Supervised Pectoral Muscle Removal in Mammography Images

Parvaneh Aliniya(✉), Mircea Nicolescu, Monica Nicolescu, and George Bebis

University of Nevada, Reno, NV 89557, USA
paliniya@unr.edu

Abstract. In this paper, we provide the segmentation masks of the pectoral muscle for INbreast, MIAS, and a CBIS-DDSM subset datasets, which will enable the development of supervised methods and the utilization of deep learning for pectoral muscle removal from mammography images. We trained AU-Net separately on the INbreast and CBIS-DDSM subset for the segmentation of the pectoral muscle. We used cross-dataset testing to evaluate the performance of the models on an unseen dataset. The experimental results show that cross-dataset testing achieves a comparable performance to the same-dataset experiments. In addition, the models were tested on the entire MIAS dataset, and they outperformed previous methods. The segmentation masks are available at https://github.com/Parvaneh-Aliniya/pectoral_muscle_groundtruth_segmentation.

Keywords: Breast Cancer · Pectoral Muscle · Mammography

1 Introduction

Breast cancer is one of the main cancer types in the female population, with a high mortality rate. Mammograms are images taken from two different views of breast. These views are Craniocaudal (CC) and mediolateral oblique (MLO). CC is the view in which the breast is compressed horizontally, and in the MLO view, the compression is diagonal. Mammography images are the most commonly used tool for breast cancer screening due to their availability and lower cost. Therefore, the development of automated cancer detection methods for these images is of high importance due to the benefits they bring to the patients by increasing survival chances when detecting the abnormalities accurately in early stages [1].

In this paper, we target the task of removing the pectoral muscle in mammography images, which is of high importance for automated density estimation. Due to the fact that normally the segmentation mask of the pectoral muscle is not available in the publicly available datasets or in the examinations for breast cancer screening in current practice, most of the pectoral removal approaches use traditional machine learning approaches. Since the location, shape, density, and position of the muscle vary between the images [2], the performance of these methods is limited. Ideally, utilizing deep learning-based approaches would help mitigate these hurdles to a great extent if annotations for the pectoral muscle were available.

J. Finkelstein et al. (Eds.): AIME 2024, LNAI 14845, pp. 126–130, 2024.
https://doi.org/10.1007/978-3-031-66535-6_14

We argue that providing annotations even for several datasets will enable researchers to train supervised pectoral muscle removal and utilize them for new unseen datasets with high performance. Hence, we provide pectoral muscle segmentation for three benchmark datasets, INbreast [3], MIAS [4], and a widely used subset of CBIS-DDSM [5]. We separately trained AU-Net [6], which is a widely used segmentation method for mammography images, on the datasets. To validate the proposed method, same-dataset and cross-dataset tests were used. In same-dataset tests, train and test sets belong to one of the datasets. In the cross-dataset experiments, train and test sets are from two different datasets. The models achieve high accuracy (99.24 ± 0.16 on average) for both same-dataset and cross-dataset experiments.

The contributions of this paper are three-fold: (1) Generating the segmentation masks of pectoral muscle for INbreast, MIAS, and CBIS-DDSM subset datasets. (2) Training pectoral muscle removal models using the AU-Net architectures separately for INbreast and CBIS-DDSM datasets. (3) Evaluating the models by same-dataset and cross-dataset testing to measure the generalizability of the supervisely trained models on the same and new datasets.

2 Related Work

The segmentation of the pectoral muscle is not available in mammography datasets; hence, most of the proposed methods are traditional machine learning-based methods. These methods, in general, aim to use the appearance of the muscle and its location to detect and eliminate it. According to a recent study [1], thresholding [7] and region growing [8] are among the widely used approaches.

The general idea for thresholding is to use the observation that the brightness of the pectoral muscle is generally higher than the neighboring regions; therefore, by eliminating pixels lower than a certain threshold, the region for the pectoral muscle will the extracted. This idea, coupled with the utilization of the orientation of the breast (whether the breast is on the left or right side of the image) and the generic shape of the muscle, has been used in the literature [9].

In general, region-growing-based methods start with initial seeds, and then, according to certain similarity metrics, they continue adding a new neighboring pixel to a region until a termination criterion is met [9, 10]. Aside from the previous methods that aim to use region growing and thresholding, graph-cut [11], Hough Transform [12], line estimation, polynomial fitting and curve estimation [13], k-means [14], active contours [15] and contour growing [16] are also used in several methods.

3 Proposed Method

3.1 Ground Truth Generation for Pectoral Muscle

For the segmentation of pectoral muscle, LabelMe [17] was used, in which polygons were fitted to the pectoral muscle for the MLO images in INbreast, a subset of CBIS-DDSM and MIAS datasets. For images with higher density or lower visibility of the boundaries of the pectoral muscle, the portion of the muscle that was clearly distinguishable from the breast tissues was selected. The segmentation masks are generated in JSON and image formats with two classes, the pectoral muscle, and background (the rest of breast tissues and image background).

3.2 Pectoral Muscle Segmentation

The main motivation for this study is to provide the segmentation of the pectoral muscle for several datasets, which enables the training of pectoral muscle removal that could also be applied to new unseen datasets. The main use-case of these segmentation masks will be in the removal of pectoral muscle from images in the preprocessing step for tasks such as the classification of images (for instance, benign/malignant), segmentation of the masses and other abnormalities, and density estimation. To use the segmentation masks for a new dataset; first, a segmentation model should be trained using the provided muscle segmentation masks, and then the model could be used for the segmentation of the pectoral muscles in the new dataset.

Given the segmentation masks of the pectoral muscles, we propose to use deep learning-based methods for pectoral muscle segmentation. To this end, we selected AU-Net which is a state-of-the-art method for segmentation in mammography images. AU-Net [6] is an improved version of the U-Net [18] in which the encoding and decoding paths are not symmetrical. ResUnit and the basic decoder proposed in the AU-Net have been used for the encoder and the decoder. The details of the novel idea of AU-Net, the Attention guided Up-sampling Block (AU Block), are presented in the AU-Net approach [6]. The binary cross-entropy loss function was used in the proposed method.

After training on the INbreast and CBIS-DDSM subsets separately, the models have been used for the cross-dataset tests to evaluate their performance on unseen datasets.

4 Experimental Results

4.1 Datasets and Preprocessing

The pectoral muscle masks for all of the MLO view images in the INbreast dataset have been provided in this study and used for the experiments. A 5-fold cross-validation is used for the INbreast dataset. CBIS-DDSM is an enhanced subset of the DDSM dataset. We provided the pectoral muscle segmentation masks for all the 447 MLO view images in a commonly used subset of CBIS-DDSM. In the experiments, the original split for testing and training is used. MIAS dataset contains only MLO view images. We also included the segmentation masks for all the images in the MIAS dataset (unless the muscle was not visible in the images). We used MIAS for cross-dataset testing using models trained on INbreast, CBIS-DDSM subset, and a combination of datasets. For all of the datasets, cropping, padding, resizing, and artifact removal have been performed as needed. The models were trained with a learning rate of 0.0001 for 100 epochs.

4.2 Results for INbreast, CBIS-DDSM Subset, and MIAS Datasets

The experimental results and comparison with state-of-the-art methods are presented in Table 1. The Dice Similarity Coefficient (DSC), Sensitivity, and Accuracy have been used as evaluation metrics in the experiments. The following format is used for the names of the experiments: "train dataset name - test dataset name". We used 'CBIS' instead of 'CBIS-DDSM subset' in the table for convenience. As shown, the proposed method has accurate prediction for same and cross-dataset tests and also outperforms the previous methods on the MIAS dataset. It should be noted that the entire MIAS dataset was used only in test time for all the experiments with the MIAS dataset.

Table 1. Results for pectoral muscle segmentation for same and cross-dataset experiments. In the name of the experiments, the first term is the training dataset, and the second is the test dataset.

Train-Test Pair	DSC↑	Sensitivity↑	Accuracy↑
[16]	–	–	98.1
[13]	–	–	96.81
[7]	–	–	98
[10]	–	–	97.8
CBIS-CBIS	96.61	98.07	99.50
INbreast-INbreast	95.09	95.54	99.55
INbreast-CBIS	91.87	97.09	99.00
CBIS-INbreast	86.89	78.87	98.93
CBIS-MIAS	95.11	95.04	99.53
INbreast-MIAS	90.64	95.44	99.03
Combined-MIAS	95.73	95.55	99.59

5 Conclusion

In this study, we provide the segmentation masks for the INbreast, MIAS, and CBIS-DDSM subset datasets. We also trained AU-Net separately on the INbreast, CBIS-DDSM subset, and a combination of both datasets. We tested both models on the entire MIAS dataset, resulting in accuracies of 99.03%, 99.53%, and 99.59% for models trained on the INbreast and CBIS-DDSM subset and the combination of both datasets, respectively.

References

1. Moghbel, M., Ooi, C.Y., Ismail, N., Hau, Y.W., Memari, N.: A review of breast boundary and pectoral muscle segmentation methods in computer-aided detection/diagnosis of breast mammography. Artif. Intell. Rev. **53**, 1873–1918 (2020)
2. Haris, U., Kabeer, V.: Automated breast pectoral muscle segmentation in digital mammograms using holistic regularization method. J. Clin. Otorhinolaryngol. Head Neck Surg.
3. Moreira, I.C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M.J., Cardoso, J.S.: Inbreast: toward a full-field digital mammographic database. Acad. Radiol. **19**, 236–248 (2012)

4. Suckling, J., Parker, J., Dance, D., et al.: Mammographic Image Analysis Society (MIAS) Database v1.21. Apollo - University of Cambridge Repository (2015)
5. Lee, R.S., Gimenez, F., Hoogi, A., Miyake, K.K., Gorovoy, M., Rubin, D.L.: A curated mammography data set for use in computer-aided detection and diagnosis research. Sci. Data (2017)
6. Sun, H., et al.: AUNet: attention-guided dense-upsampling networks for breast mass segmentation in whole mammograms. Phys. Med. Biol. **65**(5), 055005 (2020)
7. Czaplicka, K., Włodarczyk, J.: Automatic breast-line and pectoral muscle segmentation. Schedae Informaticae **20** (2011)
8. Ergin, S., Esener, İ.I., Yüksel, T.: A genuine GLCM-based feature extraction for breast tissue classification on mammograms. Int. J. Intell. Syst. Appl. Eng. **4**, 124–129 (2016)
9. Subashini, T.S., Ramalingam, V., Palanivel, S.: Pectoral muscle removal and detection of masses in digital mammogram using CCL. Int. J. Comput. Appl. 71–76 (2010)
10. Chen, Z., Zwiggelaar, R.: A combined method for automatic identification of the breast boundary in mammograms. In: 2012 5th International Conference on BioMedical Engineering and Informatics (2012)
11. Camilus, K.S., Govindan, V.K., Sathidevi, P.S.: Computer-aided identification of the pectoral muscle in digitized mammograms. J. Digit. Imaging **23**, 562–580 (2010)
12. Xu, W., Li, L., Liu, W.: A novel pectoral muscle segmentation algorithm based on polyline fitting and elastic thread approaching. In: 2007 1st International Conference on Bioinformatics and Biomedical Engineering, pp. 837–840. IEEE (2007)
13. Shen, R., Yan, K., Xiao, F., Chang, J., Jiang, C., Zhou, K.: Automatic pectoral muscle region segmentation in mammograms using genetic algorithm and morphological selection. J. Digit. Imaging **31**, 680–691 (2018)
14. Slavković, I.M., Gavrovska, A., Milivojević, M., Reljin, I., Reljin, B.: Breast region segmentation and pectoral muscle removal in mammograms. Telfor J. **8**(1), 50–55 (2016)
15. Yin, K., Yan, S., Song, C., Zheng, B.: A robust method for segmenting pectoral muscle in mediolateral oblique (MLO) mammograms. Int. J. Comput. Assist. Radiol. Surg. **14**, 237–248 (2019)
16. Rampun, A., Morrow, P.J., Scotney, B.W., Winder, J.: Fully automated breast boundary and pectoral muscle segmentation in mammograms. Artif. Intell. Med. **79**, 28–41 (2017)
17. LabelMe. https://github.com/labelmeai/labelme
18. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation.In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, pp. 234–241. Springer (2015)