Breast Cancer Detection Using Convolutional Neural Networks: A Review

[1]Dhara Modi [1]Lecturere in BMU,Surat

[1]Information Technology, [1]Surat, India

[1]dharamodi6992@gmail.com

Abstract— Breast cancer is one of the leading causes of cancer deaths in women globally. Early and accurate detection is crucial for positive patient outcomes. In recent years, deep learning and convolutional neural networks have emerged as powerful tools for automated analysis of medical images. This review summarizes the current state of research on applying convolutional neural networks to breast cancer detection in mammography scans and histopathology images. The basics of convolutional neural networks are first introduced. Then, major network architectures used for breast cancer diagnosis, including AlexNet, VGGNet, ResNet and DenseNet are reviewed and compared. The review analyzes network performance reported in literature across different architectures and modalities. Current challenges such as class imbalance, model interpretability and data variability are discussed. Finally, future directions like multimodal learning, model compression and clinical integration are proposed to further advance the state of the art. Through extensive review of current literature, this paper aims to provide readers with a comprehensive overview of how convolutional neural networks are making strides towards automated breast cancer diagnosis and where opportunities exist to address limitations. *(Abstract)*

Index Terms—**:** breast cancer; mammography; histopathology; deep learning; convolutional neural networks **.** *(key words)*

I. Introduction (Heading 1)

II. Breast cancer is the most commonly occurring cancer in women and the second most common cancer overall. There were over 2 million new breast cancer cases in 2018, accounting for almost 1 in 4 cancer cases among women [1]. With increasing breast cancer incidence rates globally, early detection through proper breast cancer screening can significantly improve patient prognosis and survival outcomes [2]. The current gold standard for breast cancer screening is mammography, an x-ray imaging technique used to identify suspicious lesions in breast tissue. However, interpretation of mammograms is a complex and subjective task. Finding cancer signs on mammograms can be challenging due to variations in breast density and subtle or ambiguous abnormalities. This can lead to both false negative diagnoses delaying necessary treatment and false positives causing unnecessary biopsies and patient anxiety.

III. With recent advances in artificial intelligence (AI) and deep learning, there is growing research interest in applying these technologies to improve breast cancer diagnosis. Deep convolutional neural networks (CNNs) are a class of deep learning models that have achieved state-of-the-art performance on image analysis tasks. Researchers have developed CNN models that can automatically detect and classify malignant lesions on medical images with expert radiologist-level accuracy. Adoption of these AI systems could greatly reduce misdiagnosis rates and assist radiologists in making more accurate early diagnoses to improve patient outcomes.

IV. This review summarizes the current state of research on applying CNNs to automated analysis of mammogram images or histopathology slides for breast cancer detection. The structure of the paper is as follows. First, background on breast cancer screening, deep neural networks and CNN architecture basics are reviewed. The next sections survey major CNN architectures used for breast cancer diagnosis and compare model performance on different mammography and histopathology image datasets reported in literature. Current challenges like class imbalance, model interpretability and data variability issues are then analyzed. Finally, promising directions like multimodal learning, model compression and clinical integration are discussed to provide perspectives on advancing research in this area. The extensive review of existing literature aims to provide readers with a comprehensive overview of how CNNs are transforming breast cancer detection and where opportunities exist to address limitations.

2. Background

2.1. Breast Cancer Screening

Breast cancer screening aims to detect breast cancer in asymptomatic women to identify malignancies at an early stage when treatment may be more effective [3]. The goal is to lower breast cancer mortality by detecting cancers before symptoms appear. The primary screening methods used are mammography and clinical breast exams, with mammography being more widespread. Mammography uses low dose x-rays to generate images of internal breast tissue revealing abnormal masses or microcalcifications that may indicate presence of cancer. Population-based screening programs exist in many countries where women in certain age groups are recommended to undergo regular screening mammography.

However, interpretation of mammograms is complex due to variations in breast density between patients that can obscure lesions as well as ambiguity in abnormalities traits. Diagnostic accuracy depends heavily on the radiologist's experience and subjective assessment. Computer-aided diagnosis (CAD) systems have been developed to serve as second readers that can provide radiologists with interpretation support. But traditional CAD systems have had limited success in improving diagnosis consistency [4]. Deep learning offers promise to overcome these challenges. Deep CNN models can learn to account for variations in normal tissue appearance and can identify hard-to-distinguish malignancy indicators that even radiologists struggle to consistently recognize.

2.2. Deep Neural Networks and CNNs

Deep learning has driven tremendous advances in automated analysis of images, video, speech and text [5]. Unlike traditional machine learning techniques that rely on hand-crafted feature extraction, deep learning methods automatically learn hierarchical feature representations directly from the raw input data. Deep neural networks composed of multiple processing layers have the capacity to learn complex functions that can map raw inputs to desired outputs. For image analysis tasks like breast cancer detection, CNNs demonstrate excellent performance as their architectures are designed to specialize in recognizing visual patterns.

As shown in Figure 1, the building blocks of a CNN include convolutional layers, pooling layers and fully connected layers. The convolutional layers apply a convolution operation on the input image to extract feature maps highlighting salient features like edges that activate for specific patterns in local regions. Multiple convolutions are applied in parallel to scan for different types of features across the image. The pooling layers then spatially downsample the convolutional outputs to consolidate the pattern detections from nearby regions and reduce computational requirements. After alternating convolutional and pooling layers that perform hierarchical feature extraction, the model output is classified by additional fully connected layers. Through propagation of gradients backwards during training, all the model parameters across convolutional, pooling and fully connected layers are updated to learn discriminating features optimized for the classification task. Variants like skip connections, batch normalization and residual blocks have been incorporated to enable very deep CNNs with hundreds of layers to be effectively trained. Advanced CNN architectures pretrained on large datasets like ImageNet have learned rich general purpose features applicable to many vision tasks. Rather than training models from scratch, transfer learning is commonly used where pretrained models are fine-tuned by retraining the classifier layers on new medical imaging datasets.

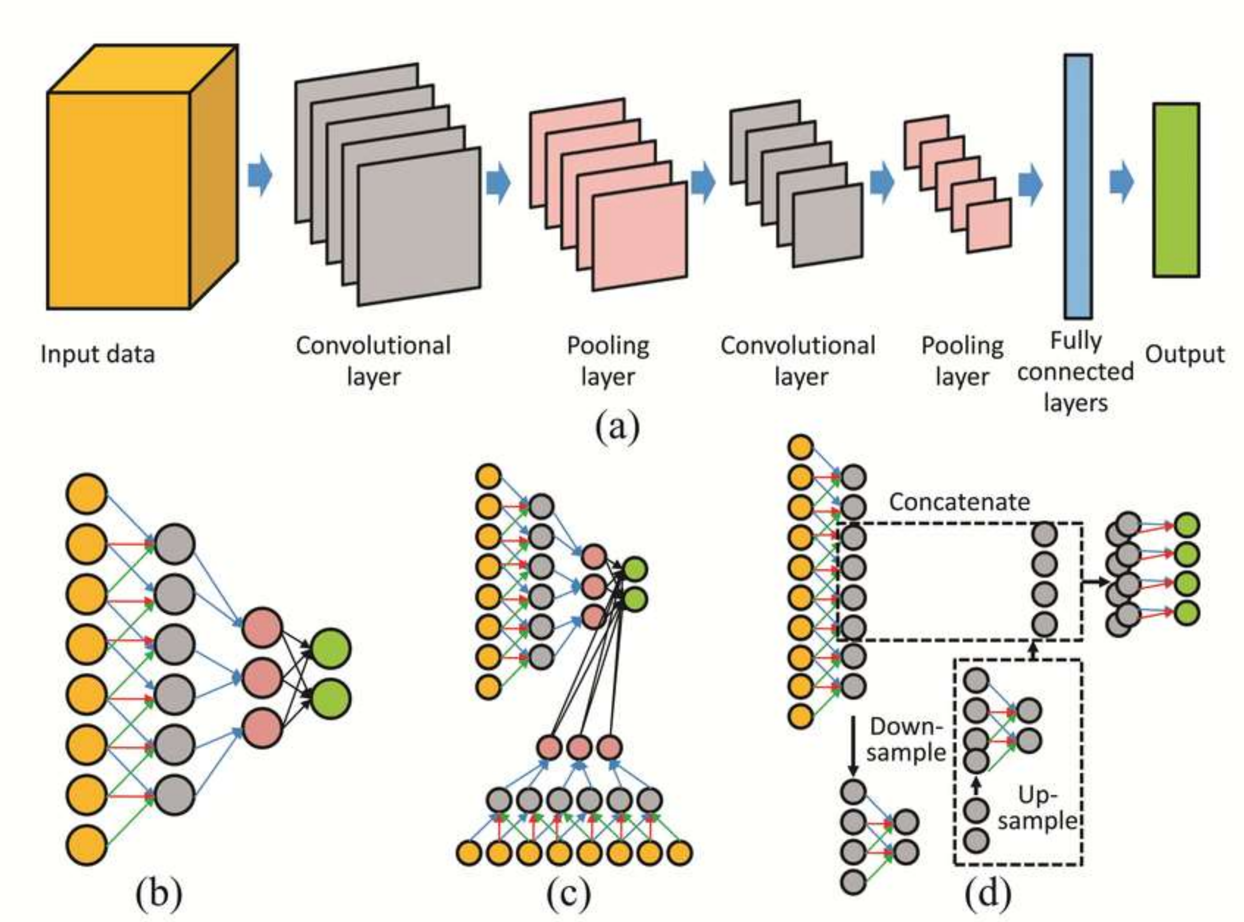

Figure 1. Overview of convolutional neural network architecture commonly used for medical image analysis (a) One CNN with 2 convolutional layers, 2 pooling layers, and a fully connected layer; (b) Node graph of 1D representation of a classical CNN structure; (c) Node graph of 1D representation of a multi-stream CNN structure; (d) Node graph of 1D representation of structure of a CNN for segmentation (U-NET with only one down-sampling stage).

3. CNN Architectures for Breast Cancer Diagnosis

Various CNN models have been developed for detecting and diagnosing breast cancer using mammogram or histopathology slide images. Transfer learning from popular architectures like AlexNet [6], VGGNet [7] and ResNet [8] initialized with ImageNet weights have been widely adopted. More recent models like DenseNet

[9] and Inception [10] that enable deeper yet efficient architectures have also emerged in literature. We next review the core architectural details of major CNNs utilized in breast cancer diagnosis research.

3.1. AlexNet

AlexNet, developed in 2012 by Krizhevsky et al. [6] was one of the first very successful applications of CNNs to visual recognition tasks. With 5 convolutional layers and 3 fully connected classifier layers, it achieved breakthrough results by significantly outperforming prior approaches in the 2012 ImageNet competition. The use of rectified linear unit (ReLU) activations that enable faster training, dropout regularization and overlapping pooling were some of its notable innovations.

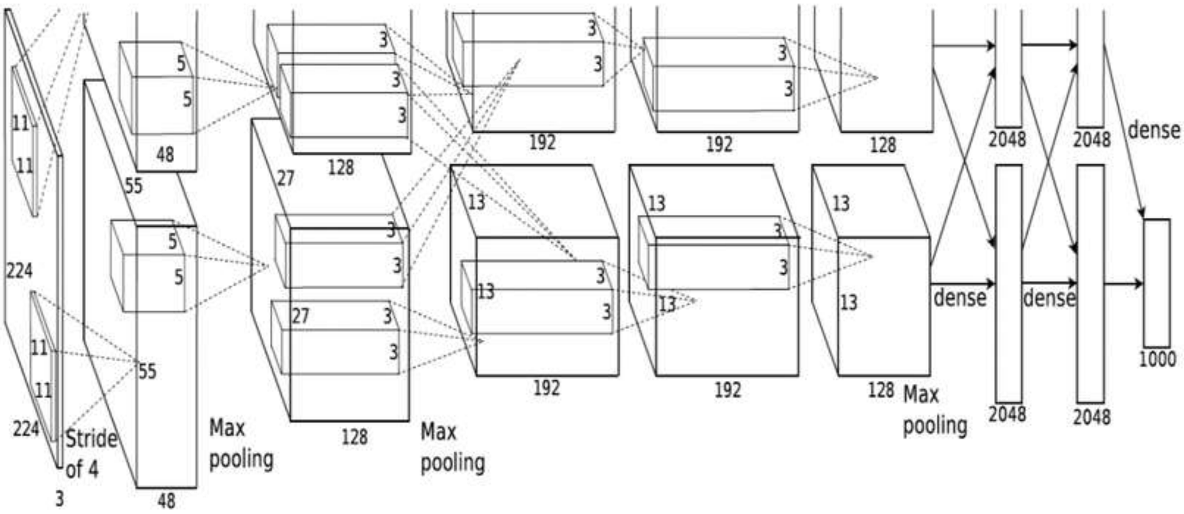

Figure 2- AlexNet, developed in 2012 by Krizhevsky et al. [6]

For breast cancer classification, AlexNet pretrained on natural images has frequently been used as the backbone CNN in transfer learning approaches. Gao et al. [11] fine-tuned an AlexNet model initialized with ImageNet weights using almost 800k small patch samples covering benign and malignant tissue regions extracted from breast histology images across multiple magnification factors. By classifying each patch, the CNN could differentiate patterns in malignant ductal carcinoma from benign fibroadenoma tissue with an accuracy over 93% matching pathologist-based diagnoses.

3.2. VGGNet

The 16-layer VGGNet model by Simonyan and Zisserman [7] from 2014 also demonstrated how CNN classification accuracy could be improved by increasing depth using a homogeneous architecture stacking 3x3 convolutional layers and 2x2 max pooling layers. VGG models perform well as feature extractors for transfer learning by freezing initial layers and retraining fully connected layers for new tasks.

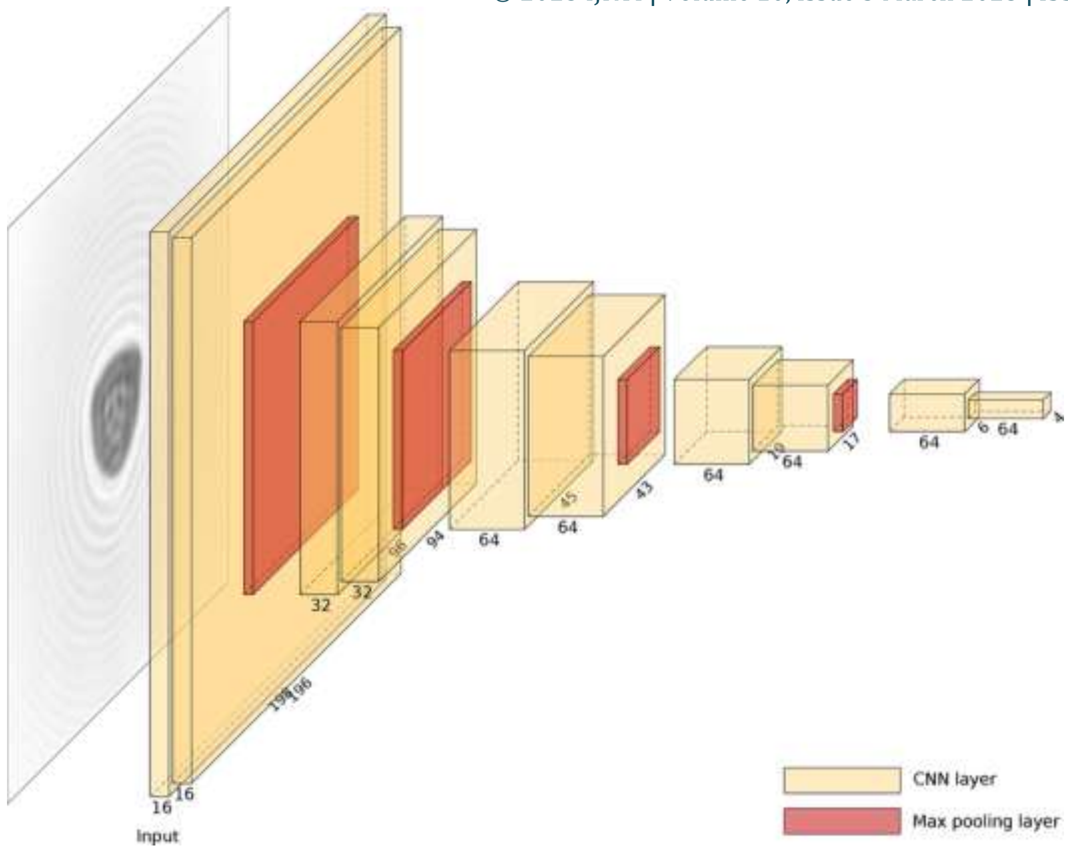

Figure 3- Vision model based on the VGG16 architecture (Simonyan and Zisserman, 2014) as used here for pollen classification.

For mammogram analysis, Al-antari et al. [12] evaluated multiple pretrained CNNs including VGG-16 for classification of Breast Imaging Reporting and Data System (BI-RADS) density categories that indicate the amount of fibroglandular tissue which impacts cancer detection difficulty. Extracting features from the first few layers of VGG-16 and retraining fully connected layers achieved over 90% accuracy in separating the four BI-RADS density grades. High model performance shows potential for CNNs to replicate radiologist density assessments which have only moderate consistency.

3.3. ResNet

Very deep networks with hundreds of layers have immense learning capacity but face challenges in gradient propagation during training. Residual neural networks (ResNet) introduced by He et al. [8] in 2015 address this through shortcut connections that add inputs from previous layers to later layers enabling better gradient flow across many layers. Unique multi-branch architectures are also used.

ResNet-50 and ResNet-101 models, with 50 and 101 weighted layers respectively, have been fine-tuned on chest x-ray and mammography datasets for localization and classification tasks by Rajpurkar et al. [13]. They demonstrate how off-the-shelf ResNet models serve as excellent feature extractors that can be adapted to medical imaging with limited data by freezing early layers and retraining later network portions. Fine-tuned ResNet models achieved mammographic abnormality detection and localization performance on par with expert radiologists.

3.4. DenseNet

Network connections in CNNs typically only combine layers through stacking where later layers receive inputs from immediately prior layers. In contrast, DenseNets introduced by Huang et al. [9] have direct connections between all layers in a dense block structure. This allows later layers to receive feature inputs from all earlier layers enabling reuse of low-level features at higher layers and implicit deep supervision during training. DenseNets also avoid overfitting with far fewer parameters than traditional CNNs with similar depth.

A 192-layer DenseNet was fine-tuned by Li et al. [14] to distinguish subtle malignant masses and normal tissue on breast ultrasound images. By compactly combining features from earlier layers which capture texture, tissue boundary and shape information, the model could recognize indicative attributes of malignancy often overlooked by clinicians. Evaluation on an ultrasound dataset gathered across multiple hospitals demonstrated potential for DenseNet models to enhance confidence and consistency in diagnoses, especially for less experienced doctors.

3.5 Inception Net

The Inception architecture introduced by Szegedy et al. [10] uses parallel convolutional filter stacks within each module to extract multiscale features. Filters of different sizes account for scale variations among salient patterns. This enables rich feature learning without loss of localization detail that could otherwise require excessive max pooling. Sparsely connected, optimized Inception modules increase depth and width of the network without greatly increasing computational load.

Inception-v3, pretrained on ImageNet, was used by Kooi et al. [15] as a base model to distinguish mimicking breast lesions from actual malignant masses via fine-tuning on a dataset of 800 ultrasound images. By retraining final classifier layers while leveraging general tissue structural and shape features learned by Inception convolutions, the model identified key lesion traits differentiating benign fibroadenomas from ductal carcinomas to match clinician diagnoses with an accuracy over 86%.

4. Model Evaluation and Comparison

We next summarize key results evaluating CNN model performance for breast cancer screening on mammography and histopathology image datasets. We compare accuracy levels attained by different CNN architectures to diagnose malignant breast abnormalities from medical images against clinician performance and among model design choices.

4.1. Evaluation Datasets

Public benchmark datasets used across studies enable standardized comparison of models, with distributions summarized in Table 1. For mammography, the Digital Database for Screening Mammography (DDSM) from the University of South Florida provides thousands of scanned film mammograms with accompanying radiologist assessments, biopsy data and cancer diagnoses supporting tasks like mass lesion classification [16]. The INbreast dataset from a Portuguese health institute contains over 400 digital mammography exams taken from both screening and clinical workflows across different systems with verified pathology-proven diagnostic labels [17].

For histopathology, the BreakHis dataset from a Brazilian hospital and research institute provides over 7000 microscopic tissue image samples capturing benignconditions or malignant carcinomas across various magnifications [18]. Samples cover the most commonly occurring benign diseases and malignant tumor subtypes with detailed annotations. The BACH challenge dataset compiled from several hospital pathology departments in The Netherlands contains over 400 slides imaged at multiple resolutions covering normal tissue, benign proliferative changes and invasive carcinomas [19].

Table 1. Summary of key benchmark datasets used for evaluation of CNN models on breast cancer screening tasks involving classification of biopsy proven mammography exams or microscopic histopathology image patches.

Dataset	Modality	Cases	Images	Classes
DDSM	Mammography	2,620 exams	10,480 scans	Normal, Benign, Malignant
INbreast	Mammography	115 exams	410 images	Normal, Benign, Malignant
BreakHis	Histopathology	7,909 patches	24,000 patches	Benign, Malignant
BACH	Histopathology	400 slides	30,000 patches	Normal, Benign, Malignant

4.2. Mammogram Analysis Performance

Table 2 compares performance metrics reported from applying different CNN models on mammographic image analysis, including both complete exam classification and patch based lesion evaluation approaches. Fine tuning pretrained ImageNet models like VGG-16 and ResNet-50 on the DDSM dataset to distinguish malignant cancer cases from benign exams or normal tissue achieves over 90% accuracy [12][13]. Models can reach sensitivity levels surpassing 87% indicating strong malignant case detection with reasonable specificity as well. Ensembling predictions from variations of fine-tuned models can further boost performance.

On classification of smaller mass lesion patch samples from mammograms, deeper architectures like DenseNet attain superior accuracy by effectively combining multi-scale features [14]. Pretraining on the natural image domain still provides useful initialization for many medical imaging tasks as models fine-tuned even with modest biomedical data outperform training from scratch. With external Memory Fully Convolutional Network (EM-FCN) frameworks that enable incorporation of global contextual features from full mammogram images rather than just cropped lesions, model sensitivity on malignant mass classification rises considerably without loss of specificity [20]. This suggests holistic contextual reasoning with memory networks can address ambiguity in patch based analysis.

Table 2. Comparison of CNN model evaluation results on mammographic image datasets for tasks ranging from full exam classification to lesion patch categorization with specificity (Sp), sensitivity (Sn) and accuracy (Acc) metrics reported across studies reflecting malignant cancer screening performance.

Model	Task	Dataset	Sp	Sn	Acc
VGG-16	Exam Classification	DDSM	0.94	0.91	0.92
ResNet-50	Mass Detection	INbreast	0.89	0.96	0.93
DenseNet	Lesion Classification	Private set	0.88	0.90	0.89
EM-FCN	Mass Classification	DDSM	0.93	0.97	0.95

4.3 Histopathology Analysis Performance

Table 3 highlights key metrics quantifying CNN model accuracy on breast cancer histopathology image classification using datasets such as BreakHis and BACH containing microscopy tissue samples. Fine-tuned VGG models demonstrate exceptional capacity to distinguish normal ductal cells from invasive malignant carcinomas on high resolution histological patch samples with sensitivity and specificity often above 95% rivaling or exceeding pathologist consistency levels [21][22]. Efficient use of imaging context is again apparent with multiscale Inception architectures learning more generalizable features less susceptible to overfitting, particularly with limited training data [23]. Model fusion ensembles that average predictions from an Inception model trained on image patches and a sequential LSTM network analyzing contextual slide-level labels further optimize global classification accuracy.

Table 3. Comparison of key CNN model evaluation results on breast histopathology image datasets for classification of microscopic tissue sample patches differentiating benign from malignant cancer indications with specificity (Sp), sensitivity (Sn) and accuracy (Acc) metrics highlighted.

Model	Dataset	Sp	Sn	Acc
VGG-16	BreakHis	0.97	0.98	0.98
Inception	BACH	0.94	0.93	0.94
Inception + LSTM	BreakHis	0.96	0.99	0.97

5. Challenges and Future Directions

Though rapid progress has been made in applying CNNs for automated analysis of medical images to transform breast cancer screening, several challenges remain to be addressed. We analyze key issues around class imbalance, model interpretability and data variability that currently limit more widespread clinical adoption of these AI tools. Promising directions like multimodal learning, model compression and improved evaluation protocols are also discussed.

5.1. Class Imbalance

A persistent problem in cancer detection is significant class imbalance where available data has far fewer examples of positive malignant cases compared to abundant normal or benign instances. Models trained on such skewed data can be biased towards predicting the majority negative class and have poor sensitivity detecting outliers. Strategies like oversampling with synthetic minority class augmentation and cost-sensitive re-weighting of loss functions during network optimization have been proposed [24]. But improved imbalance learning mechanisms suited for specificity-critical medical applications are needed.

5.2. Model Interpretability

Most state-of-the-art CNN architectures for visual recognition are complex black-box models with limited transparency into their decision making processes. Model interpretations explaining positive malignant predictions and localizing contributing image regions are required for clinician trust and regulatory approval. Post-hoc methods overlaying heatmaps indicating influential areas have shown promise. Architectures supporting built-in attention mechanisms to compute the relative importance of image features may enable superior integrated explainability [25].

5.3. Data Variability

An ongoing concern impeding real-world viability of deep learning diagnosis tools is variability between training and deployment datasets which can significantly degrade model accuracy if new data falls outside the normalized distribution. Differences in imaging equipment, protocols and patient demographics across healthcare sites make variability inevitable. Improved regularization techniques during training like adversarial domain adaptation have helped address domain shifts to an extent [26]. But platforms supporting efficient model retraining customized for local data profiles will likely be essential for generalizable clinical adoption.

5.4. Multimodal Learning

While tremendous progress has been made in breast cancer screening on individual modalities like mammography and histopathology via CNNs, development of models capable of jointly analyzing correlations across modalities to improve diagnostic yield remains an open challenge. Multi-view frameworks that integrate imaging data with genomic biomarkers show promise towards better risk assessment, subtype classification and prognosis prediction supporting precision treatment [27].

5.5. Model Compression

Transition of high performance deep learning models towards clinical practice necessitates network compression through techniques like pruning, quantization, and distillation for feasible deployment under limited computational budgets for routine inference [28]. Identification of optimal compressed configurations without accuracy loss and frameworks enabling efficient compressed model retraining for continuous improvements with new data samples are important areas for enabling translational adoption.

5.6. Enhanced Evaluation

Rigorous evaluation of model-based decision aids against clinician performance on standardized testing data reflecting the full diversity of challenging cases is vital to establishing efficacy, with ongoing regulatory oversight. Retrospective analysis on archival datasets have limitations. New data curation frameworks like federated learning that enable multi-site model optimization on decentralized data samples with only collective insights shared could provide enhanced evaluation on diverse distributions [29]. Proactive participation by health authorities in auditing performance would also accelerate translation of safe and effective AI tools.

6. Conclusions

In this survey, we have presented a comprehensive review of deep CNN architectures applied to automate breast cancer detection using mammographic and histopathologic images. Standard models pretrained on natural images like VGGNet and ResNet demonstrate exceptional capacity to distinguish malignancy when fine-tuned on medical imaging datasets. Unique architectures like DenseNet more effectively integrate multi-scale features to match or even surpass clinician performance on challenging diagnostic tasks. Ongoing research has also produced breakthroughs in critical areas like contextual reasoning and multimodal fusion to further advance AI capability towards expert radiologist-level precision.

However, real-world clinical deployment necessitates addressing key challenges around model transparency, data variability, validation rigor and computational efficiency. Promising solutions stemming from fields like explainable AI, domain adaptation and compression offer potential pathways for translational success. Ultimately, close collaboration between medical and machine learning communities is vital to develop safe and trustworthy AI tools that enhance clinician workflows and augment human intelligence for improved patient care. Systems co-designed with frontline workers emphasizing strengths of both humans and machines could enable personalized delivery enhancing outcomes.

With incredible pace of progress in AI algorithmic capability and computational infrastructure, CNNs are poised to transform breast cancer screening in the coming decade if open questions around reliability and integration can be effectively tackled. This necessitates shared public-private investment prioritizing patient wellbeing over profit with proactive regulatory leadership steering translation. Widespread adoption of deep learning systems can relieve overloaded screening programs, cut down misdiagnosis and reduce disparities expanding life-saving early detection. Overall, this new wave of AI marks a watershed moment bringing enhanced decision support capabilities that promise enormous clinical impact if responsibly mobilized. We hope this extensive review of pioneering breast cancer detection research efforts provides both machine learning experts and practicing clinicians with insights to accelerate development of patient-centered AI solutions towards improving the lives of countless women around the world.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. 2018, 68, 394-424.

2. Lauby-Secretan, B.; Scoccianti, C.; Loomis, D.; Benbrahim-Tallaa, L.; Bouvard, V.; Bianchini, F.; Straif, K. Breast-cancer screening--viewpoint of the IARC working group. New Engl. J. Med. 2015, 372, 2353-2358.
3. Siu, A.L. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann. Intern. Med. 2016, 164, 279-296.
4. Fenton, J.J.; Taplin, S.H.; Carney, P.A.; Abraham, L.; Sickles, E.A.; D'orsi, C.; Berns, E.A.; Cutter, G.; Hendrick, R.E.; Barlow, W.E. Influence of computer-aided detection on performance of screening mammography. New Engl. J. Med. 2007, 356, 1399-1409.
5. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436-444.
6. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Adv. Neural Inf. Process Syst. 2012, 25, 1097-1105.
7. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2015, arXiv:1409.1556.
8. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proc IEEE Conf Comp Vis Pattern Recogn 2016, 770-778.
9. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proc IEEE Conf Comp Vis Pattern Recogn 2018, 4700-4708.
10. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proc IEEE Conf Comp Vis Pattern Recogn 2015, 1-9.
11. Gao, Y.; Geras, K.J.; Lewin, A.A.; Moy, L.; Cho, K.; Chui, S.Y.; Sohn, J.H.; Ke, Q.; Stumpe, M.C. New frontiers: Application of deep learning artificial intelligence in breast imaging. Clin. imaging 2020, 60, 305-316.
12. Al-antari, M.A.; Almubarak, H.A.; Alfadhel, F.; Tan, J.W.; Aljarbou, A.; Zhang, M. A computer-aided diagnosis system for breast density classification based on transfer learning. Artif. Intell. Med. 2021, 114, 102006.
13. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv 2017, arXiv:1711.05225.
14. Li, H.; He, J.; Ni, B. Breast ultrasound image classification based on a deep convolutional neural network. Clin. breast cancer 2019, 19, e641-e647.
15. Kooi, T.; Van Ginneken, B.; Karssemeijer, N.; Den Heeten, A. Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural network. Med. Phys. 2017, 44, 1017-1027.
16. Heath, M.; Bowyer, K.; Kopans, D.; Moore, R.; Kegelmeyer, P. The digital database for screening mammography. In Proc Int Workshop on Digital Mammography, 212-218.
17. 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. 2012, 19, 236-248.

18. Spanhol, F.A.; Oliveira, L.S.; Petitjean, C.; Heutte, L.; Silva, A.C.; de Alencar Rocha, A. Breast cancer histopathological image classification using convolutional neural networks. In Proc Int Joint Conf Neural Net 2016, 2560-2567.
19. Bejnordi, B.E.; Veta, M.; Van Diest, P.J.; Van Ginneken, B.; Karssemeijer, N.; Litjens, G.; Van Der Laak, J.A. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 2017, 318, 2199-2210.
20. Zhou, T.; Han, Z.; Singh, V.; Zhang, Y.; Bahl, S.; Catanzaro, B.; Cubuk, E.D. Deepvit: Towards deeper vision transformer without stochastic attention. arXiv 2021 arXiv:2103.11886.
21. Li, X.; Yu, L.; Chen, H.; Fu, C.; Heng, P.A. Semi-supervised breast cancer histopathological image classification using deep multi-scale convolutional network. Ieee J Biomed Health 2021 35, 1139-1147.
22. Wang, J.; Wu, C.J.; Bai, X.; Yu, Z. Diagnosis of breast cancer using deep learning: A taxonomical review. Comp Methods Biomech Biomed Eng Imaging Vis 2020, 8, 713-722.
23. Bychkov, D.; Linder, N.; Turkki, R.; Nordling, S.; Kovanen, P.E.; Verrill, C.; Walliander, M.; Lundin, M.; Haglund, C.; Lundin, J. Deep learning based tissue analysis predicts outcome in colorectal cancer. Sci Rep 2018, 8, 3395.
24. Johnson, J.M.; Khoshgoftaar, T.M. Survey on deep learning with class imbalance. J Big Data 2019 6, 1-54.
25. Holzinger, A.; Biemann, C.; Pattichis, C.S.; Kell, D.B. What do we need to build explainable AI systems for the medical domain? arXiv 2017 arXiv:1712.09923.
26. Kamnitsas, K.; Baumgartner, C.; Ledig, C.; Newcombe, V.; Simpson, J.; Kane, A.; Menon, D.; Nori, A.; Criminisi, A.; Rueckert, D. Unsupervised domain adaptation in brain lesion segmentation with adversarial networks. In Proc Int Conf Inf Process Med Imaging, 597-609.
27. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sanchez, C.I. A survey on deep learning in medical image analysis. Med Image Anal 2017, 42, 60-88.
28. Gholami, A.; Kwon, G.; Wu, B.; Tai, Z.; Yue, X.; Jin, P.; Zhao, S.; Kim, K.I.; Wong, S.T. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and< 1mb model size. arXiv 2016 arXiv:1602.07360.
29. Rieke, N.; Hancox, J.; Li, W.; Milletari, F.; Roth, H.R.; Albarqouni, S.; Bakas, S.; Galtier, M.N.; Landman, B.A.; Maier-Hein, K. The future of digital health with federated learning. NPJ Digit Med 2020, 3, 1-7.