

Deep Learning Based Pectoral Muscle Segmentation on MIAS Mammograms

Young Jae Kim

Gachon University - Medical Campus

Eun Young Yoo

Gachon University Gil Medical Center

Kwang Gi Kim (✉ kimkg@gachon.ac.kr)

Department of Biomedical Engineering, Gachon University, Incheon, 21936, Republic of Korea.

<https://orcid.org/0000-0001-9714-6038>

Research article

Keywords: Deep learning, Mammograms, Pectoral muscle, Segmentation, Computer aided diagnosis

Posted Date: October 19th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-92779/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: The purpose of this study was to propose a deep learning-based method for automated detection of the pectoral muscle, in order to reduce misdetection in a computer-aided diagnosis (CAD) system for diagnosing breast cancer in mammography. This study also aimed to assess the performance of the deep learning method for pectoral muscle detection by comparing it to an image processing-based method using the random sample consensus (RANSAC) algorithm.

Methods: Using the 322 images in the Mammographic Image Analysis Society (MIAS) database, the pectoral muscle detection model was trained with the U-Net architecture. Of the total data, 80% was allocated as training data and 20% was allocated as test data, and the performance of the deep learning model was tested by 5-fold cross validation.

Results: The image processing-based method for pectoral muscle detection using RANSAC showed 92% detection accuracy. Using the 5-fold cross validation, the deep learning-based method showed a mean sensitivity of 95.55%, mean specificity of 99.88%, mean accuracy of 99.67%, and mean Dice similarity coefficient (DSC) of 95.88%.

Conclusions: The proposed deep learning-based method of pectoral muscle detection performed better than an existing image processing-based method. In the future, by collecting data from various medical institutions and devices to further train the model and improve its reliability, we expect that this model could greatly reduce misdetection rates by CAD systems for breast cancer diagnosis.

Background

Among breast imaging methods, mammography is widely used in screening for breast cancer [1]. However, the results can be difficult to interpret for dense breasts, which leads to a high risk of misdiagnosis [2]. For this reason, there has been ongoing research on computer-aided diagnosis (CAD) systems, in order to reduce the number of misdiagnoses and to improve the accuracy of diagnosis by radiologists using mammography [3–5]. CAD systems use computer algorithms to enable objective and accurate detection of lesions that are difficult to distinguish with the naked eye. However, accurate lesion detection by these CAD systems can be negatively affected for various reasons. In particular, in CAD systems used to diagnose breast cancer, the pectoral muscle shows a similar pixel intensity to that of lesions of the right mediolateral oblique (RMLO) and left mediolateral oblique (LMLO) muscles, which can cause misdetection [6]. To prevent this, a separate pectoral muscle detection algorithm is required.

In 2016, we developed an image processing-based automated pectoral muscle detection algorithm using the random sample consensus (RANSAC) algorithm on images from the Mammographic Image Analysis Society (MIAS) database [7]. This algorithm had a detection accuracy of 92.2%, which was higher than that found in other studies using the MIAS database [8–12]. Nevertheless, detection accuracy was poor in some images due to the complex shape of the pectoral muscle, and so the algorithm needed to be improved.

Recent advances in hardware have created a favorable environment for deep learning techniques, which have been applied in various fields. Convolutional neural networks (CNNs) especially, as one type of deep learning technique, have been used with outstanding results in various imaging fields [13, 14]. In medical imaging as well, numerous studies using CNNs have reported better performance than conventional image processing techniques [15–17]. Thus, in this study, we aimed to use deep learning in pectoral muscle detection, to improve upon the previously encountered problems for the complex-shaped pectoral muscle, and to enhance detection accuracy. The deep learning model for pectoral muscle detection was trained using the same MIAS database as before, and the performance of the algorithm was assessed in comparison to the results of the image processing-based method using RANSAC.

Materials And Methods

Data

This study used mammograms from the mini-MIAS database, which is one of several open access databases. The MIAS database consists of scans of 322 mammogram films taken as part of the United Kingdom's national breast cancer screening program [18]. All the images include RMLO or LMLO views, and the image sizes are $1,024 \times 1,024$ pixels. For annotation data to use in learning, regions of interest (ROIs) for the pectoral muscle were drawn directly by a specialist. A binary mask was made, using 0 for the background and 1 for the pectoral muscle, and this was used as the ground truth. Figure 1 shows one of the binary mask images used as the ground truth.

The size of the data set used in this study was not large, and so 5-fold cross validation was used to ensure that the model would be robust in terms of data dependency. For each cross validation, 80% of the overall data (257–258 scans) was used as learning data, and 20% (64–65 scans) was separately constructed and used as test data. Each data point was used exactly once as test data, without duplication.

Development environment

The system for deep learning consisted of 4 NVIDIA TITAN Xp (NVIDIA Corp., Santa Clara, CA, USA) graphics processing units (GPUs), a Xeon E5-1650 v4 (Intel Corp., Santa Clara, CA, USA) central processing unit (CPU), and 128 GB of random access memory (RAM). Deep learning was conducted using Python 2.7.6 and the Keras 2.1.5 framework with a TensorFlow backend in the Ubuntu 14.04 operating system.

Data augmentation

The images from the MIAS database used in this experiment were insufficient to train the deep learning model. Data augmentation was performed to acquire a sufficient quantity of learning data [19]. An arbitrary combination of flips, rotations, translations, and stacking were used to expand the learning data set 20-fold.

Training the deep learning model

For the CNN in this study, the U-Net model was used. One advantage of the U-Net model is that it has a structure that re-uses encoding and decoding phases via skip connections (Fig. 2); as a result, when images are reconstructed using the network, the original image can be reconstructed without losing even fine details, meaning that the output images have excellent quality [20, 21]. For the learning environment, the batch size was set to 8, and the number of epochs was fixed to 300 based on an Adam optimizer. The learning rate was set to 0.001 up to Epoch 100, 0.0001 from Epoch 100 to 250, and 0.00001 from Epoch 250 to 300.

Results

In this study, deep learning was used to train a model to detect pectoral muscle in MIAS database images. The trained model was applied to separately constructed test data to assess its performance. Figure 3 compares the ground truth data from the test data with the results automatically extracted using the trained model.

We performed 5-fold cross validation to make sure the model was robust in terms of data dependency. For each cross validation, 20% of the total data was used as test data, and each data point was used exactly once as test data, without duplication. For each cross validation, the model was tested using 4 statistical indices: sensitivity, specificity, accuracy, and Dice similarity coefficient (DSC). The extracted results of the deep-learning model were compared pixel-by-pixel with the ground truth data, the true positive (TP), false positive (FP), true negative (TN), and false negative (FN) rates were calculated, and the statistical indices were calculated using the equations below. From the results of the 5 cross validations, the mean sensitivity was 95.55%, the mean specificity was 99.88%, the mean accuracy was 99.67%, and the mean DSC was 95.88% (Table 1).

Table 1
Results of 5-fold cross validation for the deep learning-based pectoral muscle detection method

	Sensitivity	Specificity	Accuracy	DSC
CV1	94.01	99.91	99.57	94.50
CV2	96.40	99.87	99.68	96.78
CV3	95.65	99.90	99.72	96.45
CV4	95.97	99.83	99.66	96.18
CV5	95.73	99.89	99.71	95.49
Total	95.55	99.88	99.67	95.88
CV, cross validation; DCS, Dice similarity coefficient				

The deep learning-based pectoral muscle detection algorithm was assessed using the same method as our previous study on an image processing-based method using the RANSAC algorithm, and the results of the two models were compared. We assessed the differences between the automated detection results of the deep learning model and the manually drawn ground truth data. Concordance $\geq 90\%$ between the deep learning-based automated detection and the manual detection images was defined as “good”, concordance $\geq 50\%$ and $< 90\%$ was defined as “acceptable”, and concordance $< 50\%$ was defined as “unacceptable”. The previous method using the RANSAC algorithm showed 264 “good” results, whereas the deep learning model showed 322 “good” results (Table 2). The FP and FN rates of the previous method were, respectively, $4.51 \pm 6.53\%$ and $5.68 \pm 8.57\%$ (Table 3). In contrast, the FP and FN rates of the deep learning method were, respectively, $2.88 \pm 6.05\%$ and $4.27 \pm 8.72\%$.

Table 2
Comparison of performance (categorical) between the deep learning-based pectoral muscle detection method and the image processing-based method using RANSAC

	Good	Acceptable	Unacceptable
RANSAC method	264	36	22
Deep learning method	322	0	0
RANSAC, random sample consensus			

Table 3
Comparison of performance (detection accuracy) between the deep learning-based pectoral muscle detection method and the image processing-based method using RANSAC

Category	RANSAC method (%)	Deep learning method (%)
FP	4.51 ± 6.53	2.88 ± 6.05
FN	5.68 ± 8.57	4.27 ± 8.72
FP $< 5\%$ and FN $< 5\%$	56.5	71.0
$5\% < \text{FP} < 15\%$, $5\% < \text{FN} < 15\%$	31.5	20.7
$15\% < \text{FP}$, $15\% < \text{FN}$	12.0	8.3
FN, false negative; FP, false positive; RANSAC, random sample consensus		

Discussion

This study proposed a pectoral muscle detection method using deep learning with the MIAS database. Although the proposed method used all the images in the MIAS database, there was too little data to construct a separate validation set. We used 5-fold cross validation to supplement the shortage of

validation data and to minimize the data dependency of the model. The proposed model showed high accuracy, with a mean sensitivity of 95.55% and mean DSC of 95.88%.

We also compared the results of the deep learning-based pectoral muscle detection method with a previous image processing-based pectoral muscle detection method using the RANSAC algorithm. While the RANSAC method showed “unacceptable” results for 22 images, the deep learning method did not show “unacceptable” results for even a single image. Moreover, when the misdetection rate was inspected, the RANSAC algorithm showed FP and FN rates < 5% for only 56.5% of the images, whereas the deep learning algorithm showed a higher proportion of images (71.0%) with FP and FN rates < 5%. These results demonstrate that the deep learning algorithm achieved more accurate and more stable detection results than the RANSAC algorithm.

In image processing, RANSAC is an algorithm for approximation. Approximation cannot guarantee the detection of an accurate pectoral muscle region. Although there have been attempts to approximate the area of a curved pectoral muscle using nonlinear RANSAC, it is quite difficult to accurately approximate the pectoral muscle. This weakness is thought to have led to the 22 “unacceptable” results. In contrast, because deep learning makes an overall judgment about the shape and the relationship attenuation and pixels based on the training images, it requires generalized, diverse training data. In this study, we used the same data set in order to compare the new model objectively with the previous image processing-based pectoral muscle detection algorithm. The data set was somewhat small to train a deep learning model. The fact that we were still able to obtain relatively good results is thought to be because the shape and position of the pectoral muscle was fairly consistent, and did not differ greatly between patients. Although the possibility of overfitting can be suspected, these doubts can be partially resolved based on the results of cross validation.

This study had some limitations. In current clinical settings, digital mammography is used in most instances, but the MIAS database contains data obtained by scanning film mammograms. Therefore, there is no guarantee that a deep learning model trained on MIAS data will show good results for digital mammograms. In order to resolve this issue, it will be necessary to collect more digital mammograms, and to further train a deep learning model based on the collected data. It will also be important to collect data from diverse medical institutions and devices, and to objectively validate the deep learning model through multi-center validation. We expect that these additional studies would further enhance the clinical reliability of the deep learning-based pectoral muscle detection method proposed in this report.

Conclusions

The experiments in this study demonstrated that the deep learning-based pectoral muscle detection method achieved a more accurate and stable detection rate than a previous image processing-based method. We expect that this type of deep learning technique could be useful to overcome the limitations of conventional image processing techniques.

Abbreviations

CAD

Computer Aided Diagnosis; DSC:Dice Similarity Coefficient; RANSAC:random sample consensus; MIAS:Mammographic Image Analysis Society; RMLO:right mediolateral oblique; LMLO:left mediolateral oblique; CNNs:Convolved neural networks; ROIs:regions of interest; GPUs:graphics processing units; CPU:central processing unit; RAM:random access memory; TP:true positive; FP:false positive; TN:true negative; FN:false negative.

Declarations

Acknowledgments

This work was supported by the Gachon University Gil Medical Center (Grant number: 2018-5299, Grant number : FRD2019-11) and the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2019R1C1C1008381). Kwang Gi Kim and Eun Young Yoo equally contributed to this work.

Authors' contributions

YJ Kim Manuscript preparation and editing. Data analysis. Experiment. KG Kim Study idea. Manuscript preparation and editing. EY Yoo Study idea. Data acquisition. Manuscript editing.

Funding

Not applicable.

Availability of data and materials

The datasets analysed during this current study are available in the MIAS database.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that there is no conflict of interest regarding the publication of this article.

References

1. Lee CH, Dershaw DD, Kopans D, Evans P, Monsees B, Monticciolo D, et al. Breast cancer screening with imaging: recommendations from the society of breast imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer. *Journal of the American college of radiology*. 2010;7:18–27.
2. Linver MN. Mammographic Density and the Risk and Detection of Breast Cancer. *Breast Diseases*. 2008;18:364–5.
3. Aroquiaraj IL, Thangavel K. Pectoral muscles suppression in digital mammograms using hybridization of soft computing methods. 2014. arXiv preprint arXiv:1401.0870.
4. Subashini TS, Ramalingam V, Palanivel S. Pectoral muscle removal and detection of masses in digital mammogram using CCL. *International Journal of Computer Applications*. 2010;1:71–6.
5. Ginneken BV, Schaefer-Prokop CM, Prokop M, Computer-aided diagnosis: how to move from the laboratory to the clinic. *Radiology*. 2011;261:719–32.
6. Vaidehi K. Automatic identification and elimination of pectoral muscle in digital mammograms. *International Journal of Computer Applications*. 2013;75:15–8.
7. Yoon WB, Oh JE, Chae EY, Kim HH, Lee SY, Kim KG. Automatic detection of pectoral muscle region for computer-aided diagnosis using MIAS mammograms. *BioMed Research International*. 2016; <https://doi.org/10.1155/2016/5967580>
8. Alam N, Islam MJ. Pectoral muscle elimination on mammogram using K-means clustering approach. *International Journal of Computer Vision & Signal Processing*. 2014;4:11–21.
9. Mustra M, Grgic M. Robust automatic breast and pectoral muscle segmentation from scanned mammograms. *Signal Processing*. 2013;93:2817–27.
10. Molinara M, Marrocco C, Tortorella F. Automatic segmentation of the pectoral muscle in mediolateral oblique mammograms. in *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*. 2013; DOI: 10.1109/CBMS.2013.6627852
11. Kwok SM, Chandrasekhar R, Attikiouzel Y. Automatic pectoral muscle segmentation on mammograms by straight line estimation and cliff detection. in *The Seventh Australian and New Zealand Intelligent Information Systems Conference*. 2001; DOI: 10.1109/ANZIIS.2001.974051
12. Raba D, Oliver A, Martí J, Peracaula M, Espunya J. Breast segmentation with pectoral muscle suppression on digital mammograms. *Lecture Notes in Computer Science*. 2005;3523:471–8.
13. Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*. 2012;1–9.
14. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*. 2016;35:1285–98.
15. Suzuki K. Overview of deep learning in medical imaging. *Radiological Physics and Technology*. 2017;10:257–73.

16. Cheng JZ, Ni D, Chou YH, Qin J, Tiu CM, Chang YC, et al. Computer-aided diagnosis with deep learning architecture: applications to breast lesions in us images and pulmonary nodules in CT scans. *Scientific Reports*. 2016;6:1–13.
17. Hua KL, Hsu CH, Hidayati SC, Cheng WH, Chen YJ. Computer-aided classification of lung nodules on computed tomography images via deep learning technique. *OncoTargets and Therapy*, 2015;8:2015–22.
18. Matheus BRN, Schiabel H. Online mammographic images database for development and comparison of CAD schemes. *Journal of Digital Imaging*. 2011;24:500–6.
19. Perez L, Wang J. The Effectiveness of data augmentation in image classification using deep learning. 2017. arXiv preprint arXiv:1712.04621.
20. Norman B, Pedoia V, Majumdar S. Use of 2D U-Net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry. 2018;288:172322.
21. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells W, Frangi A. *International Conference on Medical Image Computing and Computer-assisted Intervention*. Cham: Springer; 2015. pp. 234–241.

Figures

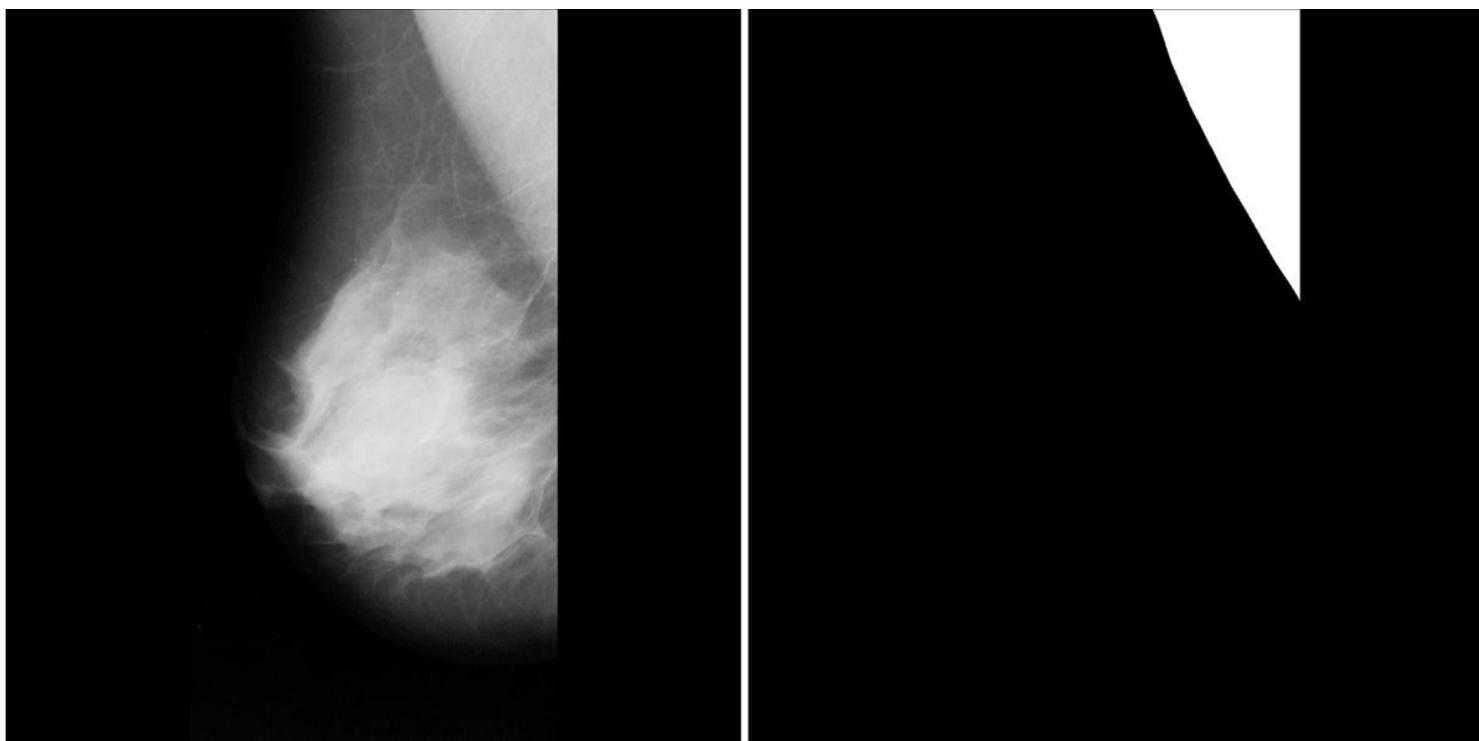


Figure 1

Examples of an image from the MIAS database (left) and annotation data made directly by a specialist (right)

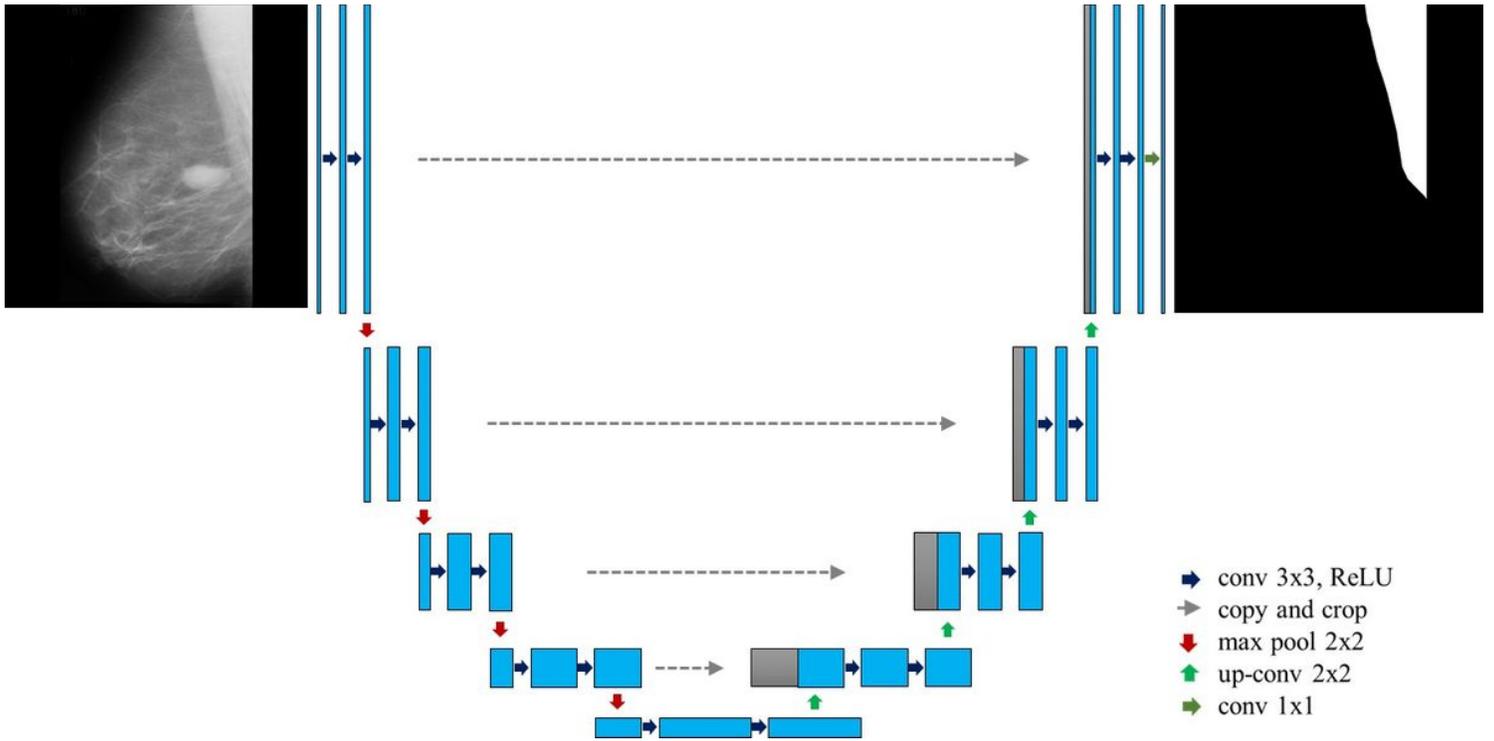


Figure 2

The U-Net architecture consisting of encoding and decoding units

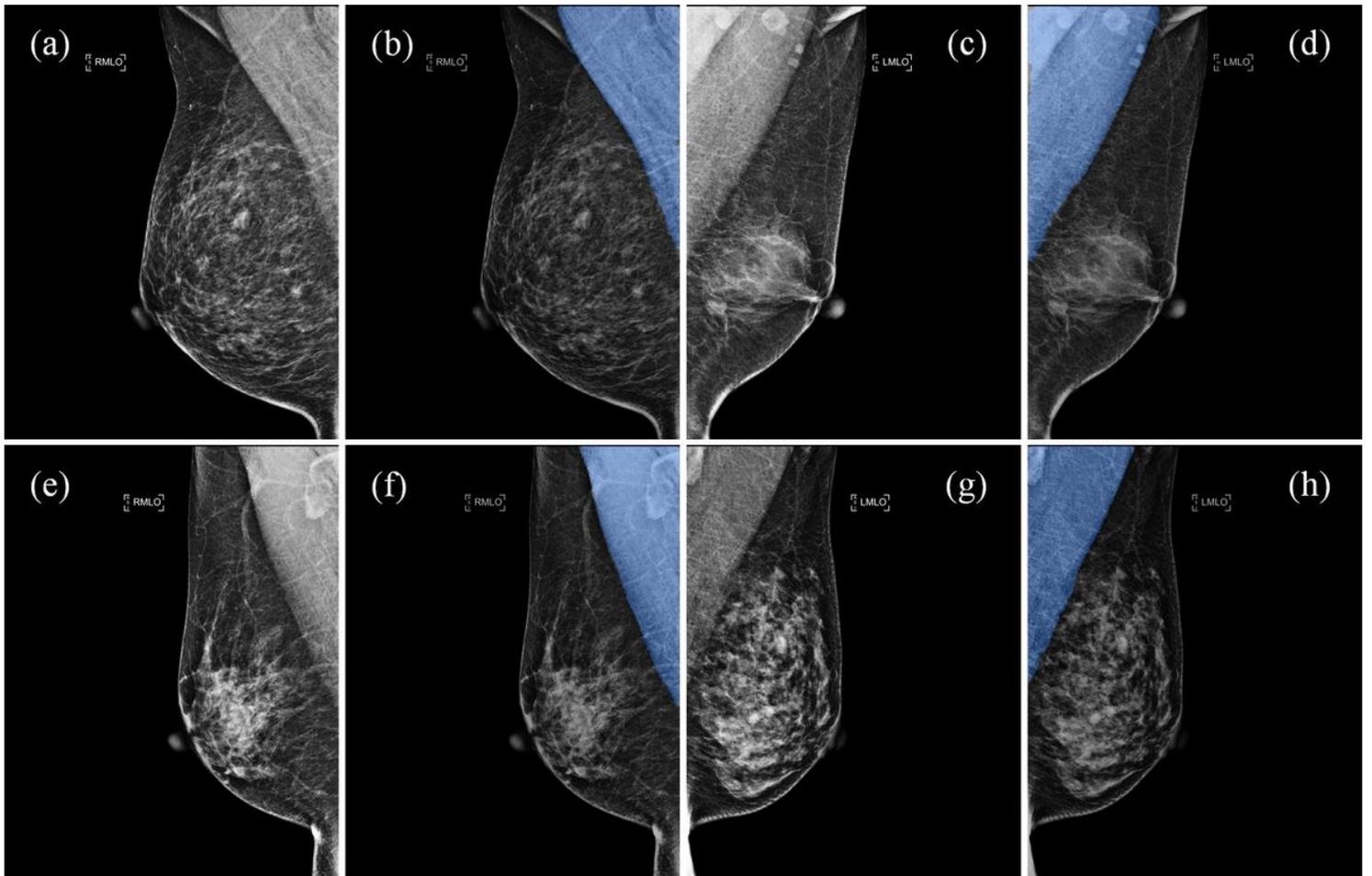


Figure 3

Examples of pectoral muscle detection results from the test data: (a)(c)(e)(g) original image, (b)(d)(f)(h) Deep learning-based automated detection results