Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)

doi:10.3233/SHTI240603

Integrated and Interoperable Platform for Detecting Masses on Mammograms

Edouard CHATZOPOULOS[a] and Sébastien JODOGNE[a,1]
[a] *ICTEAM, UCLouvain, Belgium*
ORCiD ID: Edouard Chatzopoulos https://orcid.org/0000-0001-5848-7999,
Sébastien Jodogne https://orcid.org/0000-0001-6685-7398

Abstract. The screening and diagnosis of breast cancer is a major public health issue. Although deep learning models are proving highly effective in breast imaging, these models are not yet readily accessible to a wide audience. In order to promote the widespread dissemination of such models, this article introduces a free and open-source, integrated platform for the automated detection of masses on mammograms. A state-of-the-art RetinaNet model is trained on this task and the results of the inference are encoded using the DICOM-SR interoperable format. These contributions present a significant step towards overcoming the accessibility gap in deep learning for breast imaging.

Keywords. Breast imaging, deep learning, interoperability, DICOM, open source

1. Introduction

World Health Organization (WHO) identifies breast cancer as the most common cancer in women, with over 2.3 million new cases diagnosed in 2020 [1]. Breast cancer is therefore a major public health issue. Both the screening and diagnosis of breast cancer rely on mammography. The use of artificial intelligence (AI) in breast imaging, and particularly deep learning, is currently attracting major interest. For example, a randomized clinical trial in Sweden has recently demonstrated that an AI algorithm was able to replace one of the two radiologists traditionally required as part of the second-reading procedure, without deteriorating the cancer detection rate [2]. Nevertheless, access to high-performance deep learning models remains mainly limited to physicians working in clinical departments that have purchased an institutional license. Although there is a growing trend to make more deep learning models available on an open-access basis, their quality still varies. Moreover, existing open models for medicine are in general not integrated into clinically oriented software, which severely limits their usefulness for nursing staff. Consequently, barriers to entry in deep learning for breast imaging persist, including challenges in the training of physicians, in the utilization of AI by researchers, and in the dissemination of AI models within emerging economies.

In this paper, we release a state-of-the-art deep learning model to detect masses on mammograms. The inference results of this model are encoded according to the

[1] Corresponding Author: Sébastien Jodogne, Computer Science and Engineering Department (INGI), Institute of Information and Communication Technologies, Electronics and Applied Mathematics (ICTEAM), Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium; E-mail: sebastien.jodogne@uclouvain.be.

DICOM (Digital Imaging and Communication in Medicine) standard as structured report instances (DICOM-SR), enabling their display by a wide range of professional visualization tools. In addition, we introduce a Web viewer capable of rendering such DICOM-SR instances. These contributions are integrated into the free and open-source Orthanc ecosystem [3], enabling the AI model to be easily run on any computer. To the best of our knowledge, this is the first open-source breast mass detection platform integrated in a Web environment and compliant with the DICOM-SR format.

2. Methods

2.1. Deep learning model

The detection model used in this work is a RetinaNet [4] based on ResNet-50 backbone [5] with a feature pyramid extractor (FPN) [6]. The backbone is used to derive three feature maps at various scales and levels of abstraction. These feature maps are then fed to the classification and regression heads of the neural network, respectively to identify the masses and to predict their dimensions. The output of the model is a set of bounding boxes around the tumors together with a set of confidence scores.

The dataset used to train this detection network is the "Curated Breast Imaging Subset" (CBIS-DDSM) dataset [7]. This dataset consists of 6775 scanned film mammograms containing both benign and malignant cases from 1566 patients. Only the 1696 mammograms containing masses are used to train and evaluate the model. An 80/10/10 train/validation/test split of the patients (not of the images) is performed to avoid data contamination from different views of the same tumor. Before being fed to the model, each image is resized so that its longest side has a length of 2048 pixels, while preserving the original aspect ratio. Resizing is done using bilinear interpolation.

The model is trained using the four-phase strategy that is presented in Table 1.

Table 1. The set of hyper-parameters used for the different training phases. These hyper-parameters were optimized using the validation set. Phases 1A and 1B consist in the training of the patch classifier alone. Phases 2A and 2B correspond to the training of the RetinaNet network.

Training phase	1A	1B	2A	2B
Task	Patch binary classification	Patch binary classification	Mass detection	Mass detection
Loss function	Binary cross entropy loss	Binary cross entropy loss	Focal loss	Focal loss
Trained layers	Final fully-connected layer	Complete ResNet-50	Heads of RetinaNet	Complete RetinaNet
Batch size	32	32	1	1
Learning rate	1e-4	1e-4	1e-4	1e-5
Weight decay	5e-4	5e-4	1e-4	1e-4
Optimizer	RAdam	RAdam	SGD	SGD
Learning rate scheduler	Cosine annealing	Cosine annealing	Cosine annealing	Cosine annealing
Epochs	100	100	100	100
Batch normalization state	Unfrozen	Unfrozen	Frozen	Frozen

First, the classification head of the ResNet-50 backbone pre-trained on ImageNet is fine-tuned to distinguish between 256x256 patches of mammograms that contain or that do not contain a mass, while keeping the weights of the convolutional layers frozen (phase 1A). After 100 epochs, the other layers are unfrozen and fine-tuned for a set of 100 epochs (phase 1B). The resulting patch classifier is then extended with a FPN and integrated as the backbone of a RetinaNet model to detect masses at different scales. The complete network is subsequently trained on the detection task for 100 epochs on full mammograms with the backbone frozen (phase 2A), and finally fine-tuned for an additional set of 100 epochs with the backbone unfrozen (phase 2B). Data augmentation was applied, including random affine transformations, addition of noise, as well as changes in contrast and brightness.

2.2. Interoperability of inference results

The seamless integration of AI algorithms into clinical workflows is a key topic to enable decision-making processes. The "AI Results" profile from IHE (Integrating the Healthcare Enterprise) is a recent specification that defines how AI systems can communicate and integrate their results with existing healthcare information systems. This profile notably indicates that qualitative findings corresponding to planar regions of interest (ROI) must be encoded as DICOM-SR instances, by combining the "Measurement Report" root template (known as TID 1500) with its subordinate "Planar ROI Measurement and Qualitative Evaluations" template (known as TID 1410).

The implementation of our platform is in line with this specification. Given a DICOM mammogram, the inference of the RetinaNet model described in Section 2.1 is executed using the PyTorch library. This inference is carried on inside a Python virtual environment to isolate project dependencies and to ensure consistency across different execution environments. The planar bounding boxes around the detected masses together with their detection score are then encoded as a DICOM structured report using the highdicom Python library [8]. Despite the capabilities of various professional viewers to display DICOM-SR instances, there remains a scarcity of free and open-source solutions for this purpose. To address this gap, we have extended the Stone Web viewer, a fully functional Web viewer for medical imaging developed entirely using WebAssembly [9], to include rendering capabilities for DICOM-SR instances.

3. Results

The training of the model took 30 hours on a server equipped with an Intel W9-3495W CPU and two NVIDIA RTX A4500 GPU. Figure 1 shows the Free-response Receiver Operating Characteristic Curve (FROC) of the model over the test set. This metric is taken at an intersection over union (IoU) of at least 0.2, which corresponds to the conventions of the literature for breast tumor detection tasks [10]. This figure illustrates the True Positive Rate (TPR) of the model with respect to its False Positive Rate (FPR).

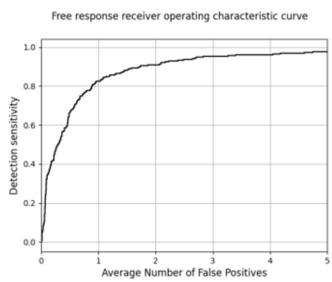

Figure 1. FROC curve of the model on the test set of the CBIS-DDSM dataset at an IoU of at least 0.2.

A comparison of the proposed model to other models that were trained and evaluated in a similar experimental setup is reported in Table 2, indicating that our model offers detection performance in line with the state-of-the-art on the CBIS-DDSM dataset.

Table 2. Comparison of the TPR at similar FPR between models that were trained and evaluated only on the CBIS-DDSM dataset. The FPR of our model is chosen to be smaller or equal to the FPR chosen in the other works to give a fair comparison with the other detection models.

	TPR at FPR
Our model	0.9283 at 2.22
Peng et al. [11]	0.9345 at 2.28
Xiang Yu et al. [12]	0.8729 at 2.86

Our deep learning model, the interoperable encoding of inference results, and the Web viewer have been integrated together as a plugin for Orthanc, an ecosystem for the storage and management of DICOM images [3]. This is made possible by the ability of Orthanc to call Python code. The source code of the plugin is released as free and open-source software (https://github.com/jodogne/orthanc-mammography). Figure 2 depicts the resulting user interface, which proposes a Web environment for the detection of masses on mammograms. The execution of the inference takes about 7 seconds on a standard laptop computer equipped with an Intel i7-1165G7 CPU.

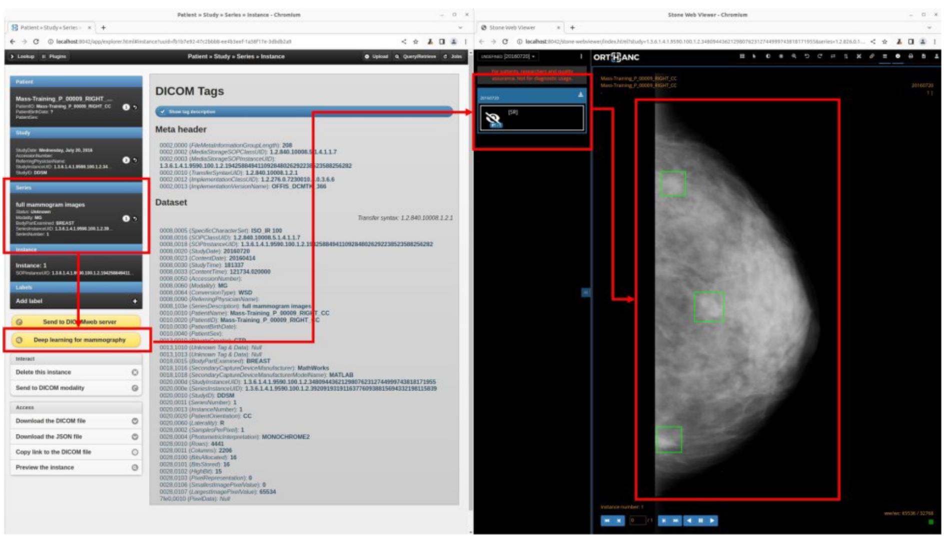

Figure 2. Mammography inference integrated within the Orthanc ecosystem. *Left:* The user selects a DICOM instance in the built-in user interface of Orthanc, and executes the inference by clicking on a button. *Right:* This operation generates a standard DICOM-SR instance that can be displayed using the Stone Web viewer.

4. Discussion

Our platform for the detection of masses in mammograms provides an effective solution for exploring the possibilities of deep learning for breast imaging without requiring the installation of a dedicated computing infrastructure. Non-technical staff could therefore make perfect use of this platform as part of their continuing education on their personal equipment. Furthermore, as our inference platform is distributed as free and open-source software, it is readily accessible to a wide audience. It could thus serve as a pedagogical tool to enhance the teaching of breast imaging. Because our model performs in line with the state-of-the-art on the CBIS-DDSM dataset, the platform could also be used as a reference implementation to support clinical research and to share technical knowledge with emerging economies. Finally, note that Orthanc provides advanced DICOM scripting capabilities, enabling the automated screening of mammograms: The deep learning model could be executed on every incoming DICOM instance, thereby providing the nursing staff with a continuously updated dashboard.

5. Conclusions

This paper introduces an open platform for the detection of masses on mammograms. This platform embeds a state-of-the-art deep learning model and is released as free and open-source software. Thanks to the fact that this solution is integrated according to the DICOM standard, it can easily be deployed in any clinical department, provided that the local regulations such as MDR in Europe are fulfilled. In the future, we plan to enhance the model to detect not only masses but also microcalcifications. The capabilities of the models will be evaluated statistically in the context of breast cancer screening. The inference code will be ported from Python to C++ or Java to further facilitate the deployment of the solution. More generally, it is planned to create a library of open, interoperable deep learning models that go beyond breast imaging.

References

[1] Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, et al. Current and future burden of breast cancer: Global statistics for 2020 and 2040. The Breast. 2022;66:15–23.

[2] Lång K, Josefsson V, Larsson AM, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI). Lancet Oncol. 2023 Aug;24(8):936–44.

[3] Jodogne S. The Orthanc Ecosystem for Medical Imaging. J Digit Imaging. 2018 May;31(3):341–52.

[4] Lin T, Goyal P, Girshick R, He K, Dollar P. Focal Loss for Dense Object Detection. In: Proc of the 2017 IEEE International Conference on Computer Vision (ICCV). Los Alamitos, CA, USA: IEEE Computer Society; 2017. p. 2999–3007.

[5] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2016. p. 770–8.

[6] Lin TY, Dollar P, Girshick R, He K, Hariharan B, Belongie S. Feature Pyramid Networks for Object Detection. In: Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2017. p. 936–44.

[7] Lee R, Gimenez F, Hoogi A, Miyake K, Gorovoy M, Rubin D. A curated mammography data set for use in computer-aided detection and diagnosis research. Sci Data. 2017 Dec;4(1).

[8] Bridge CP, Gorman C, Pieper S, Doyle SW, Lennerz JK, et al. Highdicom: a Python Library for Standardized Encoding of Image Annotations and Machine Learning Model Outputs in Pathology and Radiology. J Digit Imaging. 2022 Aug;35(6):1719–37.

[9] Jodogne S. On the Use of WebAssembly for Rendering and Segmenting Medical Images. Comms Comput Inf Sci. 2023;1814(1):393–414.

[10] Dhungel N, Carneiro G, Bradley AP. A deep learning approach for the analysis of masses in mammograms with minimal user intervention. Med Image Anal. 2017;37:114–28.

[11] Peng J, Bao C, Hu C, Wang X, Jian W, Liu W. Automated mammographic mass detection using deformable convolution and multiscale features. Med Biol Eng Comput. 2020 Jul 1;58(7).

[12] Yu X, Wang SH, Zhang YD. Multiple-level thresholding for breast mass detection. J King Saud Univ. 2023;35(1):115–30.