aug2): channel-concatenation. We used a new 3channel data augmentation method by concatenating the original image with two post-processed images as proposed in [17]. With this augmentation technique, rather than concatenating the three gray-scale images, we used two filtered images  $(I_{\gamma} \text{ with } \gamma = 0.5 \text{ and } I_{clahe} \text{ with } \alpha = 1)$  to concatenate with the original gray-scale image  $I_g$ . So, for each image in the training set, we produce a new 3-channel Image I as shown in Equation 3.

$$I = Concat(I_g, I_\gamma, I_{clahe})$$
(3)

Here, I, $I_g$ , $I_\gamma$ , and  $I_{clahe}$  is output image, image after gamma correction and image after CLAHE equalization, respectively. Figure 2 shows sample visualization different data augmentation techniques.



Figure 2. Samples for different data augmentation techniques. a. The original image, b. Outputs of Aug1 and c. The output of Aug2.

#### 2.2. Individual deep learning based detection models

To build the individual deep learning-based detection models, we employed two widely used and efficient deep learning-based object detectors: YOLO and Faster-RCNN.

YOLO [14] is an object detection model that implements all of the detection stages to detect an object using a single neural network. The YOLO algorithm has obtained impressive specifications that outperform the leading detection algorithms in terms of both speed and accuracy for detecting and determining object location. The main idea of the YOLO algorithm is to apply a single neural network to the full image and then divide the image into a grid cell with the size of  $S \times S$  (7 × 7 default). If the center of an object falls into a grid cell, that grid cell predicts bounding boxes includes confidence probabilities for the objects.

The YOLO algorithm was upgraded to five versions (including the original version) since the time of its publication. In 2020, Bochkovskiy et. al [18] presented YOLOv4 with many of the most innovative ideas coming out of the computer vision research community for each part of the model architecture. After a few months, Glenn Jocher released YOLOv5. Unlike previous YOLO versions that picked the five best-fit anchor boxes for the COCO dataset and use them as default, YOLOv5 integrates the anchor box selection process by learning the best anchor boxes automatically for the training images and use them during training. Besides, YOLOv5 uses mosaic augmentation that combines four images into four tiles of random ratio, helping the models to detect small objects.

Faster R-CNN detector [19] usually comprises four major parts: 1) a feature extractor stage–usually using a CNN, 2) a region proposal (RPN) algorithm to predict bounding boxes of possible objects in the image with a confidence score, 3) a classification layer to predict which class this object belongs to, and 4) a regression layer to make the coordinates of the object bounding box more precise. The main idea of Faster R-CNN is using a convolutional network for RPN, instead of using a selective search algorithm to generate the region proposals that yield accelerating training time and improving feature representation.

## 2.3. Implementation

In our work, Firstly, the DBT images dataset was split into training and testing sets (patient-wise splitting). In the training phase, we apply data augmentation techniques mentioned in section 2.1 to produce two training sets (training set by data augmentation technique 1 and training set by data augmentation technique 2). For each of these training sets, we train each of the mentioned detectors individually. It is worth noting that for the YOLOv5 detector, we train four models with different sizes (YOLO-Small (S), YOLO-Medium (M), YOLO-Large (L), and YOLO-XLarge (XL)).

Secondly, in the testing phase, the trained models are used to predict bounding boxes for each DBT image in the test set. Each bounding box consists of a class that the object may belong to, a coordinates list [x1,y1,x2,y2] and confidence score. The confidence represents scores reflect how confident the model is that the box contains an object, i.e. any Tumor in the box, P(Tumor). All predicted bounding boxes from a single DBT image are combined in a single bounding boxes list. Then, this list is passed to (NMS) algorithm to select the best bounding box out of a set of overlapping or duplicated boxes. The selection criteria are based on the confidence probability threshold along with the intersection over union (IoU) overlap measure threshold.

To reduce the computational complexity, we resized the original DBT images to  $640 \times 640$ . In the training phase, we implement two data augmentation techniques to increase the number of training data. It worth noting that, in this work we used the pytorch implementation of the YOLOv5 [19] and Faster-RCNN [20]. We utilized the patch size 8, and the maximum number of training epochs was 50 for the YOLO model and 20 for the faster-RCNN model.

# 2.4. Evaluation Metrics

In this study, we use three of the most popular metrics used to evaluate object detection models [21]: the true positive rate (TPR), F1-score and PASCAL VOC mean average precision (mAP) to assess the performance of the lesion detection models. The true positive rate (TPR, also called sensitivity) is the probability that an actual positive will test positive and calculated By Equation 4.

$$TPR = \frac{TP}{TP + FN} \tag{4}$$

where TP, and FN is true positive and False Negative detections, respectivly. The F1 score is the harmonic mean of precision and recall, and calculated as follows

$$Precision = \frac{TP}{TP + FP}, \quad Recall = \frac{TP}{TP + FN}$$
(5)

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$
(6)

Where FP is the false positive detections, precision is the ratio of the number of true positives to the total number of positive predictions and recall is the same as TPR. The mean average precision (mAP) is precision averages across all recall values between 0 and 1 at various IoU thresholds. By interpolating across all points, mAP can be interpreted as the area under the curve of the precision-recall curve.

## 3. Experiments and discussion

### 3.1. Dataset

In our experiments, we used the DBTex challenge dataset [22]. It contains a total of 1000 breast tomosynthesis scans from 985 patients. Of 101 patients, only 224 DBT images have annotations (the rest images are not fully annotated). It should be noted that in our experiments we divided the DBT images dataset into training and testing sets (patient-wise): 71 patients for training and 30 patients for testing.

#### 3.2. Experimental results

In the training phase, we used the two training sets produced from the two data augmentation techniques (Aug1 and Aug2). With Aug1, we produced a total of 1168 training DBT images, while with Aug2 we produced a total of 292 training DBT images.

Table 1 presents a quantitive comparison between the four YOLOv5 models (YOLO-S, YOLO-M, YOLO-L and YOLO-XL) and the faster R-CNN model trained on the training dataset produced by Aug1 in terms of TPR, F1-score, and mean average precision–mAP (IoU threshold = 5). As one can see, YOLO-S achieved the best lesion detection results when compared to the other YOLO models. However, faster R-CNN has more promising breast lesion detection results. It surpassed all YOLO models on all measures. These breast lesion detection results demonstrate that the faster R-CNN could be more suitable for DBT images, however, it has a higher computational complexity than YOLO models.

|               | YOLOv5 |      |      |      | Easter P. CNN  |  |
|---------------|--------|------|------|------|----------------|--|
|               | S      | М    | L    | XL   | Paster R-CININ |  |
| TPR           | 34.8   | 31.8 | 24.2 | 22.7 | 50             |  |
| F1-Score      | 48.5   | 45.7 | 36.1 | 35.7 | 54.1           |  |
| mAP [iou=0.5] | 31.8   | 34.1 | 26.2 | 26.7 | 45.1           |  |

Table 1. The performance of the deep learning detection methods with Aug1.

| Table 2. | The | performance | of the | deep | learning | detection | methods | with | Aug2. |
|----------|-----|-------------|--------|------|----------|-----------|---------|------|-------|
|          |     |             |        |      |          |           |         |      |       |

|               | YOLOv5 |      |      |      | Easter P. CNN  |  |
|---------------|--------|------|------|------|----------------|--|
|               | S      | М    | L    | XL   | Paster R-CIVIN |  |
| TPR           | 47     | 39.4 | 39.4 | 47   | 56.1           |  |
| F1-Score      | 52.5   | 51.7 | 56.6 | 51.4 | 57.4           |  |
| mAP [iou=0.5] | 48.7   | 38.9 | 40.4 | 41.8 | 46.8           |  |

Table 2 compares the breast lesion detection results of the models when trained on the training dataset produced by Aug2 in terms of TPR, F1-score, and mAP. As shown, when comparing values from Table 2 to values from Table1, we can argue that training the deep learning detectors based on Aug2 can yields noticeable improvements on all TPR, F1-Score, and mAP metrics. With Aug2, the performance of the YOLO-S model was also increased by 17 points in terms of mAP. Besides, the TPR and F1-score of the faster RCNN were also advanced 6% and 3.3%, respectively.

Figure 3 presents the breast lesion detection results of the YOLO models and faster-RCNN. The red boxes stand for the ground truth while the green boxes indicate the detected lesions with confidence scores with IoU greater than 0.5. As shown, the breast lesion detections performed by the models trained with the Aug2 training dataset (second row) have higher confidence scores than models trained with the Aug1 training dataset (first row). In turn, the faster R-CNN predicts the most accurate bounding boxes (IoU > 0.8).

On the basis of the above analysis, we can conclude that to second data augmentation technique (Aug2) can significantly improve the breast lesion detection results for deep learning-based breast lesion detectors like YOLO models and faster R-CNN.



**Figure 3.** Examples of lesion detection using YOLOv5 and faster R-CNN. Note that the first row includes the results of Aug1 while the second row includes the results of Aug2. a and c show the detection results of YOLO-S, YOLO-M, YOLO-L and YOLO-XL from left to right, respectively. b and d show the detection results of Faster R-CNN.

All the experiments were performed using Pytorch framework using a 64-bit Ubuntu operating system with 3.6 GHz intel core i7 with 32GB of RAM and Nvidia GTX1080 with 8GB of video RAM.

# 4. Conclusions

In this paper, we investigated the strength of two data augmentation strategies (channel-replicate and channel-concatenation) while building five breast lesion detection models based on deep learning for digital breast tomosynthesis. We demonstrated that applying the channel-concatenation data augmentation strategy helps improve the detection accuracy of all deep learning models. With a publicly available digital breast tomosynthesis dataset, our experiments showed that the mAP score of YOLO models was increased by 17% approximately, while the F1-score of the faster-RCNN detector was improved by 3%. The future work will be focused on the development of a lesion detection approach based on the combination of robust deep learning-based detectors.

## Acknowledgement

The Spanish Government partly supported this research through Project PID2019-105789RB-I00.

## References

- H.D. Cheng, Juan Shan, Wen Ju, Yanhui Guo, and Ling Zhang. Automated breast cancer detection and classification using ultrasound images: A survey. *Pattern Recognition*, 43(1):299–317, 2010.
- [2] Emine Devolli-Disha, Suzana Manxhuka-Kërliu, Halit Ymeri, and Arben Kutllovci. Comparative accuracy of mammography and ultrasound in women with breast symptoms according to age and breast density. *Bosnian journal of basic medical sciences*, 9(2):131–136, May 2009.
- [3] James T Dobbins and Devon J Godfrey. Digital x-ray tomosynthesis: current state of the art and clinical potential. *Physics in Medicine and Biology*, 48(19):R65–R106, sep 2003.
- [4] Mark A. Helvie. Digital mammography imaging: breast tomosynthesis and advanced applications. *Ra-diologic clinics of North America*, 48(5):917–929, Sep 2010.
- [5] Julianne S. Greenberg, Marcia C. Javitt, Jason Katzen, Sara Michael, and Agnes E. Holland. Clinical performance metrics of 3d digital breast tomosynthesis compared with 2d digital mammography for breast cancer screening in community practice. *American Journal of Roentgenology*, 203(3):687–693, Sep 2014.
- [6] Ming Fan, Yuanzhe Li, Shuo Zheng, Weijun Peng, Wei Tang, and Lihua Li. Computer-aided detection of mass in digital breast tomosynthesis using a faster region-based convolutional neural network. *Methods*, 166:103–111, 2019. Deep Learning in Bioinformatics.
- [7] Muhammad Imran Razzak, Saeeda Naz, and Ahmad Zaib. Deep Learning for Medical Image Processing: Overview, Challenges and the Future, pages 323–350. Springer International Publishing, Cham, 2018.
- [8] Dezső Ribli, Anna Horváth, Zsuzsa Unger, Péter Pollner, and István Csabai. Detecting and classifying lesions in mammograms with deep learning. *Scientific Reports*, 8(1):4165, Mar 2018.
- [9] Jun Bai, Russell Posner, Tianyu Wang, Clifford Yang, and Sheida Nabavi. Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: A review. *Medical Image Analysis*, 71:102049, 2021.
- [10] Ming Fan, Huizhong Zheng, Shuo Zheng, Chao You, Yajia Gu, Xin Gao, Weijun Peng, and Lihua Li. Mass detection and segmentation in digital breast tomosynthesis using 3d-mask region-based convolutional neural network: A comparative analysis. *Frontiers in molecular biosciences*, 7:599333–599333, Nov 2020.
- [11] Xiaobo Lai, Weiji Yang, and Ruipeng Li. Dbt masses automatic segmentation using u-net neural networks. Computational and Mathematical Methods in Medicine, 2020:7156165, Jan 2020.
- [12] William Lotter, Abdul Rahman Diab, Bryan Haslam, Jiye G. Kim, Giorgia Grisot, Eric Wu, Kevin Wu, Jorge Onieva Onieva, Yun Boyer, Jerrold L. Boxerman, Meiyun Wang, Mack Bandler, Gopal R. Vijayaraghavan, and A. Gregory Sorensen. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nature Medicine*, 27(2):244–249, Feb 2021.
- [13] Mateusz Buda, Ashirbani Saha, Ruth Walsh, Sujata Ghate, Nianyi Li, Albert Swieciki, Joseph Y. Lo, and Maciej A. Mazurowski. Detection of masses and architectural distortions in digital breast tomosynthesis: a publicly available dataset of 5,060 patients and a deep learning model, 2021.
- [14] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, realtime object detection. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 779–788, 2016.
- [15] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett, editors, Advances in Neural Information Processing Systems, volume 28. Curran Associates, Inc., 2015.
- [16] Kitti Koonsanit, Saowapak Thongvigitmanee, Napapong Pongnapang, and Pairash Thajchayapong. Image enhancement on digital x-ray images using n-clahe. In 2017 10th Biomedical Engineering International Conference (BMEiCON), pages 1–4, 2017.
- [17] Moi Hoon Yap, Manu Goyal, Fatima Osman, Robert Mart, Erika Denton, Arne Juette, and Reyer Zwiggelaar. Breast ultrasound region of interest detection and lesion localisation. *Artificial Intelligence* in Medicine, 107:101880, 2020.
- [18] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao. YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv e-prints, page arXiv:2004.10934, April 2020.
- [19] Glenn Jocher. Yolov5 Pytorch Implementation. https://github.com/ultralytics/yolov5. Accessed: 2021-05-29.

- [20] Faster-RCNN Pytorch Implementation. https://pytorch.org/tutorials/intermediate/ torchvision\_tutorial.html. Accessed: 2021-05-29.
- [21] Rafael Padilla, Sergio L. Netto, and Eduardo A. B. da Silva. A survey on performance metrics for objectdetection algorithms. In 2020 International Conference on Systems, Signals and Image Processing (IWSSIP), pages 237–242, 2020.
- [22] SPIE-AAPM-NCI DAIR Digital Breast Tomosynthesis Lesion Detection Challenge. https://www.aapm.org/GrandChallenge/DBTex/. Accessed: 2021-05-23.