A Review of Breast Cancer Detection Methods Using Convolutional Neural Networks in Mammography Images

Mostafa Sabzekar
Department of Computer Engineering, Birjand University of Technology
Birjand, Iran
sabzekar@birjandut.ac.ir

ABSTRACT

In recent decades, mammography has been recognized as one of the most important breast cancer screening tools. Today, the use of deep learning has increased the accuracy and efficiency of these methods, providing significant assistance to radiologists. This article provides a review of recent studies in the application of Convolutional Neural Networks (CNNs) in mammography images. Firstly, models based on Convolutional Neural Networks (CNN), one of the most important deep learning algorithms, are introduced. Subsequently, the most recent articles are discussed for four distinct applications in mammography: 1) Breast density classification, 2) Asymmetry detection and classification, 3) Mass detection and classification, and 4) Calcification detection and classification. Additionally, the article presents the use of approved models by the Food and Drug Administration (FDA) and discusses the practical applications of the mentioned algorithms in the real world. Finally, open research challenges in the field of CNN-based methods for further improving breast cancer detection are outlined.

KEYWORDS: Mammography, deep learning, Convolutional Neural Networks (CNNs).

1 INTRODUCTION

Through research conducted in the United States, breast cancer has been identified as the most common non-skin cancer diagnosed in women. Researchers estimated that over 276,000 women were diagnosed with breast cancer in 2020, and more than 42,000 of them lost their lives [1]. Mammography screening has been recommended by researchers for early detection and saving lives from breast cancer. Studies have shown that after screening, mortality has decreased by up to 40% [2] and [3]. However, mammography screening has its limitations. For every thousand women who undergo mammography screening, physicians recommend needle biopsies for 15 women, among which 10 to 13 women do not have cancer (false positives) [4]. Researchers have proposed various solutions, such as re-reading images, annual screenings [5], obtaining two views of each breast [6], and analysing previous mammograms for comparison [7] to improve the efficiency of mammography screening.

The primary goal of radiologists is to identify vital features in the breast, such as small masses (MCs), architectural distortions (ADs), and asymmetries, as biological markers of this disease or the risk of developing it. Visual detection of these features is not feasible and can reduce the accuracy of radiologists' work [8].

The first attempts to automatically identify and classify breast cancer from mammography images using Computer-Aided Detection (CAD) systems emerged in the 1990s. However, due to low accuracy, traditional CAD systems have never been able to provide satisfactory performance in screening [9, 10].

Research [11] has emphasized that deep learning using Convolutional Neural Networks (CNN) significantly improves the performance of screening and also enhances the accuracy of radiologists' interpretation of breast images. Researchers have developed several CNN-based algorithms for automated mammography analysis, some of which have been approved by the United States Food and Drug Administration (FDA) [12].

Since CNN-based algorithms have provided better performance compared to other methods, this article reviews the latest research efforts aimed at breast cancer screening from mammography images using these approaches. The following chapter will provide an overview of the CNN structure. The third chapter introduces standard and freely available datasets for breast cancer classification and detection. The fourth chapter focuses on the examination and comparison of the mentioned methods. The final chapter will summarize and provide an outlook on this research area.

2 CONVOLUTIONAL NEURAL NETWORKS

These types of networks are very powerful for image analysis because they have filters with specific patterns that can preserve spatial features in the image. Additionally, they outperform fully connected neural network methods since fully connected architectures flatten input images and do not consider essential spatial features. Figure 1 illustrates the architecture details of CNN, representing their end-to-end learning capability.

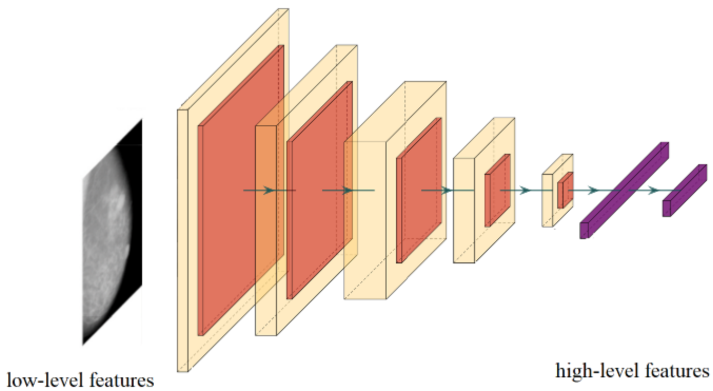

Figure 1 - Architecture of a CNN (Convolutional Neural Network) and its learning model. Different filters create various feature maps, as depicted by the orange stacks. The darker orange colour represents the output after pooling. The architecture also flattens the layers into vectors (purple) before classification. The model updates feature mapping parameters based on a loss function and propagates them through backpropagation. The feature maps expand laterally after convolution.

These types of networks require minimal feature engineering for classifying mammography images based on raw input data. As shown in Figure 2, the shallow layers detect features similar to the input image, while the deeper layers extract abstract features.

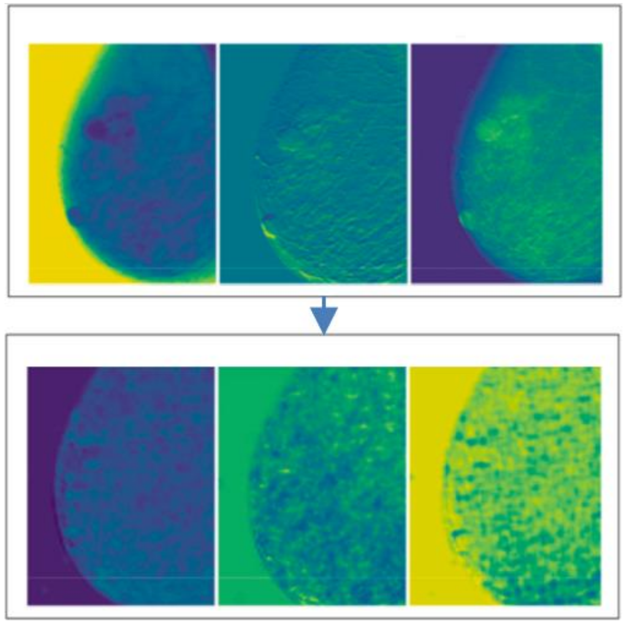

Figure 2: As the model deepens, high-level features (bottom image) are abstracted using low-level features (top image). Radiologists find it easier to recognize patterns from low-level features, while convolutional networks perform better with high-level features.

One of the key elements in CNNs is optimizers. Researchers use optimizers to minimize the loss function through backpropagation. Backpropagation updates the model's weights based on relative errors, leading to a reduction in the loss function during the algorithm. Various optimizers are used today, including Stochastic Gradient Descent (SGD), AdaGrad, and Adam, among others.

Large feature map channels with basic convolutional models can create problems such as slow training and increased computational costs. To address these issues, techniques like max-pooling and average-pooling are used for the outputs of the convolutional layers. Max-pooling and average-pooling methods consider the maximum and average values of all pixels in a two-dimensional window, summarizing the feature maps into lower-dimensional codes. By employing these operations, the number of trainable parameters is reduced, which, in turn, helps reduce overfitting and computational time.

Activation functions used as the core of neural network models transform the non-linear output into a linear one. For many applications related to mammography, especially binary classification tasks, the sigmoid activation function is frequently used in the final layer. These functions are also applied in image segmentation. Additionally, for multi-class problems, the softmax function is commonly used. It's worth noting that the ReLU activation function is highly efficient due to its minimal computational operations, leading researchers to use it in hidden layers of the network.

The dropout technique is used to reduce overfitting, improve generalization, and enhance accuracy on test and validation data. In this technique, during training, some of the neuron weights are randomly set to zero. Research has shown that this approach results in very powerful CAD models.

Batch normalization ensures that the next layer receives the data as expected and scales high values into a smaller range. Furthermore, skip connections are used to create alternative paths for error backpropagation, connecting the layers of the network. Networks like ResNets, U-Nets, DenseNets, and Squeeze-Excitation rely on skip connections. Skip connections link earlier to later layers, preserving the flow of critical information through deep networks.

3 MAMMOGRAPHY DATASETS

There are numerous mammography datasets available for researchers in the field of CNN. Some of the most common datasets used, along with their details, including publication year, image format, image views, resolution, and interpretations, are presented in Table 1. Most of these datasets consist of digital mammography images with full-field digital mammography (FFDM), which involves electronic X-ray processing.

The first dataset, containing over 300 screening mammography images, is known as the Mammographic Image Analysis Society (MIAS) dataset [13]. This dataset includes interpretations such as tissue density (fatty or dense), the presence of abnormalities such as masses or asymmetry, and the severity of abnormalities. MIAS records the locations of breast lesions in X and Y coordinates. Additionally, this dataset encompasses four classes with labels: small calcification, structural abnormality, asymmetry, and normal.

The Digital Database for Screening Mammography (DDSM) [14] archives over 2,600 scanned mammography films. A subset of this database, named CBIS-DDSM, has been organized by experts and includes well-annotated images. This dataset provides information on regions of interest and detailed pathological data related to breast lesion type, tumor grade, and disease stage. The majority of this dataset consists of scanned screening mammography films, remaining unaffected by advanced imaging methods such as FFDM and digital breast tomosynthesis (DBT).

The INBreast dataset [15] comprises over 410 mammography images from the results of around 115 individuals' screenings. This dataset contains information regarding various abnormalities, shapes, and lesion areas.

The OPTIMAM dataset [16], collected in the past two decades, is available as a downloadable open-source Python package. It includes data from the first and second OPTIMAM projects, initiated in 2008 and 2013, respectively. OPTIMAM2, a rare public dataset, consists of interpreted three-dimensional DBT images. Furthermore, this dataset features an image simulation tool capable of generating artificial two-dimensional mammography images and three-dimensional DBT images [17].

Table 1 Some of the Most Important Standard Mammography Datasets

Dataset Name	**Year**	**Number of Images**
mini-MIAS [13]	2003	322
DDSM [14]	1999	10,480
INBreast [15]	2011	410
OPTIMAM [16]	2020	Over 1 million

4 BREAST CANCER SCREENING METHODS USING CNNS

Researchers have made efforts to address the challenges in mammography applications using standard breast cancer datasets, some of which were mentioned in the previous section. They aim to overcome common clinical challenges, including low sensitivity and high false-positive rates. In this section, we will review detection methods using CNNs (Convolutional Neural Networks). Table 2 provides a summary of the presented methods, including the year of publication, application, employed architecture, advantages, and disadvantages of each method. The term "application" refers to which of the four mentioned issues the method addresses. The employed architecture indicates the type of learning used to solve the problem, including Unet, ResNet, 3D CNNs, autoencoders, and more.

Table 2 - A comparison of the presented methods for breast cancer detection using mammography images based on CNNs

Method	Year	Application	Architecture	Pros/Cons
Unsupervised Deep Autoencoders [18]	2012	Breast density classification	Autoencoders	Suitable for dimension reduction, feature detection, and reconstruction/ requires explainable abstractions due to encoding transformations.
Multi-scale patches [19]	2017	Detection and classification of calcifications	ResNets	Reduces the number of learnable parameters, but non-linearity can lead to non-interpretability/ non-linearity can result in non-explainability.
Supervised Pretrained CNN on Predicting BIRADS Classes [20]	2017	Breast density classification	Transfer Learning	Reduced training time, increased high-level feature transfer/ Lacks fine-tuning, leading to poor performance.
RetinaNet based Detector [21]	2018	Mass detection and classification	Multi-Scale CNNs	Aggregates context, larger receptive field, scalable image scales/ Choosing image scales is often arbitrary.
Attention-guided Dense Upsampling segmentation [22]	2018	Mass detection and classification	Attention	Selective focus on segments of sequential data, increased spatial complexity attention.
Transfer Learning with Deep CNN [23]	2018	Breast density classification	Transfer Learning	Reduced training time, increased high-level feature transfer/Lacks fine-tuning, leading to poor performance.
Multiscale Autoencoder Segmentation [24]	2019	Mass detection and classification	UNet	Excellent performance for small datasets, maintains information from the encoder through skip connections.
GAN based Data Augmentation [25]	2019	Mass detection and classification	Generative Models	Addresses data augmentation issues, helps balance class-imbalanced datasets/ Computational cost
DenseNet201 Transfer Learning [26]	2019	Asymmetry detection and classification	Transfer Learning	Reduced training time, increased high-level feature transfer/ Lacks fine-tuning, leading to poor performance.

Dilated Attention Guided CNN [27]	2020	Breast density classification	Attention	Selective focus on segments of sequential data, increased spatial complexity attention.
International Mass and Calcification Classification [28]	2020	Mass and calcification detection and classification	Network Ensembling	Grouping each individual model forces them to compensate for the weaknesses of others/ The challenging aspect is explaining and arguing about how a model makes decisions in complex model combinations.
SE-Attention Based Classifier [29]	2020	Detection and classification of breast asymmetry	Attention-based classifier	Focusing on selective segments of sequential data /significantly increasing spatial complexity.
Deep Belief-RF Detection [30]	2020	Mass detection and classification	Network Ensembling	Grouping each individual model forces them to compensate for the weaknesses of others/ The challenging aspect is explaining and arguing about how a model makes decisions in complex model combinations.
Automated CNN approach [31]	2021	Mass detection and classification	UNet	Excellent performance for small datasets, maintains information from the encoder through skip connections/It is challenging to understand how the UNet image transformations and traditional signal processing approaches.
Dual-Path Architecture [32]	2022	Mass detection and classification	DualCoreNet	Better performance in both image segmentation and classification compared to similar methods.
Combining Approach [33]	2022	Mass detection and classification	Network Ensembling	Achieved an increase in accuracy up to 98% on MIAS and DDSM datasets.
CNN-Wavelet Scattering [34]	2023	Breast density classification	Wavelet CNN	Combines multi-resolution analysis with CNN and leverages spectral information to improve performance.

The summarized research in Table 2 represents only a portion of the vast body of studies conducted in this field using CNN, demonstrating its special merits in problem-solving. It's essential to note that the use of this method in various image processing domains has turned it into a go-to tool for achieving high-precision results. Researchers have high hopes for it, as it has been employed effectively in a wide range of applications, thanks to its consistent advancements. For instance, in the medical field, CNN can significantly contribute to reducing diagnosis time and increasing accuracy in the detection and classification of medical images such as mammography, CT scans, and X-rays. In computer vision and image processing, CNN is used for object detection, event recognition, and even autonomous driving in vehicles.

In general, CNN has attracted considerable attention from researchers worldwide and has become a fundamental tool in the fields of computer vision and image processing, leading to significant advancements in these areas.

5 CONCLUSION AND FUTURE DIRECTIONS

In this discussion, we reviewed CNN-based methods for mammography images and their applications in various domains, with a focus on breast cancer detection and diagnosis. CNNs have proven to be a powerful tool for image processing, pattern recognition, and have been widely adopted in recent years for a range of applications. In order to achieve applications of asymmetry detection, further research is necessary, which requires datasets accompanied by more extensive interpretation. Therefore, future studies should focus on generating more extensive datasets. Despite the existence of public datasets of breast images, additional data is still needed in this area. To validate the robustness of the presented models, they need to be applied and tested on multiple datasets.

Research efforts should concentrate more on other applications of CNN-based methods in mammography images, including the classification of breast tissue density, identification and classification of asymmetry, and firmness. Additionally, there is a need for more research on the classification of various types of masses in more detail. Moreover, models generally perform better in fatty breast tissues compared to dense tissues. Therefore, research addressing these challenges may significantly enhance CAD systems.

It is worth noting that researchers must pay attention to the clinical implications of their findings and how to convey them to medical and computational audiences. The interpretation of research results in clinical training is crucial for the next generation of physicians. For example, radiologists rely on changes in imaging capabilities for accurate diagnosis and decision-making. Hence, to achieve optimal performance, computational models and educational programs should integrate with these new capabilities,

In conclusion, CNN-based CAD systems can work in synergy with medical experts, providing numerous opportunities for clinical and computational innovations. Further research in these directions will play a pivotal role in enhancing the capabilities and applicability of CNNs in mammography and other image processing domains. It is important for researchers to focus on the clinical implications of their findings and how these computational models can assist medical professionals in making more accurate diagnoses and decisions.

REFERENCES

[1] National Cancer Institute, Cancer of the breast (female) - cancer stat facts, SEER (2020), 1–1.

[2] Sannella, M. J., Constraint Satisfaction and Debugging for Interactive User Interfaces, Ph.D. Thesis, University of Washington, Seattle, WA, 1994.

[3] C. Nickson, K.E. Mason, D.R. English, A.M. Kavanagh, Mammographic screening and breast cancer mortality: a case-control study and meta-analysis, Cancer Epidemiol. Biomark. Prev. 21 (9) (Sep. 2012) 1479–1488, https://doi.org/10.1158/1055-9965.epi-12-0468.

[4] W.A. Berg, R.E.H. PhD, D.B. Kopans, P. Robert, A. Smith, Frequently asked questions about mammography and the uspstf recommendations: a guide for practitioners, Soc. Breast Imag. (2020) 1–17.

[5] S. Destounis, A. Santacroce, Age to begin and intervals for breast cancer screening: balancing benefits and harms, Am. J. Roentgenol. 210 (2) (2018) 279–284, https://doi.org/10.2214/AJR.17.18730.

[6] E. Sickles, W. Weber, H. Galvin, S. Ominsky, R. Sollitto, Baseline screening mammography: one vs two views per breast, Am. J. Roentgenol. 147 (6) (1986) 1149–1153, https://doi.org/10.2214/ajr.147.6.1149.

[7] A.A. Roelofs, et al., Importance of comparison of current and prior mammograms in breast cancer screening, Radiology 242 (1) (2007) 70–77, https://doi.org/10.1148/radiol.2421050684.

[8] A. Rimmer, Radiologist shortage leaves patient care at risk, warns royal college, BMJ Br. Med. J. (Clin. Res. Ed.) 359 (2017), https://doi.org/10.1136/bmj.j4683.

[9] J.J. Fenton, et al., Influence of computer-aided detection on performance of screening mammography, N. Engl. J. Med. 356 (14) (Apr. 2007) 1399–1409, https://doi.org/10.1056/nejmoa066099.

[10] C.D. Lehman, R.D. Wellman, D.S.M. Buist, K. Kerlikowske, A.N.A. Tosteson, D. L. Miglioretti, Diagnostic accuracy of digital screening mammography with and without computer-aided detection, JAMA Internal Med. 175 (11) (Nov. 2015) 1828, https://doi.org/10.1001/jamainternmed.2015.5231.

[11] A.D. Trister, D.S.M. Buist, C.I. Lee, Will machine learning tip the balance in breast cancer screening? JAMA Oncol. 3 (11) (Nov. 2017) 1463, https://doi.org/10.1001/jamaoncol.2017.0473.

[12] American College of Radiology, Data science Institute,, FDA Clear. AI algorith. (2020), 1–1.

[13] J. Suckling, et al., Mammographic Image Analysis Society (Mias) Database v1.21, 2015.

[14] M. Heath, et al., Current status of the digital database for screening mammography, Comput. Imag. Vision Digital Mammogr. (1998) 457–460.

[15] I. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. Cardoso, J. Cardoso, INbreast: toward a full-field digital mammographic database, Acad. Radiol. 19 (Nov. 2011) 236–248.

[16] M.D. Halling-Brown, et al., OPTIMAM mammography image database: a large scale resource of mammography images and clinical data, arXiv preprint arXiv, 2004, 04742, 2020.

[17] P. Elangovan, A. Hadjipanteli, A. Mackenzie, D.R. Dance, K.C. Young, K. Wells, OPTIMAM image simulation toolbox-recent developments and ongoing studies, in: International workshop on breast imaging, 2016, pp. 668–675, https://doi.org/10.1007/978-3-319-41546-8_83.

[18] K. Petersen, K. Chernoff, M. Nielsen, A.Y. Ng, Breast density scoring with multiscale denoising autoencoders, in: In Lecture Notes in Computer Science: Authors' Instructions, 2012, pp. 1–8.

[19] W. Lotter, G. Sorensen, D. Cox, A multi-scale cnn and curriculum learning strategy for mammogram classification, in: In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2017, pp. 169–177.

[20] A.A. Mohamed, W.A. Berg, H. Peng, Y. Luo, R.C. Jankowitz, S. Wu, A deep learning method for classifying mammographic breast density categories, Med. Phys. 45 (1) (2017) 314–321, https://doi.org/10.1002/mp.12683.

[21] H. Jung, et al., Detection of masses in mammograms using a one-stage object detector based on a deep convolutional neural network, PloS One 13 (9) (2018) p. e0203355.

[22] Sun, C. Li, B. Liu, H. Zheng, D.D. Feng, S. Wang, AUNet: Attention-Guided Dense-Upsampling Networks for Breast Mass Segmentation in Whole Mammograms, 2018, https://doi.org/10.1088/1361-6560/ab5745.

[23] N. Wu et al., "Breast density classification with deep convolutional neural networks," In 2018 Ieee International Conference on Acoustics, Speech and Signal Processing (Icassp), vol. 2018, pp. 6682–6686, https://10.1109/ICASSP.2018.8462671.

[24] Y. Yan, et al., Cascaded multi-scale convolutional encoder-decoders for breast mass segmentation in high-resolution mammograms, in: In 2019 41st Annual International Conference of the Ieee Engineering in Medicine and Biology Society (Embc), 2019, pp. 6738–6741, https://doi.org/10.1109/EMBC.2019.8857167.

[25] S. Guan, M. Loew, Breast cancer detection using synthetic mammograms from generative adversarial networks in convolutional neural networks, J. Med. Imag. 6 (3) (2019) 1–10, https://doi.org/10.1117/1.JMI.6.3.031411.

[26] X. Yu, N. Zeng, S. Liu, Y.-D. Zhang, Utilization of densenet201 for diagnosis of breast abnormality, Mach. Vis. Appl. 30 (7) (Oct. 2019) 1135–1144, https://doi.org/10.1007/s00138-019-01042-8.

[27] C. Li, et al., Multi-view mammographic density classification by dilated and attention-guided residual learning, in: IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2020, https://doi.org/10.1109/TCBB.2020.2970713.

[28] S.M. McKinney, et al., International evaluation of an ai system for breast cancer screening, Nature 577 (7788) (2020) 89–94, https://doi.org/10.1038/s41586-019-1799-6.

[29] J. Deng, Y. Ma, D.-a. Li, J. Zhao, Y. Liu, H. Zhang, Classification of breast density categories based on se-attention neural networks, in: Computer Methods And Programs In Biomedicine vol. 193, 2020, p. 105489, https://doi.org/10.1016/j.cmpb.2020.105489.

[30] T. Schaffter, et al., Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms, JAMA Netw. open 3 (3) (2020), https://doi.org/10.1001/jamanetworkopen.2020.0265 e200265–e200265.

[31] Salama, W. M., & Aly, M. H. (2021). Deep learning in mammography images segmentation and classification: Automated CNN approach. Alexandria Engineering Journal, 60(5), 4701–4709. https://doi.org//10.1016/j.aej.2021.03.048.

[32] H. Li, D. Chen, W. Nailon, M. davis, and D. Laurenson, "Dual Convolutional Neural Networks for Breast Mass Segmentation and Diagnosis in Mammography," IEEE Transactions on Medical Imaging, vol. 41, no. 1, pp. 3-13, 2022.

[33] J. Melekoodappattu,A. Dhas, B. Kandathil, and K. S. Adarsh, "Breast cancer detection in mammogram: combining modified CNN and texture feature based approach," Journal of Ambient Intelligence and Humanized Computing, 2022, https://doi.org/10.1007/s12652-022-03713-3.

[34] N. Razali, I. Isa, S. Sulaiman, N. Karim, and M. Osman, "CNN-Wavelet scattering textural feature fusion for classifying breast tissue in mammograms," Biomedical Signal Processing and Control, vol. 83, 104683, 2023.