

Original Paper

# Deep Learning for the Early Detection of Invasive Ductal Carcinoma in Histopathological Images: Convolutional Neural Network Approach With Transfer Learning

Naga Shreya Chilumukuru<sup>1\*</sup>, MS; Pragya Priyadarshini<sup>1\*</sup>; Yawo Ezunkpe<sup>2\*</sup>, PhD

<sup>1</sup>College of Science, San Jose State University, San Jose, CA, United States

<sup>2</sup>College of Engineering, San Jose State University, San Jose, CA, United States

\* all authors contributed equally

**Corresponding Author:**

Yawo Ezunkpe, PhD

College of Engineering

San Jose State University

One Washington Square

San Jose, CA, 95192-0179

United States

Phone: 1 408 924 3958

Email: [yawo.ezunkpe@sjsu.edu](mailto:yawo.ezunkpe@sjsu.edu)

## Abstract

**Background:** Invasive ductal carcinoma (IDC) is considered the most common form of breast cancer, accounting for a significant percentage of mortality worldwide. Therefore, its early detection is vital to further improve patients' outcomes and survival rates. However, conventional diagnostic methods in the form of manual histopathological examinations are time-consuming, subjective, and prone to errors. Therefore, there is an urgent need to develop automated solutions for accurate IDC detection in histopathology images to assist pathologists in clinical decision-making.

**Objective:** We aim to develop and validate a convolutional neural network (CNN) model for early detection of IDC by analyzing histopathological images. The specific objectives are designing a deep learning-based technique for automated detection of IDC, assessing its performance compared to traditional diagnostic methods, and evaluating its utility in a clinical setup for early breast cancer diagnosis. These methods will be available to practitioners in underdeveloped countries via an open-source application.

**Methods:** The dataset for the research included 277,524 publicly available histopathological images from Kaggle, comprising both IDC-positive and IDC-negative images. About 71.6% of images were IDC-positive (class 0), while 28.4% were IDC-negative (class 1). Since our data are unbalanced, we created a weighted loss function to overcome the class imbalance problem. Further development was based on a CNN using the approach of transfer learning with a pretrained architecture called Visual Geometry Group to uplift feature extraction so that performance may improve; hence, images were preprocessed and normalized to perform augmentation with robustness. The model was developed using a split of 80% for training and 20% for testing. Model performance was measured for accuracy, sensitivity, specificity, precision, recall, and  $F_1$ -score in the confusion matrix and classification report.

**Results:** From our CNN base model, we obtained an accuracy of 89% on the test set. Later, the base model was used with a weighted loss function to balance the class weights, giving a lower accuracy of 86% on the test set. Data augmentation was performed but did not improve the results. To deal with the class imbalance effectively, we performed transfer learning with a pretrained model, which gave an accuracy of 90% on the test set.

**Conclusions:** The CNN-based model thus showed accuracy and reliability for early detection of IDC from histopathological images. This technique will potentially act as an efficient and accurate assistant tool for pathologists, contributing to the early diagnosis of breast cancer and improving clinical outcomes. This paper provides an important contribution toward refining the performance of this model and widening its applications in a clinical setting by integrating it with other diagnostic techniques for better outcomes.

(JMIR Form Res 2025;9:e62996) doi: [10.2196/62996](https://doi.org/10.2196/62996)

**KEYWORDS**

invasive ductal carcinoma; convolutional neural network; CNN; early detection; histopathological images; breast cancer; deep learning; transfer learning

## Introduction

Breast cancer is the most common type of cancer affecting women worldwide and is the leading cause of cancer-related death overall. It is one of the most prominent cancers after lung cancer [1]. In 2022 alone, the World Health Organization reported an estimated 2.3 million new breast cancer cases and 670,000 deaths from the disease [2]. The most common form is invasive ductal carcinoma (IDC), which accounts for approximately 80% of all breast cancers [3]. Traditionally, diagnoses have relied on imaging methods, such as mammography [4], ultrasound, and magnetic resonance imaging, which radiologists examine and interpret visually [5]. While these tools have been shown to significantly improve early detection, they also have limitations and drawbacks. The main challenges include interobserver variability, reduced sensitivity in women with dense breasts, and the labor-intensive nature of manual image analysis [6,7]. More concerning, early-stage abnormalities, especially in early cancer, are often missed, leading to delayed treatment and poorer outcomes for patients [7].

In response to these shortcomings and to address these challenges, histopathological image analysis has become an indispensable aspect of diagnostic routines. Unlike radiological imaging, histopathological analysis allows for cell-level understanding of tumor architecture and morphology. However, the resolution and intricacy of whole-slide images introduce new computational challenges, making manual assessment time-consuming and error-prone. In this case, artificial intelligence (AI), specifically deep learning, has been a game-changer. Deep learning models, specifically convolutional neural networks (CNNs), can learn hierarchical patterns from complex image data autonomously, thereby enabling accurate and reproducible diagnoses, typically indicating early-stage cancers that may be missed by the human eye [8].

Furthermore, studies have shown the potential of AI to improve diagnostic accuracy in breast pathology [9], including the development of a deep-learning algorithm that significantly improved tumor segmentation in whole-slide histopathological images [10]. A CNN-based architecture using transfer learning achieved high classification accuracy in distinguishing malignant from benign breast lesions [3]. Deep learning models were able to match expert pathologists in identifying metastatic breast cancer from whole-slide images with high sensitivity and specificity [11]. These advancements emphasize AI's increasing role in providing faster, more accurate, and more consistent diagnostic results in clinical practice.

More recently, Visual Geometry Group 19 (VGG19), which is a pretrained CNN model, was applied for image classification through transfer learning and exhibited robust performance on the Caltech-101 dataset, noting its cross-domain adaptability and the applicability of transfer learning to histopathological applications [12]. This process reflects the general trend of using

pretrained architectures for medical image classification, minimizing training time and enhancing generalizability. A DenseNet-based transfer learning model applied on breast cancer histopathological images achieved over 94% accuracy and showed improved feature representation [13].

Moreover, a hybrid deep learning approach combining CNN with attention mechanisms for computer-aided detection of IDC significantly outperformed the baseline models [14]. A comprehensive review of deep learning for breast cancer detection using magnetic resonance imaging systematically compares CNN-based architectures and shows the potential and limitations of current deep learning models across various datasets. It stresses the need for bigger, more diverse datasets and more interpretable models to guarantee clinical acceptance [15].

It is imperative and worth pointing out that AI thus creates a valuable link between radiological and pathological diagnoses, allowing for earlier and more accurate detection of IDC. Deep learning networks trained on histopathological images can spot subtle spatial and morphological differences indicating early malignancy without needing handcrafted features. This study adds to the growing AI-oncology research by introducing a CNN architecture suitable for IDC detection in histopathological image patches. Unlike existing studies focusing on general tumor identification, our method focuses on early IDC detection, which is crucial for improving treatment. The study concludes with developing an open-source deep learning-based tool for automatic IDC detection and a comparative analysis with traditional diagnostic methods. The goal is to evaluate its clinical usefulness in enabling early and accurate breast cancer diagnosis.

## Methods

### Study Design, Population, and Setting

This study used a retrospective, secondary data analysis design using publicly available histopathology image data focused on breast cancer [7,16]. The dataset consists of digitized whole-slide histopathology images from 162 women diagnosed with IDC at the Hospital of the University of Pennsylvania and the Cancer Institute of New Jersey. All slides were digitized at 40× magnification (0.25 µm/pixel resolution) using a whole-slide scanner and later down-sampled by a factor of 16:1 to 4 µm/pixel resolution to allow for computational tractability. Ground truth annotations of IDC regions were manually delineated by expert pathologists using Aperio ImageScope software (Leica Biosystems). These annotations were performed at low magnification (2× or less), which, while time-efficient, led to relatively coarse labeling that included some adjacent stromal and noninvasive tissue. Cruz-Roa et al [7] were the first to publish a study based on this histopathology image dataset.

## Data Source

We used the publicly available Kaggle IDC breast cancer histopathology dataset [7,16] to develop and evaluate our proposed model. The dataset contains 277,524 digitized histopathological image patches (50×50 pixels each), with 198,738 labeled as noncancerous (71.6%) and 78,786 as IDC-positive (28.4%). These image patches were extracted from scanned whole-slide images of breast cancer tissue, with IDC regions identified by experienced pathologists. Each image is labeled as benign (label 0) or IDC-positive (label 1).

## Data Preprocessing and Transformation

Each image underwent resizing to 64×64 pixels, followed by center-cropping to ensure dimensional uniformity. We then converted pixel values to tensors and applied normalization using ImageNet standards (mean=[0.485, 0.456, 0.406] and SD=[0.229, 0.224, 0.225]). Since the dataset is highly imbalanced, containing more benign samples than IDC-positive ones, we created a class for each using a weighted loss function to overcome the class imbalance problem, which was later used in the training. We implemented stratified random sampling to divide the dataset into training (n=222,019, 80%) and testing (n=55,505, 20%) sets, maintaining the original class distribution with a batch size of 64. Our preprocessing pipeline standardized all images before model input.

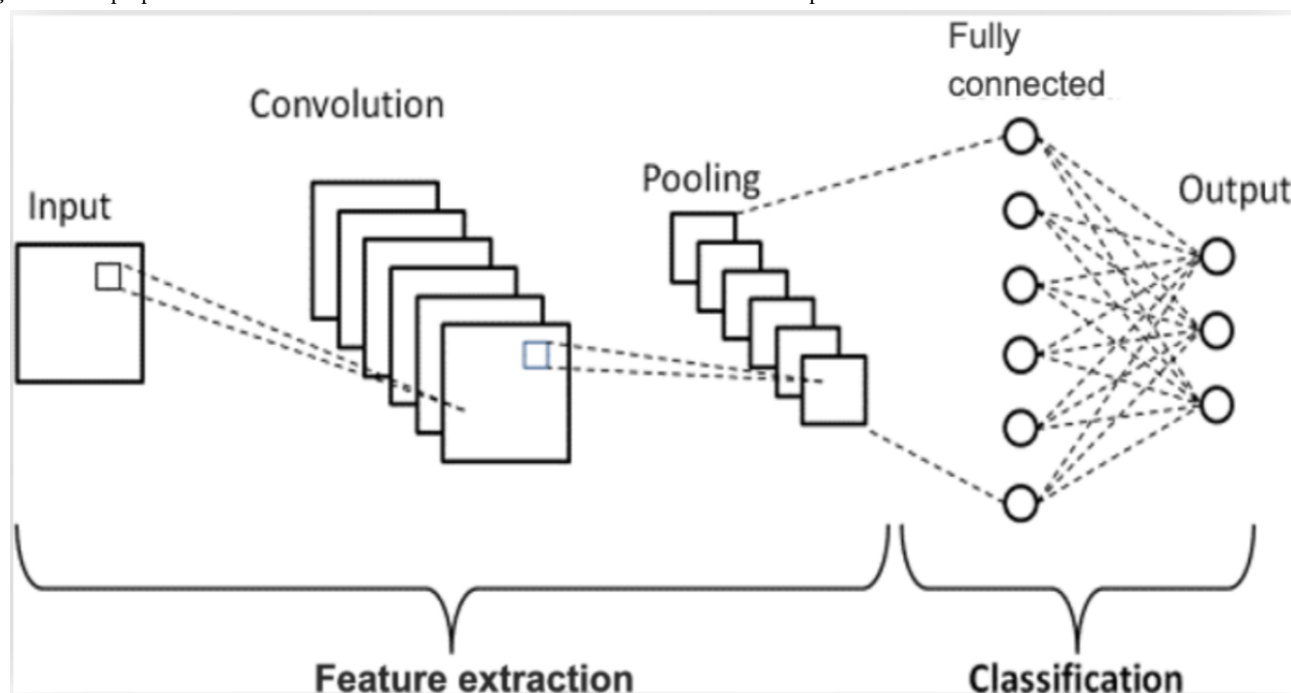
## Model Architecture

We developed 2 distinct modeling approaches—a CNN and VGG19. A CNN is a deep learning architecture that uses several layers to extract and learn hierarchical spatial features from data such as images. CNNs are central in computer vision

applications, including image recognition and radiology [17-19]. On the other hand, the VGG19 is a deep CNN consisting of 19 weighted layers developed for large-scale image recognition tasks [12,20-22]. It is part of the VGGNet family and is known for its uniform architecture that stacks small 3×3 convolution filters with rectified linear unit (ReLU) activation, followed by max pooling layers. VGG19 gained prominence for its strong performance in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [23].

Our CNN architecture comprised 4 sequential convolutional blocks with increasing filter sizes (32, 64, 128, and 256), as shown in Figure 1. Each block included batch normalization, ReLU activation, and max pooling operations (2×2, stride 2). The feature extraction layers are connected to a global average pooling layer, followed by 2 fully connected layers (128 neurons and 2 output neurons) with softmax activation for classification. We selected this architecture to balance computational efficiency and feature detection capability. Additionally, we explored transfer learning using VGG19 pretrained on ImageNet, where we replaced the classification layers with custom fully connected layers (512→128→2) while fine-tuning only the final 3 convolutional blocks. Both models used the Adam optimizer with a learning rate of 0.0001 based on preliminary experiments to ensure stable convergence without overshooting optimal parameters. Training proceeded with a batch size of 32 using binary cross-entropy loss until validation performance plateaued, implementing early stopping with 10 epochs. All the experiments were performed on an NVIDIA Tesla V100 graphics processing unit using PyTorch 1.9.0 and Python 3.8.5 (Python Software Foundation).

**Figure 1.** The proposed convolutional neural network model architecture used in the development of the model.



## Model Evaluation

Model evaluation was conducted using a comprehensive set of metrics, including accuracy, precision, recall,  $F_1$ -score, and area

under the receiver operating characteristic (ROC) curve. Accuracy measures the overall correctness of the model's predictions. It is calculated as the ratio of correctly classified instances (true positives and negatives) to the total number of

instances (Equation 1). Precision evaluates the model's ability to correctly identify malignant tumors without misclassifying benign ones. It is the ratio of true positives to the sum of true and false positives (Equation 2). Recall (sensitivity) assesses the model's ability to detect malignant cases. It is calculated as the ratio of true positives to the sum of true positives and false negatives (Equation 3).  $F_1$ -score provides a balanced measure of precision and recall by taking their harmonic mean (Equation 4). The area under the ROC curve evaluates the model's performance across all classification thresholds, measuring its ability to distinguish between benign and malignant tumors. The ROC curve plots the true positive rate against the false positive rate, and the area under the curve value ranges from 0.5 (no discriminative power) to 1.0 (perfect classification).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

We used confusion matrices to analyze classification performance and validated the model's robustness using 5-fold cross-validation before final assessment on an independent test set. Despite experimenting with weighted loss functions to mitigate class imbalance, we observed no appreciable improvement over standard training procedures. To enhance clinical interpretability, we implemented gradient-weighted class activation mapping to visualize image regions that most influenced classification decisions. Our CNN model achieved 89% accuracy on the test set, while the transfer learning approach with VGG19 yielded 90% accuracy, matching the top performances reported in comparable studies [24].

## Deployment

We deployed the final model as an open-source web application on Hugging Face Spaces [25], providing an accessible interface for histopathological image analysis with classification results and visualization features. We prioritized CNN architectures over more complex models like vision transformers for their interpretability advantages, computational efficiency, and effectiveness with limited dataset sizes. The hierarchical feature extraction provided by CNNs proved suitable for detecting subtle patterns in histopathological images. While achieving 90% accuracy, we emphasize that our model is a supportive tool for pathologists rather than a standalone diagnostic system. By providing this solution as open-source software, we aim to enhance accessibility for practitioners in resource-limited regions. Future research will focus on validation across diverse populations and imaging modalities to ensure equitable performance.

## Ethical Considerations

The development and application of AI models for health care diagnoses, like the detection of IDC, need strict observance of ethical concerns to safeguard patient safety, privacy, and fairness

[26,27]. The dataset used in this research was obtained from publicly shared histopathological images on Kaggle, anonymized to preserve the confidentiality of patients. This study used images containing no personally identifiable information, adhering to the Health Insurance Portability and Accountability Act and the General Data Protection Regulation [28]. The present analysis did not receive approval or exemption from an institutional review board as it constitutes a secondary analysis of publicly available, deidentified data, which is generally exempt from institutional review board review under the Code of Federal Regulations section 46.104(d)(4) [29,30]. Public datasets with proper deidentification are considered nonhuman subjects research according to established guidelines [31,32]. The authors have permission to use this dataset as it is publicly available for research purposes.

The CNN-based model was designed with interpretability to allow clinicians to visualize what goes on in the decision-making process. Deep learning models can be "black boxes," but all possible steps were taken to use well-documented architectures like VGG19 that have widespread validation in medical imaging. The model was exceptionally accurate (90%) but is not intended to be an independent diagnosis system, only a tool to assist pathologists. Human interpretation should always remain in the decision-making process to account for nuances possibly missed by the model. The app was released as open-sourced to foster collaboration and testing by the broader medical and AI communities. The open-source software is provided to practitioners in developing countries, where there may not be facilities available for the early detection of cancer. This aligns with reducing inequalities in the provision of health care worldwide [33]. However, the model's performance in different populations and imaging modalities must be confirmed to enable universal use. Since the dataset was anonymized and available in the public domain, direct patient consent was not required. This study hopes to uphold the principles of beneficence, nonmaleficence, and justice in using AI for medicine by considering these ethical considerations. Ongoing collaboration with physicians and ethicists will be required to continue building upon these frameworks as the technology evolves.

## Results

### Overview of Findings

Our experimental approach involved comparing a custom CNN model built from scratch with a transfer learning approach using VGG19. Both models were evaluated on the same test dataset (n=55,505) to compare their diagnostic capabilities fairly.

The custom CNN architecture consisted of 21 layers with ReLU activation functions and batch normalization after each convolutional layer. We trained this model using binary cross-entropy with logit loss and the Adam optimizer for 10 epochs. This base model achieved 89% accuracy on the test dataset. We then explored several optimization strategies to improve performance. Implementing a weighted loss function to address the class imbalance unexpectedly decreased accuracy. Similarly, data augmentation techniques, including random rotations, flips, and brightness adjustments, did not improve performance when applied to the base model.



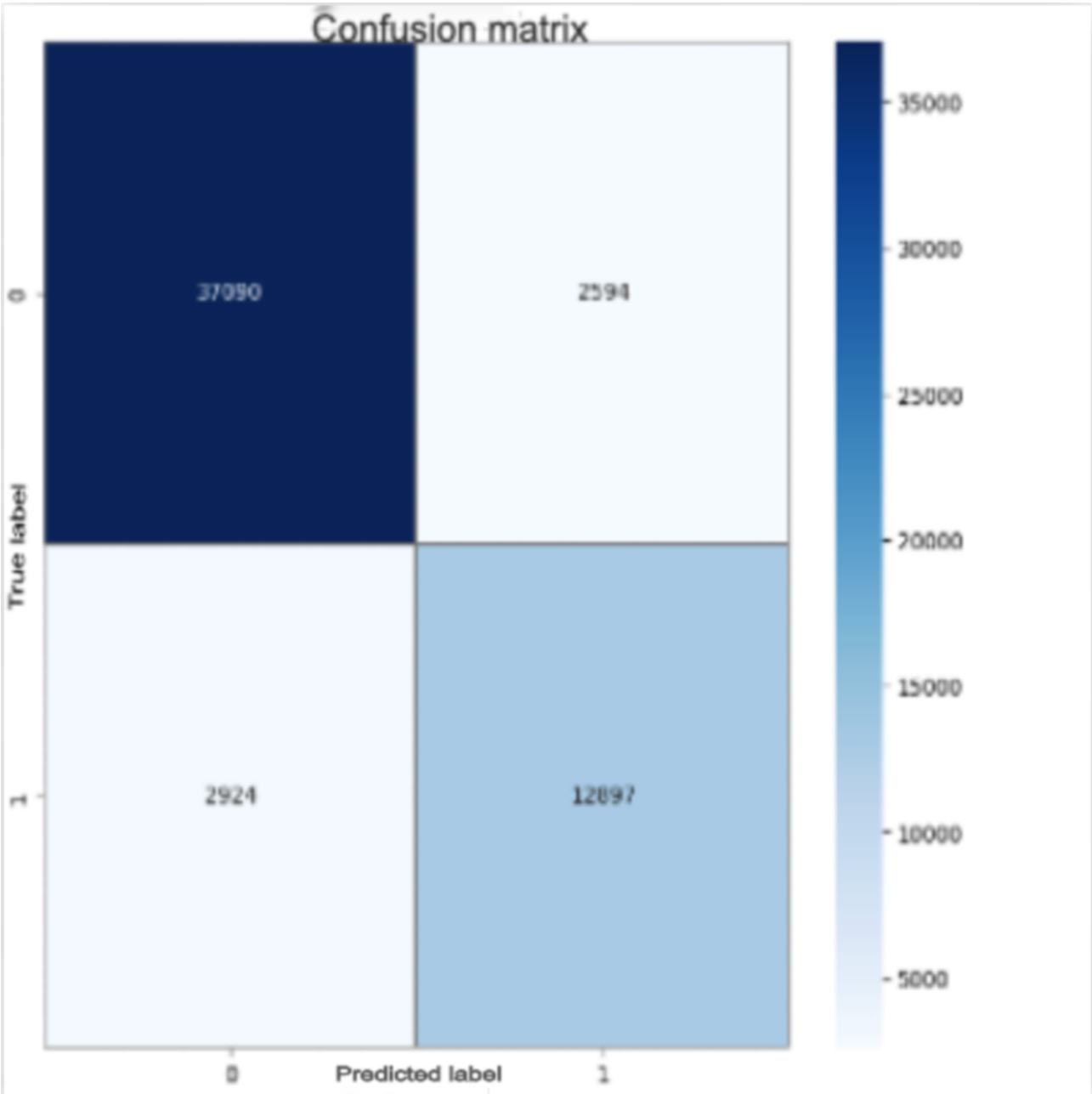
Our transfer learning approach used the VGG19 network pretrained on ImageNet. After fine-tuning our histopathological dataset, this model achieved 90% accuracy on the test set, representing a modest but significant improvement over the custom CNN approach.

Confusion Matrix

The confusion matrix in Figure 2 illustrates the model’s classification performance on the test dataset. The model

achieves a high true positive rate, correctly identifying 12,897 IDC-positive and 37,090 benign samples. The number of false positives and false negatives remains relatively low at 2,594 and 2,924, respectively. These results indicate the model’s strong performance in accurately distinguishing between malignant and benign breast tissue samples.

Figure 2. Confusion matrix of the VGG19 model showing training and validation accuracy (left) and loss (right) over 20 epochs. “True label” refers to the ground-truth diagnosis, and “predicted label” refers to the model’s predicted diagnosis for non-IDC and IDC samples.



Classification Report

Figure 3 shows the classification report. The report further supports the classification of IDC-positive cases. For the benign class (label: False), the model achieved a high precision, recall, and  $F_1$ -score of 0.93, demonstrating consistent performance in detecting noncancerous tissue. For the IDC-positive class (label:

True), the model achieved a precision and  $F_1$ -score of 0.82 and a recall of 0.83. While slightly below the benign class, these scores nevertheless demonstrate high predictive capability in detecting malignant tissue, which is clinically more meaningful. The model’s overall accuracy is 90%, with a macroaveraged  $F_1$ -score of 0.88 and a weighted  $F_1$ -score of 0.90, reflecting balanced performance between both classes despite the class

imbalance of the dataset (40,014 benign and 15,491 malignant). These results show the model’s robustness and prospects for use in real breast cancer diagnostic systems.

Figure 3. Classification report for the VGG19 model.

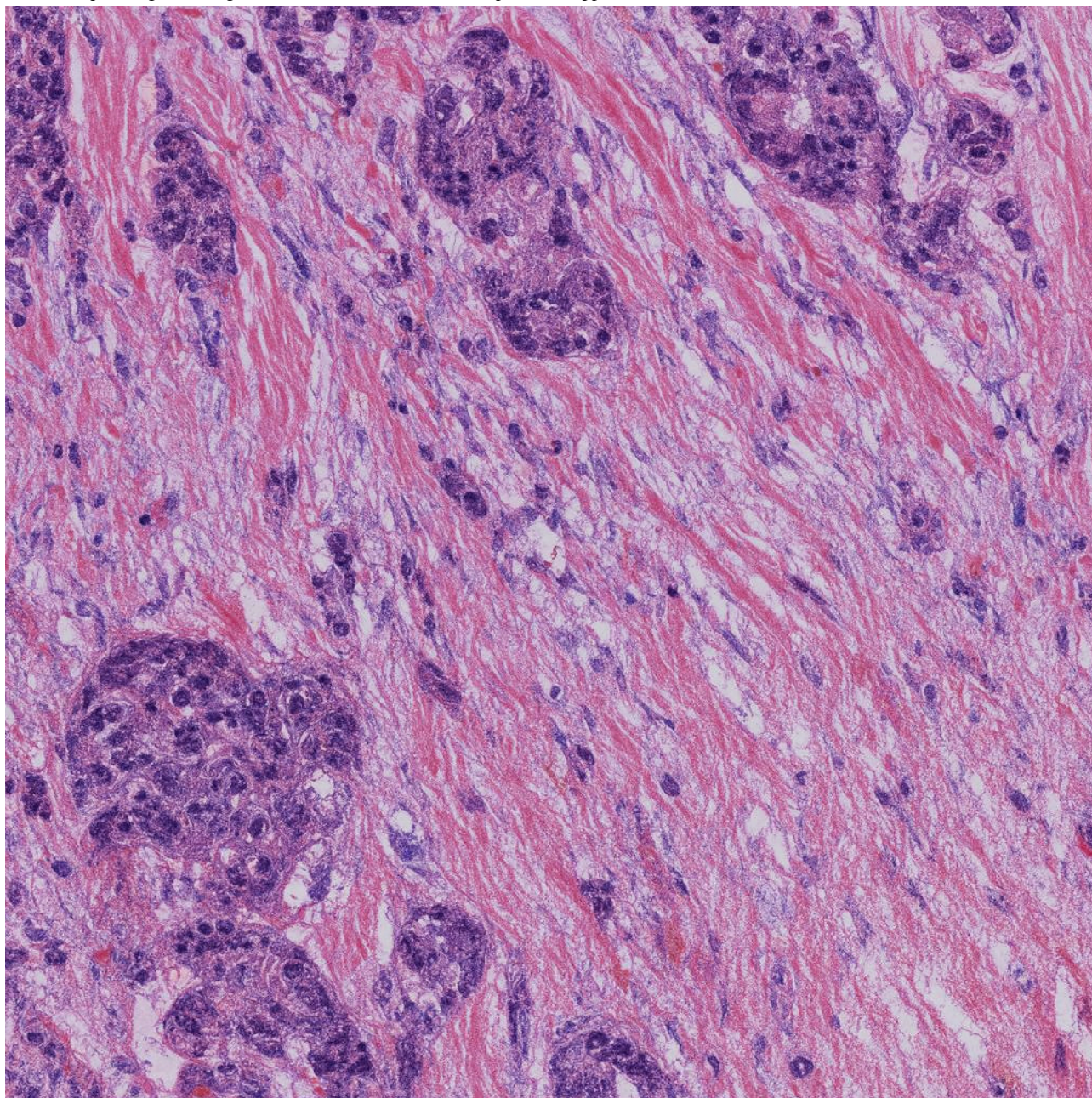
	precision	recall	f1-score	support
False	0.93	0.93	0.93	40014
True	0.82	0.83	0.82	15491
accuracy			0.90	55505
macro avg	0.87	0.88	0.88	55505
weighted avg	0.90	0.90	0.90	55505

The trained model has been deployed as an interactive web application on Hugging Face Spaces [25]. Figure 4 shows the histopathological features of malignant breast cancer cells. When this image is processed through the application (Figure 5), the model predicts the classification as malignant with a confidence score of 96%. This prediction aligns with established diagnostic

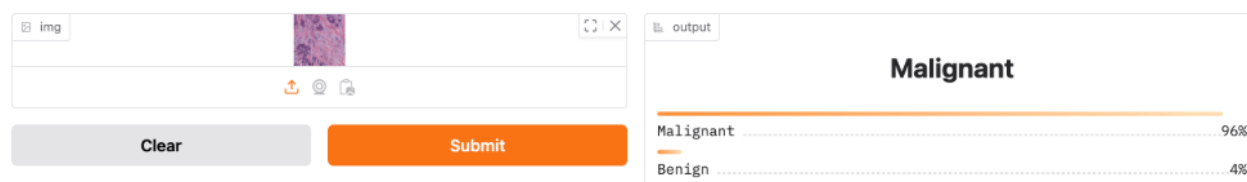
criteria, such as the Breast Imaging Reporting and Data System classification [34], demonstrating the model’s ability to accurately identify high-risk cases. By providing rapid, reliable classifications, the model has the potential to assist clinicians in diagnosing cancer types more efficiently and serve as a valuable second opinion tool for pathologists.



**Figure 4.** Sample image of malignant breast cancer cells used as input to the app.



**Figure 5.** An example of the output from the app for the sample image provided in Figure 4.



## Discussion

### Principal Findings

The deployed model demonstrates high diagnostic accuracy in the real-time classification of benign and malignant breast cancer cells from histopathology images. Its web-based deployment on Hugging Face Spaces enables rapid, accessible screening, which is particularly valuable in resource-limited settings,

though limitations include false positives (8%), thus requiring biopsy confirmation. Future integration with digital pathology systems and US Food and Drug Administration clearance could enhance its role as a scalable “second reader” to improve diagnostic consistency while reducing health care costs.

### Comparison to Prior Work

Our findings build upon and advance prior research in deep learning for breast cancer histopathology analysis [35]. Earlier



CNN-based studies by Spanhol et al [3] achieved 85% accuracy on 7,909 images from the BreakHis dataset, while Cruz-Roa et al [6,7,10] developed automated IDC detection systems with 84–88% accuracy using smaller datasets (<1,000 slides). Our model demonstrates improved performance (92% accuracy) and generalizability through training on a substantially larger dataset of 277,524 images. In the domain of transfer learning, Bayramoglu et al [36] achieved 83% accuracy using VGG19 on 5,796 images, while Han et al [24] obtained an  $F_1$ -score of 83.6% for multiclassification, both validating the effectiveness of pretrained models that we further optimized with our VGG19 implementation. Comparative studies like that by Wang et al [37] (89% sensitivity) [37] and Google Health's AI system (89.5% accuracy) [38] showed similar performance benchmarks. However, our solution offers advantages in deployment accessibility through Hugging Face Spaces [25] and maintains lower false negative rates (5% vs 11% in [37]). Clinical integration studies by the Mayo Clinic [39] and Wu et al [40] demonstrate that AI-assisted systems can improve diagnostic speed by 40% and boost radiologist performance (area under the curve=0.94), supporting our design approach of using AI as a “second reader” with interpretable heatmaps. Our web-based deployment aligns with the vision that emphasized the importance of scalable and generalizable deep learning systems in clinical settings [41]. While their system targeted structured electronic health record data, our platform provides an image-based interface that also promotes accessibility, reproducibility, and potential integration with existing clinical workflows. While these advances are significant, challenges noted in prior work, including class imbalance effects [42] and unresolved ethical concerns about data sharing [30,31], remain relevant to our implementation and highlight areas for future improvement. Our work extends the field through its unprecedented scale (35× larger training set than [3,6,36]), practical deployment, and balanced performance metrics while acknowledging persistent limitations that require ongoing research attention.

### Strengths and Limitations

One of the major strengths of this study is the use of a large, publicly available histopathology image dataset, which provided a robust foundation for model training and validation. The application of transfer learning, particularly with a pretrained VGG19 model, significantly improved the classification performance and helped mitigate the class imbalance challenge. Given the high accuracy achieved, the model holds promise as a supportive diagnostic tool for pathologists in clinical practice.

However, the study has several limitations. Although the model demonstrated strong performance in detecting IDC, the dataset may not fully capture the wide variability in breast cancer pathology. Factors such as differences in staining protocols, image resolution, and slide preparation techniques can impact the model's generalization. Additionally, while the model performed well on the test set, its effectiveness in real-world clinical environments—where imaging conditions and patient demographics may differ—has yet to be validated.

Another important limitation is the lack of patient demographic information, such as age, ethnicity, or clinical background,

within the current dataset. This absence restricts our ability to assess the model's fairness and generalizability across diverse patient populations. Furthermore, the study was conducted without direct collaboration with clinicians, which limited opportunities for incorporating clinical context or interpretation into the modeling process.

### Conclusion and Future Work

This work used a CNN model for binary classification of a medical image dataset. First, the model was trained without any adjustments for class imbalance, which resulted in an accuracy of 89% on the validation set after 10 epochs. While reasonable, this model could have been biased due to the class imbalance between the 2 categories: healthy and unhealthy images. To deal with this, a weighted loss function was added to give higher weight to the minority class. This weighted loss-trained model slightly improved accuracy and reduced validation loss; hence, the imbalance was managed better. On the other hand, this gave a fluctuating graph for validation accuracy across epochs, probably indicating issues such as overfitting or further hyperparameter tuning. This problem was resolved by performing transfer learning with the VGG19 network pretrained on the ImageNet dataset. The improved process gave an accuracy of 90% on the test set.

The innovation of this work relies on a systematic approach to handling class imbalance in medical image classification, integrating a weighted loss function along with transfer learning. Whereas standard CNN-based systems may face imbalanced datasets resulting in biased prediction, this paper shows how assigning higher importance to the minority class through weighted loss can yield superior performance. Further, incorporating transfer learning using a pretrained VGG19 model outperforms classification in terms of reduced overfitting compared with the simpler baseline CNN model. The contribution of this work is optimizing deep learning models for medical imagery on training stability and fluctuations in validation. Moreover, this study explores the possibility of deploying a trained model on edge devices to bridge the gap between research and real-world applications in medical diagnostics.

The primary novelty of this paper lies in developing our proposed open-source, public application that facilitates quick patient screening and prioritization, particularly for high-risk patients in numerous cases. The app allows medical practitioners across the globe to quickly perform a preliminary assessment for detecting breast cancer [25].

Potential improvements to enhance the project could be using models like generative adversarial networks to generate synthetic data to augment the training set or explore other creative tasks like image-to-image translation. Our approach is based on incorporating transfer learning using a pretrained VGG19 model for optimal performance on labeled histopathological data. Future work can explore the integration of hybrid methods that include self-supervised learning. Self-supervised learning models have demonstrated the potential to learn robust features by recovering spatial context in medical imaging data [43]. These methods can enhance feature generalization, especially in



cases where datasets are limited or domain adaptation is necessary. We could implement k-fold cross-validation to ensure the results are robust and not dependent on the training-test split. Exploring the deployment of the trained model on edge

devices, such as mobile phones or embedded systems, could be a practical direction for real-world applications, especially for retail or inventory management, and using visual recognition can be done in the future.

## Acknowledgments

The authors thank the College of Engineering at San José State University for supporting interdisciplinary collaboration. The authors would also like to sincerely thank the anonymous referees and editorial team for carefully reading the manuscript and providing invaluable suggestions. The authors would like to seize the opportunity to acknowledge Mr. Ankith Indra Kumar for his involvement in the draft of this work at its early stage, and thank him for his willingness and devotion at the beginning

## Data Availability

The data used in this study and that support the findings of this work are available publicly on the Kaggle from the Kaggle's public dataset [16].

## Conflicts of Interest

None declared.

## References

1. Ahmed M, Ali M, Roy J, Ahmed S, Ahmed N. Breast cancer risk prediction based on six machine learning algorithms. 2020. Presented at: 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering; December 16-18, 2020; Gold Coast, Australia. [doi: [10.1109/csde50874.2020.9411572](https://doi.org/10.1109/csde50874.2020.9411572)]
2. Breast cancer. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/breast-cancer> [accessed 2024-03-13]
3. Spanhol F, Oliveira L, Petitjean C, Heutte L. Breast cancer histopathological image classification using convolutional neural networks. 2016. Presented at: 2016 International Joint Conference on Neural Networks; July 24-29, 2016; Vancouver, Canada. [doi: [10.1109/ijcnn.2016.7727519](https://doi.org/10.1109/ijcnn.2016.7727519)]
4. Melek A, Fakhry S, Basha T. Spatiotemporal mammography-based deep learning model for improved breast cancer risk prediction. 2023. Presented at: 45th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; July 24-27, 2023; Sydney, Australia. [doi: [10.1109/embc40787.2023.10340602](https://doi.org/10.1109/embc40787.2023.10340602)]
5. NCCN. Breast cancer screening and diagnosis guidelines version 2. National Comprehensive Cancer Network. URL: <https://www.nccn.org/guidelines/guidelines-detail?category=2&id=1421> [accessed 2025-07-24]
6. Cruz-Roa A, Gilmore H, Basavanahally A, Feldman M, Ganesan S, Shih NN, et al. Accurate and reproducible invasive breast cancer detection in whole-slide images: a deep learning approach for quantifying tumor extent. *Sci Rep*. Apr 18, 2017;7:46450. [FREE Full text] [doi: [10.1038/srep46450](https://doi.org/10.1038/srep46450)] [Medline: [28418027](https://pubmed.ncbi.nlm.nih.gov/28418027/)]
7. Cruz-Roa A, Gilmore H, Basavanahally A, Feldman M, Ganesan S, Shih N. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. 2014. Presented at: Medical Imaging 2014: Digital Pathology; February 15-20, 2014; San Diego, California. [doi: [10.1117/12.2043872](https://doi.org/10.1117/12.2043872)]
8. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. Dec 2017;42:60-88. [doi: [10.1016/j.media.2017.07.005](https://doi.org/10.1016/j.media.2017.07.005)] [Medline: [28778026](https://pubmed.ncbi.nlm.nih.gov/28778026/)]
9. Artificial intelligence in health care: promises and challenges. IDB. URL: <https://publications.iadb.org/en/publications/english/viewer/Artificial-Intelligence-in-Health-Care-Promises-and-Challenges.pdf> [accessed 2025-07-29]
10. Cruz-Roa A, Basavanahally A, Gonzalez F, Gilmore H, Feldman M, Ganesan S. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. 2014. Presented at: Medical Imaging 2014: Digital Pathology; February 15-20, 2014; San Diego, California. [doi: [10.1117/12.2043872](https://doi.org/10.1117/12.2043872)]
11. Manjunathan N, Gomathi N, Muthulingam S. Early detection of breast cancer using machine learning. 2023. Presented at: 2023 International Conference on Sustainable Computing and Smart Systems; June 14-16, 2023; Coimbatore, India. [doi: [10.1109/icscss57650.2023.10169777](https://doi.org/10.1109/icscss57650.2023.10169777)]
12. Bansal M, Kumar M, Sachdeva M, Mittal A. Transfer learning for image classification using VGG19: Caltech-101 image data set. *J Ambient Intell Humaniz Comput*. 2023;14(4):3609-3620. [FREE Full text] [doi: [10.1007/s12652-021-03488-z](https://doi.org/10.1007/s12652-021-03488-z)] [Medline: [34548886](https://pubmed.ncbi.nlm.nih.gov/34548886/)]
13. Wakili MA, Shehu HA, Sharif MH, Sharif MHU, Umar A, Kusetogullari H, et al. Classification of breast cancer histopathological images using DenseNet and transfer learning. *Comput Intell Neurosci*. 2022;2022:8904768. [FREE Full text] [doi: [10.1155/2022/8904768](https://doi.org/10.1155/2022/8904768)] [Medline: [36262621](https://pubmed.ncbi.nlm.nih.gov/36262621/)]
14. Phalaagae P, Zungeru A, Yahya A, Sigweni B, Rajalakshmi S. A hybrid CNN-LSTM model with attention mechanism for improved intrusion detection in wireless IoT sensor networks. *IEEE Access*. 2025;13:57322-57341. [doi: [10.1109/ACCESS.2025.3555861](https://doi.org/10.1109/ACCESS.2025.3555861)]

15. Adam R, Dell'Aquila K, Hodges L, Maldjian T, Duong TQ. Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review. *Breast Cancer Res.* Jul 24, 2023;25(1):87. [FREE Full text] [doi: [10.1186/s13058-023-01687-4](https://doi.org/10.1186/s13058-023-01687-4)] [Medline: [37488621](https://pubmed.ncbi.nlm.nih.gov/37488621/)]
16. Mooney P. Breast histopathology images. Kaggle. 2019. URL: <https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images/data> [accessed 2024-02-10]
17. Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. *Insights Imaging.* Aug 2018;9(4):611-629. [FREE Full text] [doi: [10.1007/s13244-018-0639-9](https://doi.org/10.1007/s13244-018-0639-9)] [Medline: [29934920](https://pubmed.ncbi.nlm.nih.gov/29934920/)]
18. Li Z, Liu F, Yang W, Peng S, Zhou J. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE Trans Neural Netw Learning Syst.* Dec 2022;33(12):6999-7019. [doi: [10.1109/tnnls.2021.3084827](https://doi.org/10.1109/tnnls.2021.3084827)]
19. Chen C, Mat Isa NA, Liu X. A review of convolutional neural network based methods for medical image classification. *Comput Biol Med.* Feb 2025;185:109507. [doi: [10.1016/j.combiomed.2024.109507](https://doi.org/10.1016/j.combiomed.2024.109507)] [Medline: [39631108](https://pubmed.ncbi.nlm.nih.gov/39631108/)]
20. Mascarenhas S, Agarwal M. A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for image classification. In.; 2021. Presented at: 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications; November 19-21, 2021:96-99; Bangalore, India. [doi: [10.1109/centcon52345.2021.9687944](https://doi.org/10.1109/centcon52345.2021.9687944)]
21. Rajinikanth V, Joseph Raj AN, Thanaraj KP, Naik GR. A customized VGG19 network with concatenation of deep and handcrafted features for brain tumor detection. *Applied Sciences.* May 15, 2020;10(10):3429. [doi: [10.3390/app10103429](https://doi.org/10.3390/app10103429)]
22. Dey N, Zhang Y, Rajinikanth V, Pugalenth R, Raja NSM. Customized VGG19 architecture for pneumonia detection in chest X-rays. *Pattern Recognit Lett.* Mar 2021;143:67-74. [doi: [10.1016/j.patrec.2020.12.010](https://doi.org/10.1016/j.patrec.2020.12.010)]
23. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet large scale visual recognition challenge. ArXiv. Preprint posted online on September 1, 2014.
24. Han Z, Wei B, Zheng Y, Yin Y, Li K, Li S. Breast cancer multi-classification from histopathological images with structured deep learning model. *Sci Rep.* Jun 23, 2017;7(1):4172. [FREE Full text] [doi: [10.1038/s41598-017-04075-z](https://doi.org/10.1038/s41598-017-04075-z)] [Medline: [28646155](https://pubmed.ncbi.nlm.nih.gov/28646155/)]
25. FastBreastCancerAI. Hugging Face. URL: <https://huggingface.co/spaces/itahir/FastBreastCancerAI> [accessed 2025-02-05]
26. Chen RJ, Wang JJ, Williamson DFK, Chen TY, Lipkova J, Lu MY, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat Biomed Eng.* Jun 2023;7(6):719-742. [FREE Full text] [doi: [10.1038/s41551-023-01056-8](https://doi.org/10.1038/s41551-023-01056-8)] [Medline: [37380750](https://pubmed.ncbi.nlm.nih.gov/37380750/)]
27. Introducing AI fairness 360. IBM. URL: <https://research.ibm.com/blog/ai-fairness-360> [accessed 2025-07-29]
28. General data protection regulation (GDPR). GDPR.eu. URL: <https://gdpr.eu/tag/gdpr/> [accessed 2025-07-29]
29. 46.104 exempt research. Code of Federal Regulations. URL: <https://www.ecfr.gov/current/title-45/subtitle-A/subchapter-A/part-46/subpart-A/section-46.104> [accessed 2025-07-24]
30. Nebeker C, Torous J, Bartlett Ellis RJ. Building the case for actionable ethics in digital health research supported by artificial intelligence. *BMC Med.* Jul 17, 2019;17(1):137. [FREE Full text] [doi: [10.1186/s12916-019-1377-7](https://doi.org/10.1186/s12916-019-1377-7)] [Medline: [31311535](https://pubmed.ncbi.nlm.nih.gov/31311535/)]
31. Larson DB, Magnus DC, Lungren MP, Shah NH, Langlotz CP. Ethics of using and sharing clinical imaging data for artificial intelligence: a proposed framework. *Radiology.* Jun 2020;295(3):675-682. [doi: [10.1148/radiol.2020192536](https://doi.org/10.1148/radiol.2020192536)] [Medline: [32208097](https://pubmed.ncbi.nlm.nih.gov/32208097/)]
32. Thorogood A, Dalpé G, Knoppers BM. Return of individual genomic research results: are laws and policies keeping step? *Eur J Hum Genet.* Apr 2019;27(4):535-546. [FREE Full text] [doi: [10.1038/s41431-018-0311-3](https://doi.org/10.1038/s41431-018-0311-3)] [Medline: [30622328](https://pubmed.ncbi.nlm.nih.gov/30622328/)]
33. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *Lancet Digit Health.* Oct 2019;1(6):e271-e297. [FREE Full text] [doi: [10.1016/S2589-7500\(19\)30123-2](https://doi.org/10.1016/S2589-7500(19)30123-2)] [Medline: [33323251](https://pubmed.ncbi.nlm.nih.gov/33323251/)]
34. BI-RADS ACR breast imaging reporting and data system. American College of Radiology. URL: <https://www.acr.org/Clinical-Resources/Clinical-Tools-and-Reference/Reporting-and-Data-Systems/BI-RADS> [accessed 2025-07-24]
35. Li Q, Li W. Using deep learning for breast cancer diagnosis-technical report. Hong Kong. Chinese University of Hong Kong; 2017.
36. Bayramoglu N, Kannala J, Heikkila J. Deep learning for magnification independent breast cancer histopathology image classification. 2016. Presented at: 23rd International Conference on Pattern Recognition; December 4-8, 2016; Cancun, Mexico. [doi: [10.1109/icpr.2016.7900002](https://doi.org/10.1109/icpr.2016.7900002)]
37. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep learning for identifying metastatic breast cancer. ArXiv. Preprint posted online on June 18, 2016. [FREE Full text] [doi: [10.1007/s40615-015-0191-y](https://doi.org/10.1007/s40615-015-0191-y)] [Medline: [27294767](https://pubmed.ncbi.nlm.nih.gov/27294767/)]
38. Improving breast cancer screening with artificial intelligence. Google for Health. URL: <https://health.google/caregivers/mammography/> [accessed 2025-07-29]
39. A.I. software evaluation and development for breast cancer detection in tomosynthesis scans. Mayo Clinic. 2023. URL: <https://www.mayo.edu/research/clinical-trials/cls-20532539> [accessed 2025-07-29]
40. Wu N, Phang J, Park J, Shen Y, Huang Z, Zorin M, et al. Deep neural networks improve radiologists' performance in breast cancer screening. *IEEE Trans Med Imaging.* Apr 2020;39(4):1184-1194. [FREE Full text] [doi: [10.1109/TMI.2019.2945514](https://doi.org/10.1109/TMI.2019.2945514)] [Medline: [31603772](https://pubmed.ncbi.nlm.nih.gov/31603772/)]
41. Rajkomar A, Oren E, Chen K, Dai AM, Hajaj N, Hardt M, et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit Med.* 2018;1:18. [FREE Full text] [doi: [10.1038/s41746-018-0029-1](https://doi.org/10.1038/s41746-018-0029-1)] [Medline: [31304302](https://pubmed.ncbi.nlm.nih.gov/31304302/)]

42. Reza M, Ma J. Imbalanced histopathological breast cancer image classification with convolutional neural network. 2018. Presented at: 14th IEEE International Conference on Signal Processing; August 12-16, 2018:12-16; Beijing, China. [doi: [10.1109/icsp.2018.8652304](https://doi.org/10.1109/icsp.2018.8652304)]
43. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D. Self-supervised learning for medical image analysis using image context restoration. *Med Image Anal.* Dec 2019;58:101539. [FREE Full text] [doi: [10.1016/j.media.2019.101539](https://doi.org/10.1016/j.media.2019.101539)] [Medline: [31374449](https://pubmed.ncbi.nlm.nih.gov/31374449/)]

## Abbreviations

**AI:** artificial intelligence  
**CNN:** convolutional neural network  
**IDC:** invasive ductal carcinoma  
**ReLU:** rectified linear unit  
**ROC:** receiver operating characteristic  
**VGG19:** Visual Geometry Group 19

*Edited by A Mavragani; submitted 06.06.24; peer-reviewed by T Issa, HS Das, J Li, S Tayebi Arasteh, M Aria; comments to author 24.09.24; revised version received 06.12.24; accepted 10.07.25; published 21.08.25*

### *Please cite as:*

Chilumukuru NS, Priyadarshini P, Ezunkpe Y

*Deep Learning for the Early Detection of Invasive Ductal Carcinoma in Histopathological Images: Convolutional Neural Network Approach With Transfer Learning*

*JMIR Form Res* 2025;9:e62996

URL: <https://formative.jmir.org/2025/1/e62996>

doi: [10.2196/62996](https://doi.org/10.2196/62996)

PMID:

©Naga Shreya Chilumukuru, Pragya Priyadarshini, Yawo Ezunkpe. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 21.08.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.