An Abnormity AI Detection Method of Breast Mammography

Peng Guo, Jie Zheng, Yong zhen Wang and Yang Hu

ABSTRACT

We propose a method to detect the abnormity in mammograms by using a modified Faster-region convolution neural network (Faster R-CNN). In our approach, the Saliency detection was used to reserve the details of mass and improve the contrast between masses and normal tissues. The size of the anchor is modified to improve the mass detection accuracy. To better utilize the multi-level convolution features and enrich the discriminant information of each bounding box, the multi-level region-of-interest pooling (MLRP) method substitutes for the original region-of-interest pooling method of Faster R-CNN. We demonstrate experimentally the good performance of the proposed method in abnormity detection of real clinical mammograms that relatively high detection accuracy is obtained with false positive rate reduced.

KEYWORDS
Mass Detection, Faster R-CNN, Saliency Detection, MLRP.

INTRODUCTION

Cancer has threatened people health all over the world. With the gradual aging population, the acceleration of industrialization and urbanization, and the accumulation of unhealthy lifestyle, the prevention and control of cancer is grim. Breast cancer as the second deadly cancer of woman, has aroused public attention. It has been a consensus that early detection and treatment could greatly improve the cure rate and reduce the recurrence rate of breast cancer[1].

The abnormal area in mammograms includes mass, calcification, bilateral asymmetry and structure distortion, in which the mass and calcification cluster are the most common symptoms of breast cancer[2]. Because of being indistinguishable from the surrounding normal tissues, Mass detection is of paramount importance in computer-aided detection (CAD) systems[3]. Therefore, if not emphasized, the abnormal area in this paper refers to the mass.

The traditional mass detection CAD systems always contain three steps. Firstly, image preprocessing of the digital mammograms, it plays an important role in suppressing noises and improving the contrast of the image. Secondly, image segmentation is supposed to locate the suspicious regions, the existed algorithms involve a variety of methods that are region-based, contour-based and clustering[4].
Finally, effective features are extracted[5]-[7] and selected for classifying[8], [9] the suspicious region as mass or normal tissue. There are two main problems in traditional mass detection CAD systems:(1) the robustness and generalization ability of hand-crafted features are not great;(2) the methods to obtain the region of interest are mainly based on specific shape and gray value, which are lack of universality.

Since 2012, as deep learning techniques developing rapidly, the use of convolutional neural networks (CNN) comes to a superior status in object detection field. The main factors that CNN is superior to other traditional algorithms are: (1) large-scale training data;(2) the network can discover good features;(3) the acceleration of GPU makes large-scale parallel computation become a reality and accelerates the training process of model. In 2014, Girshick et al.[10] proposed the R-CNN object detection model, combining Region Proposals with CNN, applying SVM to classify deep features and linear regression to get final refined boxes. In 2015, Girshick[11] integrated Spatial Pyramid Pooling[12] method into R-CNN and proposed Fast R-CNN model, which greatly reduced the number of operations of CNN and saved calculation and storage resources. In 2016, in order to reduce the computational pressure of extracting candidate boxes, Ren et al.[13] proposed Faster R-CNN model to generate candidate boxes using Region Proposal Networks, which realized end-to-end object detection.

MATERIALS AND METHODS

GENERAL INFORMATION

The first part of mammograms used in this study are available on the Digital Dataset for Screening Mammography (DDSM) provided by the University of South Florida and the second part come from INbreast provided by a Breast Centre located in a University Hospital in Portugal. The rest of the data are from cooperative hospitals collected by the Luna DR machine of Lanmage company, which are diagnosed and analyzed by experts. Each case of the database includes two images from each breast with CC and MLO views. There are 4190 mammograms used as the evaluated dataset of the proposed system. 2514 mammograms from the evaluated are randomly selected to establish the training dataset.

DATA AUGMENTATION

Data augmentation techniques such as cropping, rotation, padding are commonly used to train large neural networks. We firstly take the sub-images around the mass according to the designed ratio with the position information of the mass marked by the expert in the dataset, and generate 3 new samples using 90, 180 and 270 degrees rotation transformations for each sub-images. That is, the total number of training set is 7983 after filtering the training data. The rest of mammograms in the evaluated dataset are used as the test dataset to evaluate the performance.
SALIENCY DETECTION

Frequency-tuned algorithm[14] is considered on the basis of its capability to generate saliency map from an input image. The method defines the saliency of each pixel in the image by comparing the local color and brightness features, and preserves the boundary of the ROI by retaining more frequency information, which is simple and efficient, and has good performance on image detection without complicated background. The steps involved in computing the saliency map are explained below:

Gaussian blur is applied to the original image, while converting the color space from RGB to Lab.

The mean of L, a and b channels in transformed image are calculated.

The Euclidean distance of image after Gaussian blurring and averaging is calculated to generate the final saliency map.

FASTER R-CNN DETECTION MODEL

Faster R-CNN is one of the mainstream frameworks for object detection in the field of deep learning. This detection framework integrates several steps of object detection methods, including generating region proposals, extracting proposals’ features, classifying region proposals and refining by bounding box regression into a convolutional neural network. Faster R-CNN evolved from the R-CNN and Fast R-CNN models. Compared with the time-consuming selective search strategy adopted by the latter two for generating region proposals, Faster R-CNN truly realizes real-time detection and improves processing speed while ensuring detection accuracy. The pipeline of our proposed method includes several processes presented in Figure 1.

(1) Region Proposal Network (RPN): Anchors with different scales and aspect ratios of each pixel in the feature map generated by the basic network are set up, and IoU strategy is used to select positive and negative samples for RPN training: an anchor that has an IoU overlap higher than 0.7 with any ground-truth box is assigned a positive label, a non-positive anchor if its IoU ratio is lower than 0.3 for all ground-truth boxes is assigned a negative label. Anchors that are neither positive nor negative do not contribute to the training objective. The loss function for an image is defined as

\[ L((p_i), (t_i)) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p^*_i) + \lambda \frac{1}{N_{neg}} \sum_i p^*_i L_{reg}(t_i, t^*_i) \]  

(1)

Here, \( i \) is the index of an anchor in a mini-batch and \( p_i \) is the predicted probability of anchor \( i \) being an object. The ground-truth label \( p^*_i \) is 1 if the anchor is positive, and is 0 if the anchor is negative. \( t_i \) is a vector representing the 4 parameterized coordinates of the predicted bounding box, an \( t^*_i \) is that of the ground-truth box associates with a positive anchor. The classification loss \( L_{cls}() \) is log loss over two classes. For the regression loss \( L_{reg}(t_i, t^*_i) = \sum_{x,y,w,h} \text{smooth}_{t_i}(t_i - t^*_i) \). Here, the \( \text{smooth}_{t_i} \) loss function is defined as
\[
\text{smooth}_t(x) = \begin{cases} 
0.5x^2, & |x| < 1 \\
|x| - 0.5, & |x| \geq 1 
\end{cases}
\] (2)

In (1), the term \( p_i^* L_{reg} \) means the regression loss is activated only for positive anchors (\( p_i^* = 1 \)) and is disable otherwise (\( p_i^* = 0 \)). The two terms are normalized by \( N_{cls} \) and \( N_{reg} \) and weighted by a balancing parameter \( \lambda \). For bounding box regression, the parameterizations of the 4 coordinates are adopted following.

\[
t^x_\alpha = \frac{(x^\ast - x_\alpha)}{w_\alpha},
\]

(3)

\[
t^y_\alpha = \frac{(y^\ast - y_\alpha)}{h_\alpha},
\]

(4)

\[
t^w_\alpha = \log\left(\frac{w^\ast}{w_\alpha}\right),
\]

(5)

\[
t^h_\alpha = \log\left(\frac{h^\ast}{h_\alpha}\right),
\]

(6)

where \( x^\ast, y^\ast, w^\ast, h^\ast \) denote the ground-truth box's center coordinates and its width and height. Variable \( x_\alpha \) is for the predicted box's (likewise for \( y^\ast, w^\ast, h^\ast \)).

(2) Fast R-CNN: The network first processes the whole image with several convolutional and max pooling layers to produce a conv feature map. Then, for each object proposal a region of interest (RoI) pooling layer extracts a fixed-length feature vector from the feature map. Here, in the mammographic mass detection system, we applied the multi-level region-of-interest pooling (MLRP)[15] to extracting the fixed-length feature vector. To better utilize the multi-level convolutional features and enrich the discriminant information of each bounding box, MLRP is performed over the Conv4_3 as well as Conv5_3 convolutional feature maps of the VGG16 network. Specifically, channel concatenation is applied on each pooled feature and concatenated feature is encoded by convolutional layer. The convolutional layer combines the multi-level pooled features and learns the fusion weights in the training process as well as reducing the dimensions to match VGG16’s first fully-connected layer. The multi-level weighted fusion feature is then propagated to the follow-up bounding box classification and regression model.
RESULTS

In this paper, the Tensorflow+Keras framework in Windows environment was adopted for model training, and the Nvidia 1080Ti GPU was used for accelerating. The pre-training model parameters of VGG16 based on ImageNet were used as the initial training parameters.

EVALUATION OF RESULTS

In the mammographic mass detection system, what to evaluate the performance of the system are sensitivity (Sens) and false positive per Image (FPI). Computational methods of these two measurements are as follows.

\[
Sens = \frac{\text{Numbers of True Positive Marks}}{\text{Numbers of Lesion}}\quad (7)
\]

\[
FPI = \frac{\text{Numbers of False Positive Marks}}{\text{Numbers of Images}}\quad (8)
\]

where Sens means the ratio between the number of detected true positive marks and lesions marked by experts in dataset. FPI refers to the ratio between the number of detected false positive marks and the number of images in dataset.

EVALUATION OF SALIENCY DETECTION

Figure 2(b) shows an example of saliency detection result. From Figure 2(b), it is obvious that the contrast between the mass and surrounding tissues is significantly
improved, and the RoI is made stand out relative to its neighbors in some ways. The necessity of saliency detection before Faster R-CNN algorithm is testified with the Table I. As shown in Table I, the second column indicates the detection framework without saliency detection. The third column represents the combination of Faster R-CNN and saliency detection. It has been proved that the saliency detection method is valid for mammograms from the results.

![Image of saliency detection result](image)

Figure 2. Example of saliency detection result: (a) Original image, and (b) the saliency detection result.

### TABLE I. COMPARISON OF FASTER R-CNN AND SALIENCY+FASTER R-CNN.

| Contents | Faster R-CNN | Saliency+Faster R-CNN |
|----------|--------------|-----------------------|
| Sens     | 0.9112(1232/1352) | 0.9246(1250/1352) |
| FPI      | 1.6516(2768/1676) | 1.6623(2786/1676) |

### ADAPTED ANCHORS

According to the size of the mass marked by experts in the training dataset, the largest mass reaches approximately $1000 \times 1000$ pixels, the smallest mass is only $50 \times 50$ pixels. To better detecting the widely varied range of mass in dataset, we set the scales and aspect ratios as $(50, 100, 200, 300, 500, 650, 800, 1000, 1050)$ and $(1:1)$ respectively. The validity of the adapted anchors is verified with the Table II. With regard to Table II, the original anchors and adapted anchors are used to make the comparison experiment.

### TABLE II. COMPARISON OF ORIGINAL ANCHORS AND ADAPTED ANCHORS.

| Contents | Original Anchors | Adapted Anchors |
|----------|------------------|-----------------|
| Sens     | 0.9246(1250/1352) | 0.9334(1262/1352) |
| FPI      | 1.6623(2786/1676) | 1.6504(2766/1676) |
MULTI-LEVEL REGION-OF-INTEREST POOLING (MLRP)

The MLRP method integrates multi-level convolutional features and enriches the discriminant information of each bounding box. Therefore, it has obvious effect on improving the performance of mammographic mass detection system. The effect of RoI Pooling and MLRP method are compared in Table III.

The final detection results of the proposed method on the example are presented in Figure 3. Each detection result is labeled with the predicted class and the precision value. 3A and 3B are from DDSM database, 3C and 3D are from INbreast database, while 3E and 3F are the clinic mammograms caputured by Luna DR machine. We can see the proposed mass detection method works well in detecting various types of masses based on different mammograms.

| Contents | ROI Pooling | MLRP |
|----------|-------------|------|
| Sens     | 0.9334(1262/1352) | 0.9467(1280/1352) |
| FPI      | 1.6504(2766/1676) | 1.6521(2769/1676) |

CONCLUSION

Recent advances in machine learning, especially in deep learning, are critical to recognizing classifying and quantifying the medical images. In this paper, we apply the adapted faster r-cnn algorithm to mammographic mass detection system. First of all, augmentation is applied to enlarging the training dataset by means of taking subimages and rotation transformations. Secondly, a saliency model based on FT method is adopted to improve the contrast between masses and normal tissues and reserve the details of mass. And then, according to the variable mass size in the database, the original anchor scales and aspect ratios in Faster R-CNN algorithm are modified. Finally, the multi-level region-of-interest pooling strategy is a substitute for original RoI pooling method aims to extract the fixed-length feature vector. Our proposed mammographic mass detection system is effective proved by the experimental results. However, how to establish large-scale training dataset and
work out mass segmentation approach of mammograms to locate mass more accurately will continue to be our focus.

REFERENCES

1. J. Ma and A. Jemal, Breast Cancer Metastasis and Drug Resistance: Chapter 1, Breast Cancer Statistics, A. Ahmad, Ed. New York: Springer, 2013.
2. R. A. Castellino, “Computer aided detection (cad): An overview,” Cancer Imaging: The Official Publication of the International Cancer Imaging Society, vol. 5, no. 1, pp. 17–19, 2005.
3. M. P. Sampat, M. K. Markey, and A. C. Bovik, “Computer-aided detection and diagnosis in mammography,” Handbook of Image and Video Processing, pp. 1195–1217, 2005.
4. A. Oliver, J. Freixenet, J. Martí, et al, “A review of automatic mass detection and segmentation in mammographic images,” Medical Image Analysis, vol. 14, no. 2, pp. 87–110, 2010.
5. H.-P. Chan, D. T. Wei, M. A. Helvie, et al, “Computer-aided classification of mammographic masses and normal tissue: Linear discriminant analysis in texture feature space,” Physics in Medicine and Biology, vol. 40, no. 5, pp. 857–876, 1995.
6. N. R. Mudigonda, R. Rangayyan, J. E. L. Desautels, “Gradient and texture analysis for the classification of mammographic masses,” IEEE Transactions on Medical Imaging, vol. 19, no. 10, pp. 1032–1043, 2000.
7. L. H. Li, Y. Zheng, L. Zhang, R. A. Clark, “False-positive reduction in cad mass detection using a competitive classification strategy,” Medical Physics, vol. 28, no. 2, pp. 250–258, 2001.
8. R. Campanini, D. Dongiovanni, E. Iampieri, et al, “A novel featureless approach to mass detection in digital mammograms based on support vector machines,” Physics in Medicine and Biology, vol. 49, no. 6, pp. 961–975, 2004.
9. X. S. Zhang, M. H. Wang, M. Ji, “Sparse representation for detection of microcalcification clusters,” Telkommika, vol. 10, no. 3, pp. 545–550, 2012.
10. R. Girshick, J. Donahue, T. Darrell, J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2014.
11. R. Girshick, “Fast r-cnn,” in The IEEE International Conference on Computer Vision (ICCV), December 2015.
12. K. He, X. Zhang, S. Ren, and J. Sun, “Spatial pyramid pooling in deep convolutional networks for visual recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1904–1916, 2015.
13. S. Ren, K. He, R. Girshick, J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” in Advances in Neural Information Processing Systems 28, D. D. L. M. S. R. G. C. Cortes, N. D. Lawrence, Ed. Curran Associates, Inc., 2015.
14. A. Radhakrishna, H. Sheila, E. Francisco, S. Sabine, “Frequency-tuned salient region detection,” IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009), pp. 1597–1604, 2009.
15. Z. Zhong, L. Jin, S. Zhang, Z. Feng, “Deeptext: A unified framework for text proposal generation and text detection in natural images,” CoRR, vol. abs/1605.07314, 2016. [Online]. Available: http://arxiv.org/abs/1605.07314.