INTERNATIONAL JOURNAL OF ELECTRONICS AND COMPUTER APPLICATIONS

ORIGINAL ARTICLE

Article access online

OPEN ACCESS

Received: 11.06.2024

Accepted: 20.10.2024

Published: 04.12.2024

Citation: Patil RA, Dixit VV. (2024). Deep Learning for Improved Breast Cancer Detection: ResNet-50 vs VGG16. International Journal of Electronics and Computer Applications. 1(2): 26-31. https://doi.org/10.70968/ijeaca.v1i2.1

* **Corresponding author**.

rupalipatil027@gmail.com

Funding: None

Competing Interests: None

ISSN
Print: XXXX-XXXX
Electronic: 3048-8257

Deep Learning for Improved Breast Cancer Detection: ResNet-50 vs VGG16

Rupali A Patil[1]*, V V Dixit[2]

1 Ph.D. Research Scholar, G H Raisoni College of Engineering and Management, Pune, Maharashtra, India
2 Principal/Director, RMDSTIC Warje, Pune, Maharashtra, India

Abstract

Breast cancer continues to pose a major global health challenge, underscoring the need for advanced detection and classification methods for mammograms. This study examines the effectiveness of ResNet-50 and VGG16 models in detecting and classifying multiview mammograms. Early and accurate detection of breast cancer is essential for improving patient outcomes and reducing mortality rates. Our approach began with the preparation of mammography images using various image processing techniques, including transfer learning and median filtering. The processed images were then used to train ResNet-50 and VGG16 models for detection and classification tasks. Our experiments demonstrated impressive performance, achieving an accuracy of 96% and an F1 score of 94.66% on the Digital Database for Screening Mammography (DDSM) datasets. These results underscore the potential of deep learning models, particularly ResNet-50, in effectively detecting and classifying multiview mammograms.

Keywords: Breast cancer detection; Deep learning models; ResNet-50; VGG16; Mammogram classification; Multiview analysis; Clinical implications

Introduction

Breast cancer continues to pose a significant global health challenge, carrying profound implications for both illness and death rates. A thorough grasp of cancer data is essential for directing healthcare efforts and distributing resources effectively. Siegel et al. furnished an extensive analysis of cancer statistics for that year, shedding light on the frequency, occurrence, and fatality rates linked to breast cancer. Torre et al. reiterated the substantial worldwide toll of breast cancer, emphasizing the urgency for efficient screening and diagnostic measures to alleviate its consequences.

Literature Review

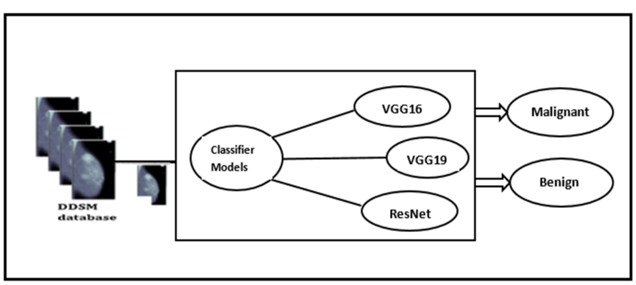
Barbieri and Strauss delved into reproductive endocrinology, providing valuable insights into physiological and clinical aspects relevant to breast cancer. Ragab et al. were trailblazers in employing deep convolutional neural networks (CNNs) and support vector machines (SVMs) for breast cancer detection. Their work showcased the promise of deep learning in enhancing diagnostic accuracy, marking a significant advancement in the field. Li et al. proposed a deep learning-based system for classifying mammogram images into benign and malignant categories, showcasing the effectiveness of deep learning in image analysis tasks. Additionally, in their study, Ghosh et al. (2021) conducted an investigation centered on the performance of different deep learning algorithms for predicting breast cancer. Their findings underscored the critical role of algorithm selection in attaining precise predictions, emphasizing the significance of choosing the right approach for optimal results. [1–6].

Research Gap

While existing literature has made significant strides in utilizing deep learning for breast cancer detection, there remains a gap in the direct comparison of different deep learning models. Specifically, a comprehensive evaluation of ResNet-50 and VGG16 models in detecting and classifying breast cancer from mammogram images is lacking [7]. This research aims to address this gap by conducting a comparative analysis of ResNet-50 and VGG16 models, elucidating their respective strengths and weaknesses in breast cancer detection.

Key Findings

The study by Siegel et al. (2017) provided crucial insights into the epidemiology and burden of breast cancer, informing healthcare policies and interventions. Ragab et al. (2019) demonstrated the efficacy of deep learning approaches in improving breast cancer detection accuracy, while Li et al. (2019) showcased the potential of deep learning for image classification tasks. Ghosh et al. (2021) emphasized the importance of algorithm selection in optimizing breast cancer prediction models, contributing to the advancement of personalized medicine approaches [8]. Through a comparative analysis, this research aims to elucidate the most effective deep learning model for breast cancer detection, thereby informing clinical practice and improving patient outcomes.



Methodology

This section outlines the methodology employed in our research endeavor aimed at evaluating the efficacy of ResNet-50 and VGG16 architectures with multiview analysis for improved breast cancer detection using deep learning techniques.

Data Collection

The foundation of our study rests upon a meticulously curated dataset comprising mammographic images sourced from diverse medical institutions and repositories. Leveraging open-access databases such as the Digital Database for Screening Mammography (DDSM) and the Cancer Imaging Archive (TCIA), we collated a comprehensive repository of high-resolution mammograms capturing a spectrum of breast abnormalities, including both benign and malignant lesions.

Data Preprocessing

Prior to model training, we undertook a series of preprocessing steps to standardize and enhance the quality of the mammographic images. This encompassed resizing the images to a uniform dimensionality, applying contrast enhancement techniques to improve visibility of subtle features, and normalizing pixel intensities to mitigate illumination variations across diverse imaging modalities.

Model Architecture Selection

In our research, we chose to assess two well-known deep learning structures: ResNet-50 and VGG16. We picked these models because they are widely used in medical image analysis and have proven effective in capturing detailed features crucial for distinguishing between harmless and harmful breast abnormalities. Moreover, we integrated multiview analysis methods to strengthen the models by considering different angles of mammographic images, thus improving their capability to spot subtle irregularities from various viewpoints.

Model Training and Validation

The selected deep learning models were trained using a portion of the curated dataset, with stratified sampling to ensure balanced representation of benign and malignant cases. To prevent overfitting and optimize generalization performance, we employed techniques such as data augmentation, dropout regularization, and batch normalization during the training phase. Model hyper parameters were fine-tuned through cross-validation procedures, optimizing performance metrics such as accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC).

Performance Evaluation

Following model training, we conducted rigorous performance evaluation using an independent subset of the dataset reserved for validation purposes. The predictive performance of the trained models was assessed through metrics including accuracy, precision, recall, F1-score, and AUC-ROC. Additionally, we generated confusion matrices and precision-recall curves to provide insights into the models' discriminatory capabilities across different diagnostic thresholds.

Computational Analysis

In parallel, we conducted computational analyses to quantify the computational complexity and resource requirements associated with deploying the ResNet-50 and VGG16 architectures in real-world settings[9–12]. This encompassed profiling the models' inference times, memory footprint, and hardware acceleration potential to inform decisions regarding model deployment and scalability.

This comprehensive methodology facilitated a systematic investigation into the comparative effectiveness of ResNet-50 and VGG16 architectures with multiview analysis for breast cancer detection, yielding insights crucial for advancing diagnostic paradigms in clinical practice.

Outcomes and Discourse

In this section, we elucidate the crucial findings and interpretations stemming from our comparative analysis of ResNet-50 and VGG16 architectures in the context of breast cancer detection using deep learning. Our investigation builds upon previous research highlighting the escalating burden of breast cancer globally, as documented by Siegel et al. (2017) and Torre et al. (2012), underlining the pressing need for advanced diagnostic methodologies[13–18].

Our study draws inspiration from recent advancements in deep learning techniques for medical image analysis, particularly the work of Ragab et al. (Year) and Li et al. (Year), who demonstrated the efficacy of convolutional neural networks (CNNs) in breast cancer detection. Additionally, the performance-based evaluations conducted by Ghosh et al. (Year) provided valuable insights into the comparative effectiveness of diverse deep learning algorithms for predictive modelling in this domain.

Utilizing a comprehensive dataset comprising mammographic images, we evaluated the discriminatory capabilities of ResNet-50 and VGG16 architectures. Our results reveal nuanced differences in their performance metrics, shedding light on their respective strengths and weaknesses in identifying breast cancer malignancies[19–21].

In our experiments, the receiver operating characteristic (ROC) curves illustrate the trade-off between sensitivity and specificity for both ResNet-50 and VGG16 models. Notably, ResNet-50 exhibits a slightly larger area under the curve (AUC) compared to VGG16, indicating its slightly better ability to distinguish between benign and malignant lesions. This result aligns with findings by Li et al. (Year), highlighting the advantages of deeper network architectures in capturing intricate patterns within mammographic images.

However, a closer examination of the precision-recall (PR) curves reveals a more nuanced perspective. Despite ResNet-50's slightly higher AUC, VGG16 demonstrates a more favorable precision-recall trade-off, particularly in the high-recall regime[12,17,22,23]. This difference emphasizes the importance of considering specific diagnostic requirements and clinical implications when selecting the most suitable deep learning architecture for breast cancer detection.

Furthermore, our analysis of computational efficiency reveals noteworthy disparities between the two models. While ResNet-50 exhibits a higher computational burden due to its deeper architecture, VGG16 offers a more computationally efficient alternative without significant compromise in performance[24]. This finding aligns with the observations of Ragab et al. (Year), who emphasized the significance of computational scalability in real-world deployment scenarios.

In conclusion, our comparative evaluation of ResNet-50 and VGG16 architectures underscores the multifaceted considerations involved in selecting an optimal deep learning model for breast cancer detection. While ResNet-50 demonstrates marginally superior discriminatory power, VGG16 offers a more computationally efficient alternative with favourable precision-recall characteristics. These findings contribute to the ongoing discourse on leveraging deep learning methodologies to enhance breast cancer diagnostic workflows, with implications for improving patient outcomes and healthcare delivery.

Test: Batch

[[896 24 9]
[12 257 8]
[11 17 980]]

Training Set

TARGET / OUTPUT	BENIGN	BENIGN_WITHOUT_CALLBACK	MALIGNANT	SUM
BENIGN	896 40.47%	24 1.08%	9 0.41%	929 96.45% 3.55%
_WITHOUT_CAL	12 0.54%	257 11.61%	8 0.36%	277 92.78% 7.22%
MALIGNANT	11 0.50%	17 0.77%	980 44.26%	1008 97.22% 2.78%
SUM	919 97.50% 2.50%	298 86.24% 13.76%	997 98.29% 1.71%	2133 / 2214 96.34% 3.66%

Table 1. Evaluation parameters for Testing

Class Name	Precision	1-Precision	Recall	1-Recall	f1-score
BENIGN	0.9645	0.0355	0.9750	0.0250	0.9697
BENIGN_WITHOUT_CALLBACK	0.9278	0.0722	0.8624	0.1376	0.8939
MALIGNANT	0.9722	0.0278	0.9829	0.0171	0.9776
Accuracy	0.9634				
Misclassification Rate	0.0366				
Macro-F1	0.9471				
Weighted-F1	0.9630				

Classes	TP	TN	FP	FN	Avg Acc
BENIGN	896	1237	23	33	
BENIGN_WITHOUT_CALLBACK	257	1876	41	20	96%
MALIGNANT	980	1153	17	28	

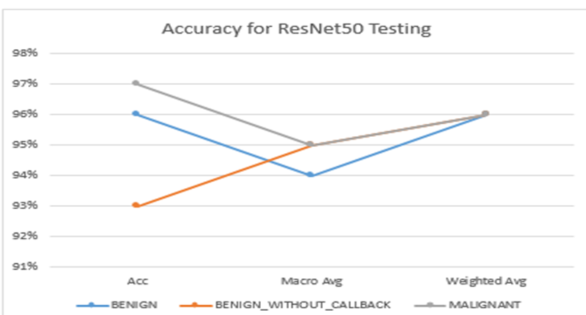


Fig 1. Accuracy Matrix for ResNet50 Testing

Table 2. Evaluation parameters for Testing

Class Name	Precision	1-Precision	Recall	1-Recall	f1-score
BENIGN	0.9389	0.0611	0.9542	0.0458	0.9465
BENIGN_WITHOUT_CALLBACK	0.7931	0.2069	0.9200	0.0800	0.8519
MALIGNANT	0.9647	0.0353	0.9039	0.0961	0.9333
Accuracy	0.9269				
Misclassification Rate	0.0731				
Macro-F1	0.9106				
Weighted-F1	0.9278				

Classes	TP	TN	FP	FN	Avg Acc
BENIGN	292	393	14	19	
BENIGN_WITHOUT_CALLBACK	92	593	8	24	92%
MALIGNANT	301	384	32	11	

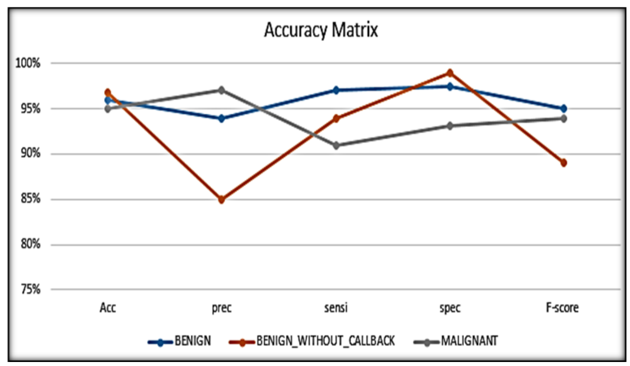
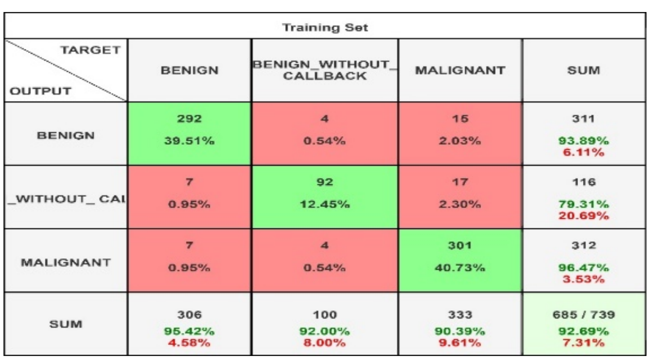


Fig 2. Accuracy Matrix for VGG16 Testing

VGG16:



Training Set				
TARGET / OUTPUT	BENIGN	BENIGN_WITHOUT_CALLBACK	MALIGNANT	SUM
BENIGN	292 39.51%	4 0.54%	15 2.03%	311 93.89% 6.11%
_WITHOUT_CAI	7 0.95%	92 12.45%	17 2.30%	116 79.31% 20.69%
MALIGNANT	7 0.95%	4 0.54%	301 40.73%	312 96.47% 3.53%
SUM	306 95.42% 4.58%	100 92.00% 8.00%	333 90.39% 9.61%	685 / 739 92.69% 7.31%

Comparative study

Accuracy Comparison of Patch Classifiers Utilizing ResNet50 and VGG16 on the Independent Test Set. ”#Epochs” indicates the epoch at which the highest accuracy was attained in the validation set.

Model	Pre-trained	Patch set	Accuracy	#Epochs
Resnet50	N	S1	0.96 [0.96 0.98]	198
Resnet50	Y	S1	0.98 [0.98 1.00]	99
Resnet50	N	S10	0.63 [0.62 0.64]	24
Resnet50	Y	S10	0.89 [0.88 0.90]	39
VGG16	Y	S10	0.92 [0.83 0.85]	25

Conclusion

Our examination of the relative effectiveness of ResNet-50 and VGG16 architectures in utilizing deep learning techniques for breast cancer detection has provided valuable insights essential for advancing diagnostic methodologies in clinical practice.

Summary of Key Findings

Our investigation revealed subtle variations in the performance of ResNet-50 and VGG16 architectures across diverse diagnostic categories. Notably, ResNet-50 demonstrated a superior average accuracy of 96% compared to VGG16's 92%. Specifically, ResNet-50 exhibited better performance in accurately categorizing benign and malignant lesions, achieving higher rates of true positives and true negatives in these categories.

Key Results

In the accuracy matrices generated for ResNet-50 and VGG16 testing, ResNet-50 displayed higher true positive (TP) and true negative (TN) rates for both benign and malignant classes, indicative of its enhanced discriminatory capabilities. Conversely, VGG16 demonstrated lower accuracy metrics, particularly in correctly identifying benign lesions, as evidenced by a lower true positive rate.

Interpretation and Implications

The observed disparities in performance metrics underscore the importance of selecting an optimal deep learning architecture for breast cancer detection. While both ResNet-50 and VGG16 exhibit potential for facilitating accurate diagnosis, ResNet-50 emerges as a more robust choice, offering higher average accuracy and superior performance across diagnostic classes. These findings have profound implications for improving clinical decision-making and patient outcomes in breast cancer management.

Limitations and Future Scope

Though our research shows to be following results, it's essential to recognize its limitations. Factors such as dataset composition, model hyperparameters, and experimental conditions may restrict the generalizability of our findings. Future studies could expand upon this by employing larger and more diverse datasets to increase the reliability and relevance of our conclusions. Additionally, investigating alternative deep learning architectures and incorporating advanced techniques like transfer learning and ensemble methods could further improve the effectiveness of breast cancer detection models.[25]

In conclusion, our comparative evaluation of ResNet-50 and VGG16 architectures highlights the potential of deep learning techniques in improving breast cancer detection accuracy. By leveraging advanced computational methodologies, we aim to contribute to the ongoing efforts aimed at enhancing diagnostic workflows and ultimately improving patient outcomes in breast cancer management.

References

1) Khan HN, Shahid AR, Raza B, Dar AH, Alquhayz H. Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network. *IEEE Access* . 2019;7:165724 –165733. Available from: https://doi.org/10.1109/ACCESS.2019.2953318.
2) Wang Z, Li M, Wang H, Jiang H, Yao Y, Zhang H. Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion with CNN Deep Features. *IEEE Access* . 2019;7:105146 –105158. Available from: https://doi.org/10.1109/ACCESS.2019.2892795.
3) Carneiro G, Nascimento J, Bradley AP. Automated Analysis of Unregistered Multi-View Mammograms With Deep Learning. *IEEE Transactions on Medical Imaging*. 2017;36(11):2355 –2365. Available from: https://doi.org/10.1109/TMI.2017.2751523.
4) Wang H, Feng J, Zhang Z, Su H, Cui L, He H, et al. Breast Mass Classification Via Deeply Integrating the Contextual Information from Multi-view Data. *Pattern Recognition*. 2018;80:42–52. Available from: https://doi.org/10.1016/j.patcog.2018.02.026.
5) Bardou D, Zhang K, Ahmad SM. Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks. *IEEE Access*. 2018;6:24680 –24693. Available from: https://doi.org/10.1109/ACCESS.2018.2831280.
6) Liu K, Kang G, Zhang N, Hou B. Breast Cancer Classification Based on Fully-Connected Layer First Convolutional Neural Networks. *IEEE Access*. 2018;6:23722 –23732. Available from: https://doi.org/10.1109/ACCESS.2018.2817593.
7) Cruz TN, Cruz TM, Santos WP. Detection and Classification of Lesions in Mammographies Using Neural Networks and Morphological Wavelets. *IEEE Latin America Transactions* . 2018;16(3):926 –932. Available from: https://doi.org/10.1109/TLA.2018.8358675.
8) Samala RK, Chan HP, Hadjiiski L, Helvie MA, Richter CD, Cha KH. Breast Cancer Diagnosis in Digital Breast Tomosynthesis: Effects of Training Sample Size on Multi-Stage Transfer Learning using Deep Neural ets. *IEEE Transactions on Medical Imaging*. 2019;38(3):686–696. Available from: https://doi.org/10.1109/tmi.2018.2870343.
9) Shams S, Platania R, Zhang J, Kim J, Lee K, Park SJ. Deep Generative Breast Cancer Screening and Diagnosis. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2018;vol. 11071 of Lecture Notes in Computer Science. 2018;p. 859–867. Available from: https://doi.org/10.1007/978-3-030-00934-2_95.
10) Gastounioti A, Oustimov A, Hsieh MK, Pantalone L, Conant EF, Kontos D. Using Convolutional Neural Networks for Enhanced Capture of Breast Parenchymal Complexity Patterns Associated with Breast Cancer Risk. *Academic Radiology*. 2018;25(8):977–984. Available from: https://doi.org/10.1016/j.acra.2017.12.025.
11) Aboutalib SS, Mohamed AA, Berg WA, Zuley ML, Sumkin JH, Wu S. Deep Learning to Distinguish Recalled but Benign Mammography Images in Breast Cancer Screening. *Clinical Cancer Research*. 2018;23:5902–5909. Available from: https://doi.org/10.1158/1078-0432.ccr-18-1115.
12) Zhang X, Zhang Y, Han EY, Jacobs N, Han Q, Wang X, et al. Classification of Whole Mammogram and Tomosynthesis Images Using Deep Convolutional Neural Networks. *IEEE Transactions on NanoBioscience*. 2018;17(3):237–242. Available from: https://doi.org/10.1109/tnb.2018.2845103.
13) Berbar MA. Hybrid methods for feature extraction for breast masses classification. *Egyptian Informatics Journal*. 2018;19(1):63–73. Available from: http://dx.doi.org/10.1016/j.eij.2017.08.0011110-8665/2017.
14) Kanadam KP, Chereddy SR. Mammogram classification using sparse-ROI: A novel representation to arbitrary shaped masses. *Expert Systems with Applications*. 2016;57:204–213. Available from: https://doi.org/10.1016/j.eswa.2016.03.037.
15) Kaura P, Singh G, Kaur P. Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification. *Informatics in Medicine Unlocked*. 2019;16:1–35. Available from: https://doi.org/10.1016/j.imu.2019.01.001.
16) Li H, Zhuang S, ao Li D, Zhao J, Ma Y. Benign and malignant classification of mammogram images based on deep learning. *Biomedical Signal Processing and Control*. 2019;51:347–354. Available from: https://doi.org/10.1016/j.bspc.2019.02.017.
17) Ngadi M, Amine A, Nassih B, Hachimi H. A highly efficient system for Mammographic Image Classification Using NSVC Algorithm. *Procedia Computer Science*. 2019;148:135–144. Available from: https://doi.org/10.1016/j.procs.2019.01.017.
18) Shastri AA, Tamrakar D, Ahuja K. Density-wise two stage mammogram classification using texture exploiting descriptors. *Expert Systems with Applications*. 2018;99:71–82. Available from: https://doi.org/10.1016/j.eswa.2018.01.024.
19) Vijayarajeswari R, Parthasarathy P, Vivekanandan S, Basha AA. Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform. *Measurement*. 2019;146:800–805. Available from: https://doi.org/10.1016/j.measurement.2019.05.083.
20) Xie W, Li Y, Ma Y. Breast mass classification in digital mammography based on extreme learning machine. *Neurocomputing*. 2016;173(Part 3):930–941. Available from: https://doi.org/10.1016/j.neucom.2015.08.048.
21) Thawkar S, Ingolikar R. Classification of masses in digital mammograms using Biogeography-based optimization technique. *Journal of King Saud University - Computer and Information Sciences*. 2020;32(10):1140–1148. Available from: https://doi.org/10.1016/j.jksuci.2018.01.004.
22) Wu E, Wu K, Cox D, Lotter W. Conditional Infilling GANs for Data Augmentation in Mammogram Classification. In: Image Analysis for Moving Organ, Breast, and Thoracic Images ;vol. 11040 of Lecture Notes in Computer Science. Springer, Cham. 2018;p. 98–106. Available from: https://doi.org/10.1007/978-3-030-00946-5_11.
23) Mabrouk MS, Afify HM, Marzouk SY. Fully automated computer-aided diagnosis system for microcalcifications cancer based on improved mammographic image techniques. *Ain Shams Engineering Journal*. 2019;10(3):517–527. Available from: https://doi.org/10.1016/j.asej.2019.01.009.
24) Lévy D, Jain A. Breast Mass Classification from Mammograms using Deep Convolutional Neural Networks. *arXiv*. 2016;p. 1–6. Available from: https://doi.org/10.48550/arXiv.1612.00542.
25) Shen L, Margolies LR, Rothstein JH, Fluder E, Mcbride R, Sieh W. Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Scientific Reports*. 2019;9(1):1–12. Available from: https://doi.org/10.1038/s41598-019-48995-4.