

# Integrating hybrid transfer learning with attention-enhanced deep learning models to improve breast cancer diagnosis

Sudha Prathyusha Jakkaladiki and Filip Maly

Faculty of Informatics and Management, University of Hradec Králové, Hradec Kralove, Hradec Kralove, Czech Republic

## ABSTRACT

Cancer, with its high fatality rate, instills fear in countless individuals worldwide. However, effective diagnosis and treatment can often lead to a successful cure. Computer-assisted diagnostics, especially in the context of deep learning, have become prominent methods for primary screening of various diseases, including cancer. Deep learning, an artificial intelligence technique that enables computers to reason like humans, has recently gained significant attention. This study focuses on training a deep neural network to predict breast cancer. With the advancements in medical imaging technologies such as X-ray, magnetic resonance imaging (MRI), and computed tomography (CT) scans, deep learning has become essential in analyzing and managing extensive image datasets. The objective of this research is to propose a deep-learning model for the identification and categorization of breast tumors. The system's performance was evaluated using the breast cancer identification (BreakHis) classification datasets from the Kaggle repository and the Wisconsin Breast Cancer Dataset (WBC) from the UCI repository. The study's findings demonstrated an impressive accuracy rate of 100%, surpassing other state-of-the-art approaches. The suggested model was thoroughly evaluated using F1-score, recall, precision, and accuracy metrics on the WBC dataset. Training, validation, and testing were conducted using pre-processed datasets, leading to remarkable results of 99.8% recall rate, 99.06% F1-score, and 100% accuracy rate on the BreakHis dataset. Similarly, on the WBC dataset, the model achieved a 99% accuracy rate, a 98.7% recall rate, and a 99.03% F1-score. These outcomes highlight the potential of deep learning models in accurately diagnosing breast cancer. Based on our research, it is evident that the proposed system outperforms existing approaches in this field.

Submitted 2 August 2023  
Accepted 10 January 2024  
Published 28 February 2024

Corresponding author  
Sudha Prathyusha Jakkaladiki,  
sudha.jakkaladiki@uhk.cz

Academic editor  
Jyotismita Chaki

Additional Information and  
Declarations can be found on  
page 19

DOI 10.7717/peerj-cs.1850

© Copyright  
2024 Jakkaladiki and Maly

Distributed under  
Creative Commons CC-BY 4.0

## OPEN ACCESS

**Subjects** Bioinformatics, Artificial Intelligence, Computer Vision, Data Mining and Machine Learning, Data Science

**Keywords** Attention mechanism, Breast cancer, Deep learning, Feature fusion, Transfer learning

## INTRODUCTION

Breast cancer is the most lethal disease in women and remains a severe public health concern in many countries (*The Early Breast Cancer Trialists' Collaborative Group (EBCTCG), 1990*). The American Cancer Society's projections for 2017 anticipated that more than 250,000 women would receive a diagnosis of invasive breast cancer, with over 40,000 succumbing to the disease (*Smith et al., 2017*). Given its complexity and diverse clinical

outcomes ([Rakha et al., 2010](#); [Rivenbark, O'Connor & Coleman, 2013](#)), different cases of this intricate illness exhibit molecular, behavioral, and physical variations, influencing therapeutic responses.

Invasive breast cancer has become progressively challenging to diagnose and treat, primarily due to its intricate nature and varying clinical manifestations ([Martin et al., 2005](#)). Thus, having the ability to accurately predict cancer prognosis can significantly benefit both patients and doctors in determining the most suitable treatment approach for individuals with breast cancer. Prognostication is particularly important in clinical practice, especially when dealing with patients with a low chance of survival. Clinicians regularly rely on prognostic prediction information to facilitate clinical decision-making ([Sun et al., 2007](#); [Gevaert et al., 2006](#)), determine a patient's eligibility for specific treatment programs ([Xu et al., 2012](#)), and plan, conduct, and evaluate clinical trials with the availability of precise prognosis estimates. When a patient is expected to have limited survival time, doctors can offer them the opportunity to consider their treatment options and make realistic preparations for their end-of-life care ([Stone & Lund, 2007](#)).

People visit an oncologist if they experience any symptoms that might be related to cancer. The oncologist has various tools, such as mammograms, breast magnetic resonance imaging (MRI) scans, breast ultrasounds, breast X-rays, and tissue biopsies, to diagnose and detect breast cancer early. Regular sentinel node biopsies are performed for patients with confirmed breast cancer to check for malignant cells in the lymph nodes. Machine learning algorithms can be employed to classify tumors as either benign or malignant.

An early breast cancer diagnosis can improve survival rates and better prognoses, facilitating timely therapy for patients ([Sun et al., 2017](#)). Moreover, patients with benign tumors can avoid unnecessary procedures. In the realm of healthcare, data mining techniques have the potential to enhance patient treatment quality, reduce healthcare costs, and even save lives. Deep learning approaches are particularly effective in handling complex and extensive datasets, outperforming traditional machine learning algorithms in accuracy. The research on breast cancer currently concentrates on several datasets related to the subject, furthering our understanding and potential for better diagnosis and treatment.

The rapid growth of breast cancer cases underscores the significance of investing in research and developing innovative approaches for its treatment. This drive urges scientists to discover new methods for swift and accurate diagnoses, ultimately leading to improved patient care and extended lives ([Hamidinekoo et al., 2018](#); [Kousalya & Saranya, 2021](#)). The key to early and consistent detection lies in revising preliminary diagnostic information and gathering relevant data from previous findings, which can be enhanced through medical imaging and deep learning (DL) strategies. Medical imaging plays a critical role in clinically diagnosing diseases, evaluating therapies, and detecting abnormalities in various parts of the body, including the eye ([Akbar et al., 2018](#)), brain ([Rajinikanth et al., 2017](#)), breast ([Fonseca et al., 2015](#)), and stomach ([Houssein et al., 2021](#)). The primary objective of medical imaging research is to classify affected organs' location, size, and characteristics, offering a promising means of extracting valuable information from extensive datasets. Medical imaging techniques such as mammograms, histopathological

images, MRI scans, ultrasounds, and thermograms are commonly employed in diagnosing breast cancer ([Houssein et al., 2021](#)).

Over the past decade, thermal imaging (or thermography) has shown tremendous promise in early breast cancer diagnosis. Thermography photos reveal physiological changes that can assist in other forms of diagnostics ([Hossam, Harb & Abd El Kader, 2018](#)). Another tool used by young women for identifying breast cancer is ultrasound imaging. However, this method may have difficulty detecting microcalcifications and deeper breast tissue due to noise levels ([Qi et al., 2019](#)), MRI screening, along with ultrasonography and thermography, is yet another method employed for early cancer detection. Magnetic resonance imaging (MRI) enables the creation of exact three-dimensional (3D) transverse images, surpassing X-rays in accuracy ([Houssein et al., 2021](#)). Mammography stands as a crucial diagnostic imaging technique and serves as the gold standard for breast cancer screening, having demonstrated a significant reduction in breast cancer mortality. As an X-ray examination, mammography is widely regarded as the most reliable and precise tool for detecting breast cancer ([Dhawan, 2011](#)). In this study, breast cancer was categorized based on mammograms. Our methodology combines DenseNet and a convolutional neural network (CNN), leveraging the strengths of both architectures. This fusion allows us to achieve better classification accuracy and diagnostic performance compared to existing methods on the same dataset. One key innovation in our methodology is the fusion of feature maps from DenseNet and CNN at a standard layer. This approach captures a richer set of features, enhancing the model's ability to differentiate between benign and malignant breast tumors. This feature fusion technique is unique to our work and directly impacts the model's performance.

The notable contributions of the studies, as mentioned above,

1. Examine the Fusion feature fusion methods for better feature extraction.
2. With an accuracy of 99%, the suggested hybrid transfer learning model (a combination of CNN and DenseNet) outperforms state-of-the-art deep learning methods
3. In this work, we employed a pre-trained neural network as a starting point for our model development, leveraging the features and representations learned from diverse datasets during its initial training. Subsequently, we fine-tuned and adapted the model to our specific task of tumor saliency detection using our dataset. This process allowed us to benefit from the knowledge encapsulated in the pre-trained model while tailoring it to our unique application.
4. The optimization method raises the breast cancer detection rate and boosts the efficiency of the suggested system.

## RELATED WORKS

Numerous researchers have explored artificial intelligence, expert systems, and neural networks to enhance the accuracy of breast cancer (BC) screening. Traditionally, hospitals relied on X-rays for breast cancer diagnosis. Still, recent advancements in intelligent modeling have significantly improved the efficiency and accuracy of mammography imaging, leading to the widespread replacement of X-rays. The datasets utilized in [Kiyan &](#)

*Yildirim (2004)* and *Gonzalez-Angulo, Morales-Vasquez & Hortobagyi (2007)* were obtained from the Kaggle repository. These studies presented a model aimed at determining the likelihood of someone having breast cancer and aiding in early detection and diagnosis. The researchers in these papers compared the effectiveness of various models for predicting breast cancer, including Naive Bayes classifiers and logistic regression. Through machine learning methodologies, their research demonstrated an accuracy ranging from 52.63% to 98.24% in predicting breast cancer illness. Over the past few decades, scientists have dedicated their efforts to studying thermographic breast cancer diagnosis using machine learning techniques. Some researchers have focused on identifying tumor size and location, while others have honed in on different features, such as attainment protocols and breast quadrants. Deep learning, a type of machine learning employing CNNs with multiple hidden layers (*Wehle, 2017*), has shown promising results. With a training dataset, deep learning can automatically extract relevant features. Researchers have made significant progress utilizing CNNs to detect breast cancer in recent years. Interestingly, CNNs were not frequently employed for breast cancer detection using thermal imaging, possibly due to the computational load or the efficiency of CNNs when compared to texture or statistical features (*Zuluaga-Gomez et al., 2021*). Nonetheless, convolutional neural networks have emerged as one of the most promising tools for pattern identification in recent years.

*Tiwari et al. (2020)* and *Naji et al. (2021)* employed a deep learning methodology, a form of deep learning known as convolutional neural networks (CNNs), to classify breast cancer mammography pictures from the public dataset BreakHis. The proposed technique utilized a feedforward network and preliminary trials, where picture patches were extracted for CNN training and then combined for final classification. The accuracy rates for classifying screening mammograms using CNNs were notably high, with the highest accuracy achieved on the Digital Database for Screening Mammography's digitized film mammograms reaching 88% in an independent test set.

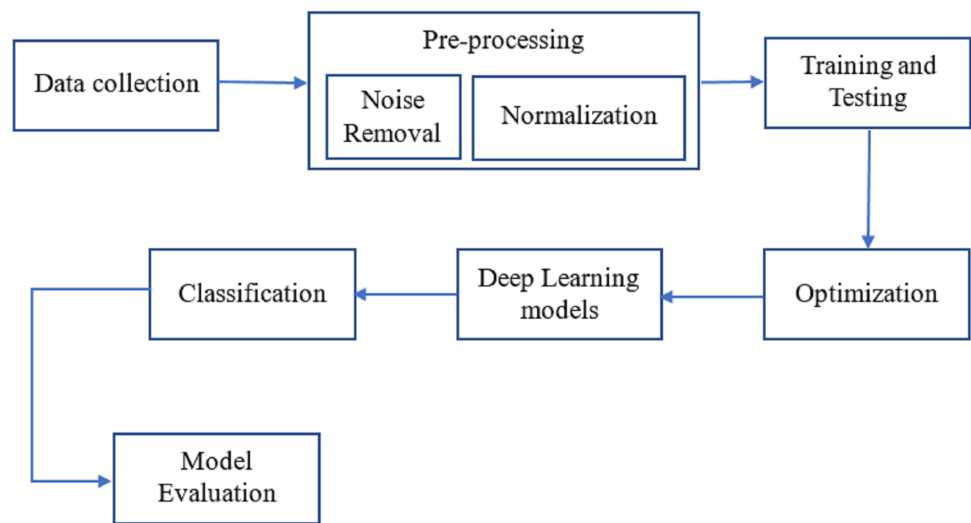
In research by *Agnes et al. (2020)* and *Joo et al. (2004)*, the primary objective was to evaluate how well artificial neural networks (ANNs) could categorize tumors into distinct prognostic groups based on gene expression profiles. Small round blue cell tumors (SRBCTs) were used to train the ANN, as clinical practitioners frequently encounter diagnostic challenges related to these four categories of cancer. The ANN correctly classified all samples, and the study identified crucial genes associated with each classification. The experimental findings suggested that the novel approaches introduced in the research might significantly enhance sample classification accuracy and consistency of selection outcomes. As a result, the new algorithms improved the categorization accuracy to almost 99%.

*Chiang et al. (2018)* suggested a computer-assisted screening method for identifying tumors in breast ultrasonography. The technique utilizes a 3D convolutional neural network and a weighted sum of the best candidates. Initially, a sliding window technique is applied to isolate the relevant volumes, and then a 3D convolutional neural network estimates the likelihood of malignancy in each volume of interest. Those with a high likelihood estimate are flagged as potential tumor patients, and their situations may intersect. An original method was developed to accumulate candidates and handle overlaps, ranking

potential candidates based on the likelihood of tumor formation during the aggregation process. Experimental results with 171 tumors showed that the suggested model achieved 95% sensitivity in less than 22 s, demonstrating its superior efficiency compared to current methods. [Agnes et al. \(2020\)](#) proposed the multiscale all CNN (MA-CNN) model for breast cancer detection. In this model, mammograms were categorized using a CNN-based method, and multiple scale characteristics extracted from mammography pictures enhanced the CNN classifier's performance. The suggested MA-CNN model proved to be a valuable tool for the early identification of breast cancer from mammograms, as demonstrated by experimental findings, which involved classifying mammograms into benign and malignant categories. Additionally, [Houssein et al. \(2021\)](#) developed a deep learning (DL) technique for modeling breast cancer detection, although further details about the DL technique were not provided in the provided text.

Chronic care involves managing and treating long-term medical conditions or diseases ([Shen et al., 2023](#)). Metal-organic frameworks are porous materials with a highly ordered structure, typically composed of metal ions or clusters connected by organic ligands ([Zeng et al., 2020](#); [Liu et al., 2021](#)). It includes patient vital signs, lab results, ECG readings, or other medical data that can vary over time ([Sun et al., 2023](#)). This method is likely designed to assist in medical image analysis, diagnosis, or research, where identifying images with similar features can be valuable for tasks like disease detection or treatment planning, particularly in lung-related medical conditions ([Zhuang et al., 2022](#); [Zhuang, Jiang & Xu, 2022](#)). This approach likely involves matching key features or points in the images to align them properly in the mosaic. It provides a more comprehensive view of the area being examined during the endoscopy procedure ([Zhang et al., 2022b](#); [Lu et al., 2023b](#)). This suggests using a deep learning approach involving a neural network for matching and tracking soft tissue features ([Lu et al., 2023a](#)). This can be a crucial tool in cancer research and medical diagnostics, enabling automated analysis of blood samples to identify and count circulating tumor cells, which can provide valuable information about a patient's cancer status and progression ([He et al., 2020b](#); [Xie et al., 2021](#); [He et al., 2020a](#)). This part implies that the overarching goal of the process is to facilitate the repair of the endothelium in the context of hypertension, a condition characterized by high blood pressure and associated vascular damage ([Li et al., 2021](#); [Zhang et al., 2022a](#)). These networks can learn from and predict large, complex datasets. Deep learning has widespread application across various fields, including image recognition, natural language processing, and medical research ([Zhu et al., 2021](#); [Gao et al., 2022](#)). This refers to a long non-coding RNA (lncRNA) called PVT1. lncRNAs are RNA molecules that do not encode proteins but play vital roles in cellular regulation ([Chang et al., 2019](#); [Wen et al., 2015](#)). The primary objective of this research or framework is to determine or predict the specific tissue or organ within the body from which a cancer sample has originated ([He et al., 2020c](#); [Wang et al., 2023](#); [Jiang et al., 2022](#)). Haplotypes are genetic variations or alleles that are commonly inherited together. In this case, two specific haplotypes, GCA and ACA, within the ESR1 gene are under examination ([Liu et al., 2022](#); [Tang et al., 2022](#)).

The suggested model for mammography screening employs end-to-end training methods, requiring labels at the picture level throughout the system's resting periods



**Figure 1** The workflow of the proposed methodology for the process of finding breast cancer.

Full-size [DOI: 10.7717/peerjcs.1850/fig-1](https://doi.org/10.7717/peerjcs.1850/fig-1)

after the first training phase. This model was trained and tested on a dataset for categorizing mammography images using CNN. The findings indicated that the proposed model outperformed previous efforts in heterogeneous mammography, particularly regarding accuracy. In another study by [Zhou et al. \(2018\)](#), a radionics approach based on convolutional neural networks (CNNs) was introduced for identifying breast cancer. The proposed technique utilized shear-wave elastography data to train a CNN and extract morphology-related features. The model demonstrated high accuracy with training on 540 images, 318 of which were cancerous and 222 benign. Experimental data showed a sensitivity of 96.2%, a specificity of 95.8%, and an overall accuracy of 95.8%. Furthermore, [Qi et al. \(2019\)](#) conducted a study highlighting how multiparametric magnetic resonance imaging improved radiologists' ability to diagnose breast cancer. This research used a pre-trained CNN model to extract structures from 927 photos. CNN characteristics were then employed to train an SVM classifier, successfully identifying healthy and cancerous tissue. The study explored the impact of adding more fusion to the mix and discovered that combining features, images, and classifiers played a crucial role in improving accuracy.

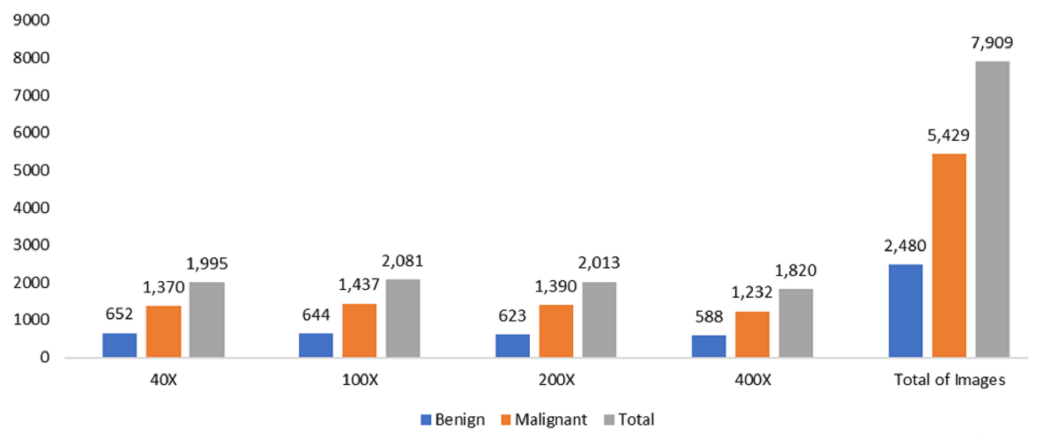
## METHODS AND MATERIALS

As can be seen in [Fig. 1](#), breast cancer detection is done. The DenseNet and feature fusion are two important innovations we tested in this method. The effectiveness of neural networks motivates using Dense Net for perfect classification.

### Dataset

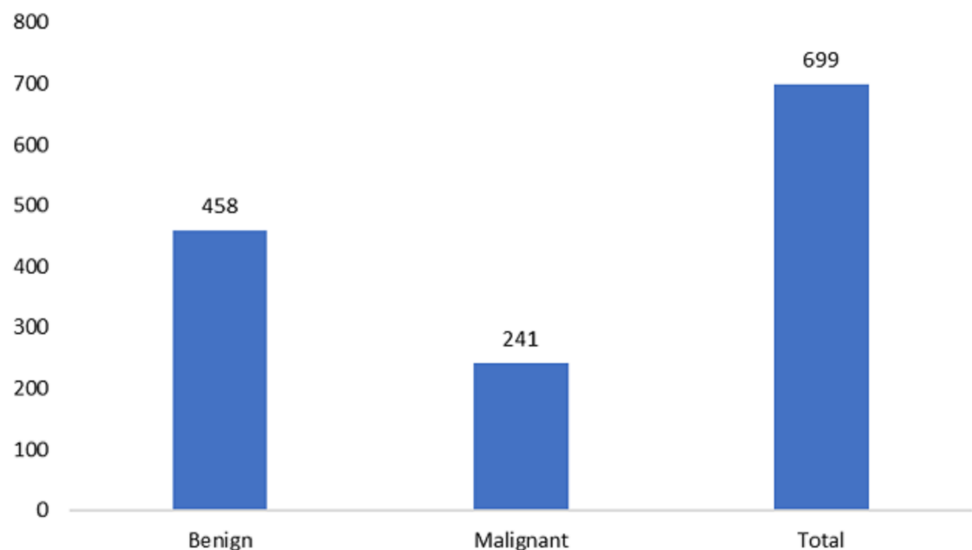
The data was obtained from the Kaggle data repository, comprising 7,909 microscopic photos of breast cancer tissue from 82 individuals in the BreakHis dataset. These images were captured at magnifications of 40X, 100X, 200X, and 400X. Among the samples, there were 2,480 healthy samples and 5,429 cancerous ones. The dataset development involved





**Figure 2** Data distribution from BreakHis dataset.

[Full-size](#) DOI: 10.7717/peerjcs.1850/fig-2



**Figure 3** Data distribution from the Wisconsin-Breast Cancer (Diagnostics) (WBC) dataset.

[Full-size](#) DOI: 10.7717/peerjcs.1850/fig-3

collaboration with the P&D Laboratory of Pathological Anatomy and Cytopathology in Parana, Brazil.

The measurements for breast cancer patients are recorded in the original Wisconsin Breast Cancer (WBC) dataset, available directly from the UCIML library. The dataset consists of both noncancerous and cancerous varieties. Specifically, this dataset's malignant category, containing outliers, comprises 21 points. In contrast, the inliers belong to the safe group. The Wisconsin cancer dataset includes 699 cases, with 458 noncancerous (accounting for 65.5% of the cases). [Figures 2](#) and [3](#) illustrate the data distribution from the datasets.

## Pre-processing

As discussed previously, noise reduction from input pictures is a crucial issue in medical imaging. Maintaining picture edges' maximum sharpness and integrity is vital during the noise-reduction process. Each pixel in the output is determined by calculating the median of the brightness values of the surrounding pixels in the input (Zhang & Hong, 2019). Median filtering establishes a pixel's significance by averaging its neighboring pixels' importance. The central filter is not highly influenced by outliers, enabling it to eliminate these values without compromising the overall image quality. This filter preserves edge form and position while reducing light intensity variance (Sharifrazi et al., 2021). The  $m \times n$  neighborhood filter utilizes an ascending sort on the sorted data before replacing the center pixels with a new set of values. Additionally, the median filter effectively filters out salt and pepper noises (Song, Jia & Ma, 2019). For this reason, we applied this filter as the first step in the processing pipeline to the input photos. The median value of the surrounding pixels is used to replace the original one in median filtering.

$$a_{(x,y)} = \text{median}(b_{m,n} : (m,n)) \in \tau \quad (1)$$

where  $(x,y)$  is the set of closest neighbors defined by  $\tau$ . The filter size in this example is  $5 \times 5$  pixels.

After eliminating the background noise, the images must be standardized by being scaled between 0 and 1 to make the dataset more stable. Here, we employed the min-max technique of normalization. Given the following restriction,

$$M_n = [A \subseteq \mathbb{R}^n] \rightarrow [x, \dots, y] \quad (2)$$

on the grayscale picture and its dimensions, the normalized image,  $M^*$ , can be described as follows.,

$$M^* = x_{new} + \frac{y_{new} - x_{new}}{y - x} \times (M - x) \quad (3)$$

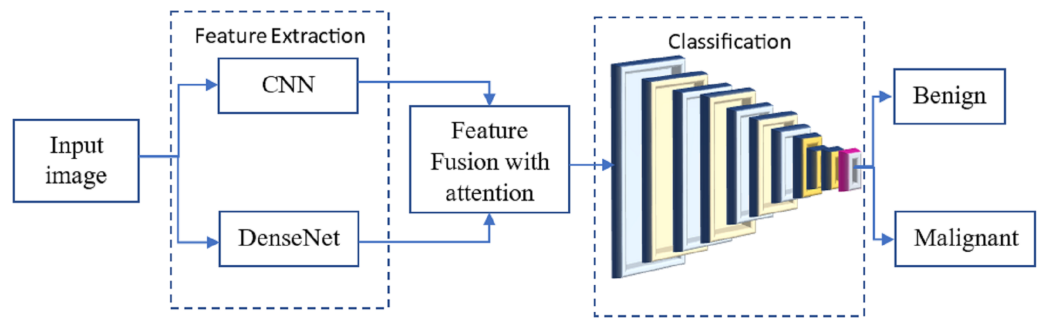
in where  $x$  and  $y$  are image intensities.

$$M^* = [A \subseteq \mathbb{R}^n] \rightarrow [x_{new}, \dots, y_{new}] \quad (4)$$

in where  $x_{new}$  and  $y_{new}$  are normalized image intensities.

Various pre-processing techniques were applied before inputting the pictures into the customized variational deep-learning algorithm. All micro pictures in BreakHis are saved as PNGs with an 8-bit resolution per channel and three RGB channels. Using machine learning methods, the size of numerous high-resolution photos was reduced to 224 by 224. The proposed technique converts the collected images into Numpy arrays, facilitating faster model training with minimal space requirements. Several data augmentation approaches were utilized to reinforce the model to address overfitting issues. The dataset size increased from 7,909 to 54,403 through the data preparation method. In the pre-processing phase, a zoom factor of 2, a rotation angle of 90 degrees, a shearing factor of 0.5, and a shifting factor of 0.4 were applied to both the width and height of the image.





**Figure 4** The proposed methodology of the breast cancer prediction with feature fusion on CNN and DenseNet.

Full-size DOI: [10.7717/peerjcs.1850/fig-4](https://doi.org/10.7717/peerjcs.1850/fig-4)

## Proposed methodology

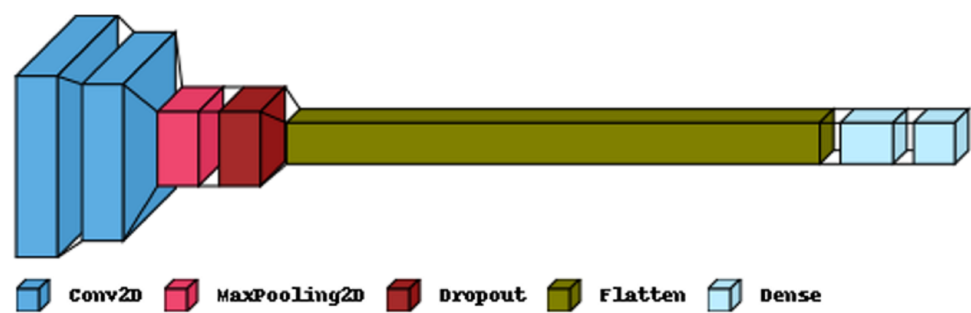
Many researchers have investigated the use of CNN for breast cancer diagnosis, an essential topic in medical picture processing. One kind of CNN that has performed well in picture classification is called DenseNet. The term "feature fusion" describes merging features from several models to boost performance. We may experiment with a DenseNet and a CNN to identify breast cancer in this scenario. One approach to feature fusion involves merging the feature maps of these two models at a standard layer, and the combined output is then passed through a fully connected layer to produce the final classification result. Alternatively, selective feature fusion can be achieved using attention techniques. The proposed model, as depicted in Fig. 4, incorporates these feature fusion strategies to enhance breast cancer diagnosis performance.

The resulting model can be trained on numerous annotated mammograms through supervised learning. Both DenseNet and CNN weights are adjusted concurrently to achieve optimal performance during training. Once the model is introduced, it can be utilized to determine whether a new mammogram is cancerous or benign.

## Convolutional neural network

CNNs uncover hidden patterns in images by convolving over the picture. In the early stages of CNNs, the network identifies simple patterns like straight lines and sharp angles. However, as we delve deeper using our neural net, we can capture more complex features. This characteristic makes CNNs highly effective in detecting objects in images. The suggested technique employs convolutional neural networks to identify breast cancer based on photographic evidence.

Figure 5 illustrates the primary layers constituting a CNN's architecture: the convolutional, pooling, and fully connected layers. In the first layer, neurons connected to nearby areas compute their output. Each output is determined by the dot product of the weights and the corresponding area. Standard filter sizes for input images are typically  $3 \times 3$ ,  $5 \times 5$ , or  $8 \times 8$  squares. By moving a window over the image, these filters learn patterns that appear throughout the picture. The stride represents the distance between successive filters. If the stride hyperparameter exceeds the filter dimension, the convolution expands



**Figure 5** The proposed CNN model structure for the feature extraction.
Full-size DOI: 10.7717/peerjcs.1850/fig-5

**Table 1** The parameter setting for the convolutional neural network model.

Layer (type)	Output shape	Param
Conv2D	(22, 22, 32)	1,568
Conv2D	(19, 19, 32)	16,416
MaxPooling2	(9, 9, 32)	0
Dropout	(9, 9, 32)	0
Flatten	(2,592)	0
Dense	(256)	663,808
Dense	(2)	514
Total params: 682,306		
Trainable params: 682,306		
Non-trainable params: 0		

to overlap adjacent windows. Table 1 comprehensively describes neural networks (NNs) in terms of their parameter numbers.

The CNN model is commonly employed for data training in medical image analysis, examination, and various applications. CNN is crucial in the successful computer aided diagnosis (CAD) systems used for medical imaging. This makes CNN a suitable choice for our intended application of breast cancer detection, although there is still potential for improvement. In a typical CNN, the kernels consist of hundreds of neurons arranged in multiple layers. Small kernels are used to maintain a constant depth for the input picture. Neurons are connected to a receptive field, representing a relatively narrow region. Especially when dealing with high-dimensional input images, connecting all neurons directly to the previous outputs becomes exceedingly challenging. As a result, alternative approaches are utilized to address this issue.

The design of our model is built upon the Sequential model, allowing the stacking of input and output layers sequentially. The convolution layer employs convolutional filters to perform a full-image scan of the input data. Its hyper-parameters include the filter size and stride (the distance between two successive receptive filters). The result of this operation is referred to as an activation map or feature map. We first create a 2D convolutional layer to process the input breast pictures. The number of output channels (16 channels in our

case) is the first input to the convolution layer function. We utilize a spatial convolution algorithm with a 3x3 filter kernel and a stride of 1, allowing the kernel to slide horizontally and vertically across the entire image. Throughout our trials, we experimented with various kernel sizes, ranging from 1 to 7, to assess their impact on the results. We observed that the 1x1 kernel size limited the system's performance. We have considered padding to ensure that the filter and stride defined by this model completely cover the input picture.

The input, denoted by  $x$ , has been used to activate a system according to the rectified linear activation function (RELU), which is written as,

$$f(x) = (0, x) \quad (5)$$

The spatial size of the feature map is shrunk *via* max pooling. The Maxpooling function downsamples the input depiction by selecting the most significant value along each dimension of the feature's axis within the window given by `pool_size`. Two further convolutional layers are added, each with 32 or 64 output channels, and the process is repeated. Initially, we employed a  $2 \times 2$  maximum pooling filter. Our approach to fixing the degradation issue is based on the same idea as the deep residual learning architecture utilized. It operates on recursive units of  $1 \times 1$ ,  $3 \times 3$ , and  $1 \times 1$  convolution filters. The activation of each feature map was calculated using a global average pooling method.

### DenseNet model

To test and analyze our dataset, we employed the DenseNet CNN framework. DenseNet offers several advantages over previous pretraining CNN techniques, including improved handling of the vanishing-gradient problem, fewer parameters, increased feature reuse, and enhanced feature propagation. CNN is a series of feedforward layers that utilize convolutional filters and pooling layers. The CNN employs multiple fully connected layers after the final pooling layer to convert the 2D feature maps learned in earlier layers into a 1D vector for classification. The notation for this is,

$$F(y) = f_m(f_m - 1(\dots(f_1(y)))) \quad (6)$$

Our method employs a total of  $m$  hidden layers.  $y$  is the input data and  $f_m$  is the function of layer  $m$ . The convolutional layer in a standard CNN model is a function  $f$  that takes in a series of convolutional kernels  $(l_1, l_2, \dots, l_k)$ , each performing a different task. For each  $k$ ,  $l_k$  represents a linear function in the  $k$ th kernel:

$$l_k(M, N) = \sum_{q=-x}^x \sum_{r=-y}^y \sum_{s=-h}^p W_k(q, r, s) Y(y - q, z - r, x - s) \quad (7)$$

The input  $y$ 's input pixel is defined by  $(q, r, s)$ . The kernel weight is denoted by  $W_k$ .

### Attention mechanism

In the process of making predictions, a deep learning model can pinpoint relevant aspects of the input through the use of attention mechanisms. The attention mechanism functions by assigning varying weights to different input features based on their predictive importance. Subsequently, the prediction is made using the weighted sum of these input characteristics.

This approach enhances the model's performance by allowing it to focus on the most crucial elements while disregarding less significant ones.

The attention mechanism is a vital concept in deep learning, especially when dealing with long input sequences and the risk of information loss due to a fixed intermediate vector length. Since its development for the seq2seq paradigm in Natural Language Processing (NLP), the attention mechanism has gained rapid adoption in various domains. The output of the attentional mechanism can be expressed as a way to effectively address this issue as,

$$Attention(X, Y, Z) = softmax\left(\frac{XY^T}{\sqrt{k_m}}\right) \quad (8)$$

where,  $K = IW_k$ ,  $K \in (X, Y, Z)$  and  $I$  is the input,  $W_k$  is the learnable weight matrix. The dimension of the keys is denoted by  $k_m$ . The softmax is the activation function, which is calculated as,

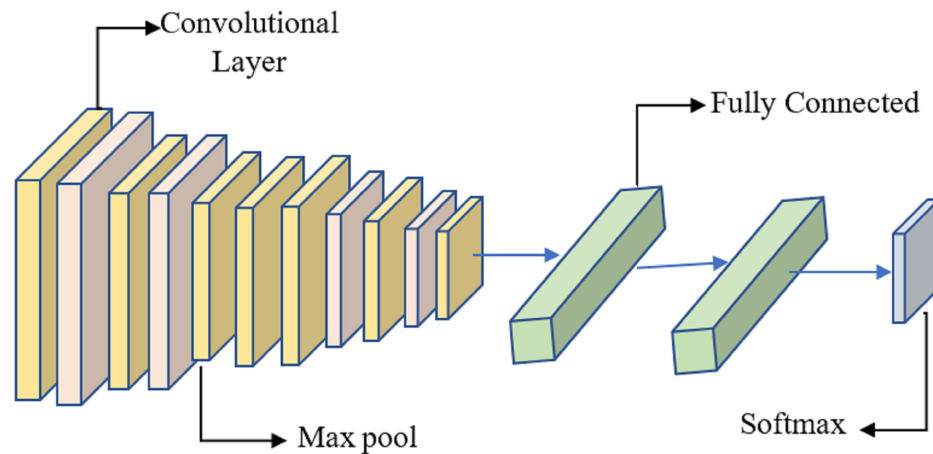
$$softmax(y_i) = \frac{e^{y_i}}{\sum_i e^{y_i}} \quad (9)$$

## Classification

A typical CNN model comprises five layers: the convolution layer, the rectified linear activation function (RELU) layer, the max pooling layer, the fully connected layer, and the dropout layer. The convolution layer holds utmost significance in a CNN, utilizing trainable filters whose settings are adjusted in each cycle. The RELU layer is a popular choice in CNN designs as it reduces training time. The Max pooling layer is often employed to manage overfitting and decrease parameter size. In the fully connected layer, neurons form a typical neural network structure. Lastly, the dropout layer is utilized to avoid overfitting.

It is proposed to employ a deep learning model based on a CNN to distinguish between healthy and diseased breast tissue. Figure 6 depicts the network's nine layers, of which the first six are convolutional, and the last three are fully linked. The proposed model has a first layer that filters a  $228 \times 228$  input picture using  $64 \times 7 \times 7$  kernels spaced by a stride of 6 pixels. The depth = 3 of the first layer's kernels determines the number of color channels in the source thermogram. The output of the first layer is sent into the second layer, which filters it using 128 kernels of size  $3 \times 3 \times 64$ , and then max-pooling is used to increase resilience and decrease computation. The third, fourth, and fifth levels are joined without pooling layers. The size of each kernel in the third layer is  $3 \times 3 \times 128$ , and there are 256 of them. There are 256,  $3 \times 3 \times 256$  kernels in the fourth layer and the same number in the fifth layer. The sixth layer contains 256,  $3 \times 3 \times 256$  kernels and is coupled to the fifth layer through a max-pooling layer. Two layers, each with 1024 neurons, are fully linked on top of the convolutional layers. Class size is proportional to the number of neurons in a neural network's last, fully-connected layer.

The "Optimizer RMSprop" was employed to evaluate various activation functions in TensorFlow, the 1D tensors generated by the convolutional layers, and the output "Flatten" layer, all contributing to weight optimization. The remaining dataset images not used in training were utilized during the test phase. The desired image feature vector is obtained



**Figure 6** System for benign and malignant data classification using fused features from a transfer learning framework.

Full-size DOI: [10.7717/peerjcs.1850/fig-6](https://doi.org/10.7717/peerjcs.1850/fig-6)

from the layer output in this phase. Subsequently, the image feature vector is compared to the feature matrix, and certain deep neural network layers capture specific data sections. However, a set of possibilities is needed for the final judgment of data categorization. To achieve this, the softmax function is often used to normalize the probability values to the interval  $[0, 1]$ . Convolutional layer output is determined by,

$$Z_c = \frac{I_s - K_s + 2d}{Stride} + 1 \quad (10)$$

The output of the pooling layer is calculated as,

$$Z_p = \frac{I_s - P_s}{Stride} + 1 \quad (11)$$

where the  $I_s$  signifies the size of the input and  $K_s$  symbolizes the kernel size, the padding is designated by  $d$  and the size of the pooling layer is signified by  $P_s$ . As noted above, the “Optimizer RMSprop” technique was used to determine the weights optimally. Reducing the cross-entropy is the basis of the system’s optimization functioning, which has the following mathematical form.

$$L = \sum_{l=1}^K \sum_{p=1}^J -v_p^l \log y_p^l \quad (12)$$

The sample number is signified by  $K$ . The output vector is denoted by  $v_p = (0, \dots, 0, 1, \dots, 1, 0, \dots, 0)$ . The predicted output is characterized by  $y_p$  and the class is formulated by,

$$k_p^l = \frac{e^{g_p}}{\sum_{l=1}^J e^{g_p}} \quad (13)$$

With the weight penalty,

$$L = \sum_{l=1}^K \sum_{p=1}^J -v_p^l \log y_p^l + \frac{1}{2} \mu \sum_C \sum_L U_{c,l}^2 \quad (14)$$

The total number of layers  $L$ , and the connections of layer  $L$  are denoted by  $C$ . The weight is described  $U_c$ .

## Hyper parameter tuning

This section discusses the hyperparameters associated with the fine-tuned DenseNet model, such as learning rate and batch processing loss. Choosing appropriate parameters and avoiding overfitting and underfitting the model is crucial for achieving the best training and testing results. This paper examines the amount of data lost during training and the accuracy of the test results. The model's divergence begins when it has already surpassed the optimal learning rate range. Ideally, the loss should continue to decrease when the learning rate is selected. Regarding the optimizer's L2 penalty (weight decay), the author suggests using the highest possible learning rate, allowing for faster training than grid-search when using weight decays of 0.01 (the minimum), 0.0001 (the next lowest), and 0.000001 (the highest). To address the overfitting issue, two dropout layers are utilized. During model training, a significant amount of data is lost after the first dropout layer, and progressively less data is lost with each succeeding dropout layer.

## EXPERIMENTAL SETUP

The tests in this study were conducted using Python 3 and a GPU. Keras and the scikit-learn package were utilized to successfully create the optimal deep-learning models. The dataset was divided into two sections: a training dataset consisting of 80% of the total data used for model enhancement and registering cross-validation (CV) results, and a testing dataset consisting of 20% of the total data used for model assessment and registering testing results. We employed two feature-selection techniques based on correlation to narrow down the potential features to eight. Then, the features selected *via* correlation, the features selected *via* univariate, and the features chosen *via* RFE were used as input for the optimal deep learning models. A few settings in the deep learning optimization were adjusted for each set of 32 batches and 100 epochs used in the experiment. The tests were repeated a total of four times.

## RESULT AND DISCUSSION

The recital of the proposed framework is assessed using several numerical indicators, including precision, recall, false-positive rate, true-negative rate ( $TR_{NG}$ ), F1-score, and Matthew's correlation coefficient (MCC). The final result of the confusion matrix determines values for parameters like true positive ( $TR_{PS}$ ), true negative ( $TR_{NG}$ ), false positive ( $FP_{PS}$ ), and false negative ( $FP_{NG}$ ). " $TR_{PS}$ " denotes a result where the proposed. All available metrics for assessment are listed below.

$$Accuracy(AC) = \frac{TR_{PS} + TR_{NG}}{TR_{PS} + TR_{NG} + FP_{PS} + FP_{NG}} \quad (15)$$

**Table 2** Sixfold cross-validation results (accuracy) on BreakHis and WBC datasets.

Dataset/Fold	BreakHis	WBC
Fold 1	99.7	99.4
Fold 2	99.6	99.5
Fold 3	99.8	99.3
Fold 4	99.7	99.6
Fold 5	99.9	99.4
Fold 6	99.8	99.7

$$Sensitivity(SE) = \frac{TR_{PS}}{TR_{PS} + FP_{NG}} \quad (16)$$

$$Specificity(SP) = \frac{TR_{NG}}{TR_{NG} + FP_{PS}} \quad (17)$$

$$Precision(PR) = \frac{TR_{PS}}{TR_{PS} + FP_{PS}} \quad (18)$$

$$Recall(RE) = \frac{TR_{PS}}{TR_{PS} + FP_{NG}} \quad (19)$$

$$F1 - score(F1s) = 2 \times \frac{PR \times RE}{PR + RE} \quad (20)$$

$$MCC = \frac{TR_{PS} \times TR_{NG} - FP_{PS} \times FP_{NG}}{\sqrt{(TR_{PS} + FP_{PS})(TR_{PS} + FP_{NG})(TR_{NG} + FP_{PS})(TR_{NG} + FP_{NG})}} \quad (21)$$

Table 2 presents the results of a sixfold cross-validation experiment conducted on two datasets: the BreakHis dataset and the WBC dataset. Cross-validation is a widely used technique in machine learning and data analysis to assess the performance and generalization ability of a model on unseen data. Each cross-validation fold involves splitting the dataset into multiple subsets, training the model on some subsets, and testing it on others. This process is repeated multiple times, and the results are averaged to obtain a more robust evaluation of the model's performance.

Table 3 demonstrates that the conceptual approach is superior to the other five models in identifying breast cancer subtypes. Pre-trained DenseNet and ResNet50 models from a five-transfer deep learning framework obtain an estimation accuracy of greater than 95.

The model's training performance is excellent even before it has been fine-tuned. Defrosting the lowest layers pre-trained with other information and re-training the model with our cancer data allows us to fine-tune the model by changing the weights of these



**Table 3** Performance analysis of the baseline models and the proposed model executed on the BreakHis and WBC datasets.

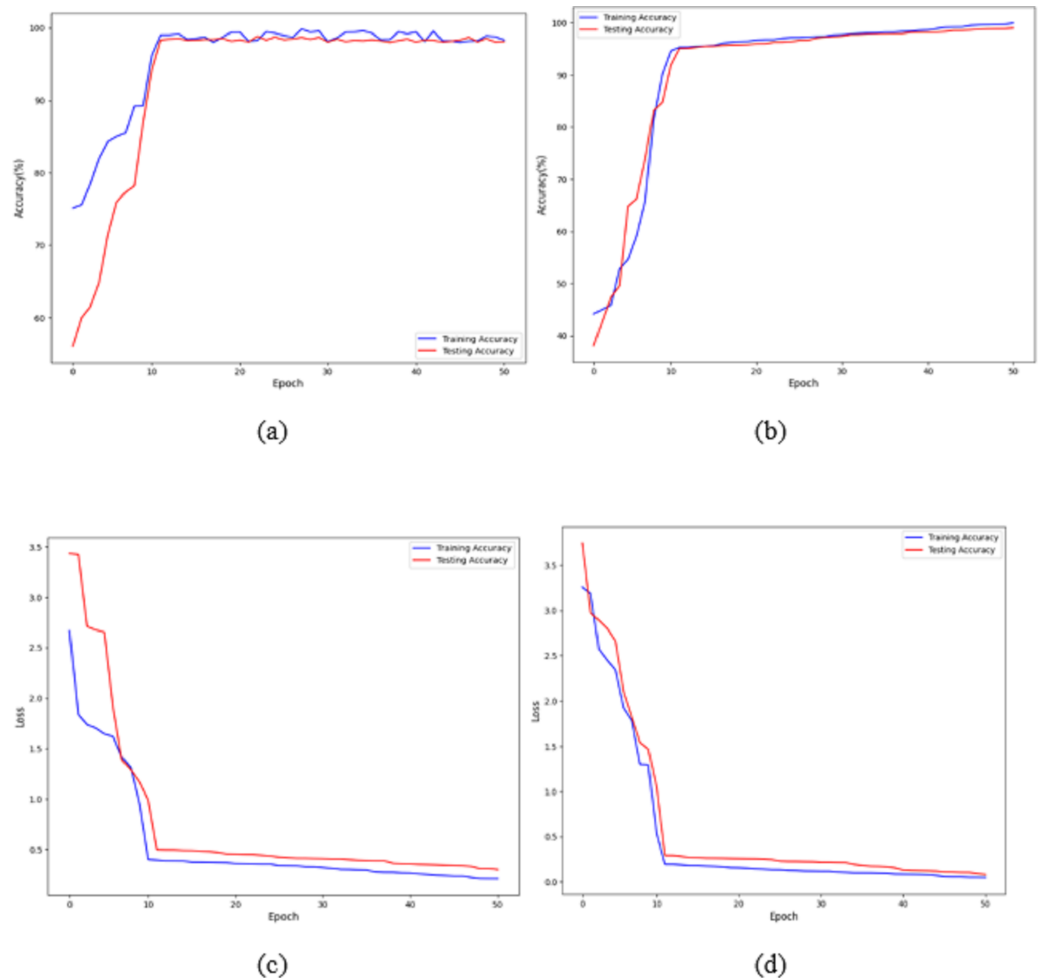
Dataset	Model	Precision	Recall	Accuracy	F1-Score	MCC
BreakHis	DenseNet	0.98	0.98	0.97	0.98	0.97
	VGG16	0.95	0.94	0.95	0.95	0.95
	MobileNet	0.94	0.93	0.94	0.94	0.94
	ResNet-50	0.96	0.97	0.96	0.96	0.97
	CNN	0.96	0.96	0.96	0.96	0.97
	Proposed model	0.99	0.99	0.99	0.99	0.99
WBC	DenseNet	0.97	0.97	0.97	0.98	0.98
	VGG16	0.96	0.97	0.96	0.96	0.96
	MobileNet	0.94	0.93	0.94	0.93	0.94
	ResNet-50	0.96	0.97	0.96	0.96	0.95
	CNN	0.95	0.96	0.95	0.94	0.94
	Proposed model	0.99	0.99	0.99	0.99	0.99

layers. Once the model has been unfrozen, it is trained at a significantly slower learning rate.

Figures 7A and 7B shows that the training and validation losses remain consistent across iterations. This indicates that with a low learning rate, the model starts to overfit. Further training would result in overfitting, where the model learns only from the characteristics present in the training set, leading to improved validation results. However, despite the model being trained to a certain extent, it performs poorly on real-world data. The previous figures clearly show that, with some adjustments, the model achieves impressive results on the data. The training loss and accuracy measures assess the model's effectiveness. As observed, the model's loss starts near zero and gradually increases to approximately 20 training batches. After that, the loss rapidly drops to almost zero, concluding the training for that cycle. The model maintains a constant momentum of 0.9 throughout the training process. However, after 20 epochs, the learning rate experiences a sharp increase. This learning process reaches a threshold represented by a variable. As the training progresses, the pace slows down again. The increased rate has a regularising effect, pushing the model away from unstable minor local minima and towards more stable wide minima. As we approach the midpoint of the cycle, the learning rate is slowed down in the hope of reaching a steady state. The next step is to search for the lowest values in the region.

The proposed architecture required 50 epochs to train and attained 100% accuracy after ten epochs on the BreakHis dataset and 99% accuracy after 12 epochs on the WBC dataset. Consequently, the overall accuracy of the validation process is 99%. The loss value remains close to 0 with no significant variations during training and validation on the BreakHis dataset, and it is 0.3 on the WBC dataset, as illustrated by the error function graph in Figs. 7C and 7D.

True positives ( $TR_{PS}$ ), true negatives ( $TR_{NG}$ ), false positives ( $FA_{PS}$ ), and false negatives ( $FA_{NG}$ ) are included in a table called the confusion matrix. When both the predicted and observed values for an outcome are positive, the former is referred to as a  $TR_{PS}$ . When a

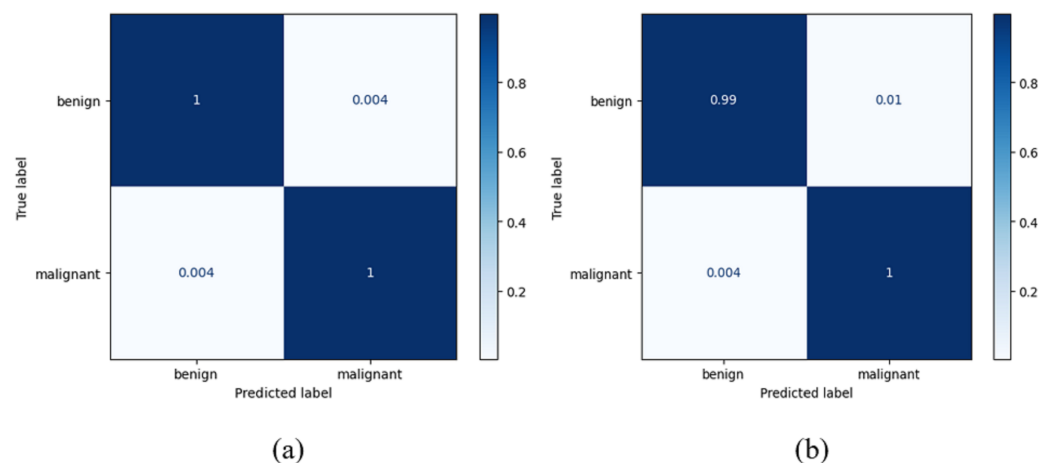


**Figure 7** Accuracy and loss of the proposed model to detect breast cancer: (A, C) BreakHis, and (B, D) WBC.

Full-size [DOI: 10.7717/peerjcs.1850/fig-7](https://doi.org/10.7717/peerjcs.1850/fig-7)

model correctly identifies an instance as negative (or 0), and the actual output is negative, we refer to this as a TruN output. When a model predicts a positive (or 1) outcome while the actual outcome is negative, we refer to the expected outcome as  $FA_{PS}$ . When a model predicts a negative (or 0) outcome while the actual result is positive, the expected result is referred to as  $FA_{NG}$ . It follows that the model's accuracy improves when the number of  $TR_{PS}$  and  $TR_{NG}$  increases (or decreases, in the case of  $FA_{PS}$  and  $FA_{NG}$ ). True positive indicates a tumour was present on the slide being examined, whereas true negative indicates no cancer. When a tumour is incorrectly diagnosed as not malignant, the result is a false negative. After training on their own, the DenseNet-169 and the improved DenseNet-169 were compared using confusion matrices. The DenseNet 169 confusion matrices are shown in Fig. 8.

Table 4 presents an analysis of the false negative rate (FNR) and false omission rate (FOR) (Gonzales-Martinez & van Dongen, 2023) for a set of baseline models and a proposed



**Figure 8** Confusion matrix for the classification of benign and malignant data from the dataset: (A) BreakHis (B) WBC.

[Full-size](#) [DOI: 10.7717/peerjcs.1850/fig-8](https://doi.org/10.7717/peerjcs.1850/fig-8)

**Table 4** FNR and FOR analysis of the baseline models and the proposed model executed on the BreakHis and WBC datasets.

Dataset	Model	FNR	FOR
BreakHis	DenseNet	0.02	0.03
BreakHis	VGG16	0.06	0.05
BreakHis	MobileNet	0.07	0.06
BreakHis	ResNet-50	0.03	0.04
BreakHis	CNN	0.04	0.05
BreakHis	Proposed model	0.01	0.01
WBC	DenseNet	0.03	0.03
WBC	VGG16	0.03	0.04
WBC	MobileNet	0.07	0.06
WBC	ResNet-50	0.03	0.04
BreakHis	CNN	0.04	0.05
BreakHis	Proposed model	0.01	0.01

model when executed on two different datasets: BreakHis and WBC. This analysis provides valuable insights into the models’ performance, particularly their ability to correctly identify positive cases and minimise the risk of false negatives and omissions.

In a study conducted by [Martinez & van Dongen \(2023\)](#), a comparison of deep learning algorithms was performed to evaluate error rates. The researchers reported a false negative rate (FNR) of 0.078 and a false omission rate (FOR) of 0.0983. This study introduces Advance Scheduling, a groundbreaking model that outperforms previous findings by achieving significantly lower false negative rate (FNR) and false omission rate (FOR). Specifically, our model demonstrates exceptional performance, with FNR and FOR recorded at an impressive 0.01 each. The findings indicate a notable enhancement in

the precision and dependability of the suggested model compared to the previously evaluated deep learning algorithms.

## CONCLUSION

Most cases can be cured if diagnosed early and treated appropriately. Computer-assisted diagnosis, especially in artificial intelligence, has dramatically facilitated early cancer identification and diagnosis. This research focused primarily on breast cancer, utilising a deep neural network to provide reliable diagnoses and classifications of breast malignancies. The study's contributions include investigating feature fusion techniques for improved feature extraction and creating a hybrid transfer learning model that outperforms current deep learning techniques by 99% accuracy. An attention-enhanced deep learning model was developed to further enhance the system's efficacy and increase the breast cancer detection rate, incorporating historical data on tumour saliency. The results showed that the proposed model achieved outstanding performance, surpassing state-of-the-art methods with 100% accuracy, 99.8% recall, and 99.06% F1-score on BreakHis, and 99% accuracy, 98.7% recall, and 99.03% F1-score on WBC. The evaluated results on FNR and FOR for Breakhis and WBC datasets show that the proposed model achieves fewer false negatives of 0.01 for both datasets. Overall, this study demonstrates the potential of deep learning to enhance the precision and timeliness of breast cancer detection and diagnosis. With further refinement, this technique could significantly save lives by enabling earlier breast cancer diagnosis and treatment. The outcomes of this system show improvement compared to earlier models. However, the system's training time is a bottleneck due to the in-depth training of the neural network. The process will run faster on computers with GPUs than on traditional hardware. Therefore, users are expected to have access to more computationally capable devices for testing and analysing their data. These strategies will play a crucial role in cancer diagnostic and prediction tasks. However, further testing and validation on larger datasets are necessary for clinical use. Their performance might benefit from more investigation into data augmentation techniques, learning in domains like the frequency domain, and implementing innovative designs like graph convolutional networks.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This study was financially supported by the Super Doctorand program and SPEV project, run by the faculty of informatics and management, at the University of Hradec Kralove, Czech Republic. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

Super Doctorand program and SPEV project, run by the faculty of informatics and management, at the University of Hradec Kralove, Czech Republic.

## Competing Interests

The authors declare there are no competing interests.

## Author Contributions

- Sudha Prathyusha Jakkaladiki conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Filip Maly conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, authored or reviewed drafts of the article, and approved the final draft.

## Data Availability

The following information was supplied regarding data availability:

The data are available at Kaggle at UC Irvine: <https://www.kaggle.com/datasets/ambarish/breakhis>.

The Diagnostic Wisconsin Breast Cancer Database (William Wolberg, Olvi Mangasarian, Nick Street, W. Street) is available at Wolberg, William, Mangasarian, Olvi, Street, Nick, and Street, W.. (1995). Breast Cancer Wisconsin (Diagnostic). UCI Machine Learning Repository. <https://doi.org/10.24432/C5DW2B>.

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.1850#supplemental-information>.

## REFERENCES

- Agnes SA, Anitha J, Pandian SIA, Peter JD. 2020. Classification of mammogram images using multiscale all convolutional neural network (MA-CNN). *Journal of Medical Systems* 44:1–9 DOI 10.1007/s10916-019-1494-z.
- Akbar S, Akram MU, Sharif M, Tariq A, Khan SA. 2018. Decision support system for detection of hypertensive retinopathy using arteriovenous ratio. *Artificial Intelligence in Medicine* 90:15–24 DOI 10.1016/j.artmed.2018.06.004.
- Chang Q-Q, Chen C-Y, Chen Z, Chang S. 2019. LncRNA PVT1 promotes proliferation and invasion through enhancing Smad3 expression by sponging miR-140-5p in cervical cancer. *Radiology and Oncology* 53(4):443 DOI 10.2478/raon-2019-0048.
- Chiang T-C, Huang Y-S, Chen R-T, Huang C-S, Chang R-F. 2018. Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation. *IEEE Transactions on Medical Imaging* 38(1):240–249.
- Dhawan AP. 2011. *Medical image analysis*. Hoboken: John Wiley & Sons.
- Fonseca P, Mendoza J, Wainer J, Ferrer J, Pinto J, Guerrero J, Castaneda B. 2015. Automatic breast density classification using a convolutional neural network architecture search procedure. In: *Medical imaging 2015: computer-aided diagnosis, volume 9414*. Cergy: SPIE, 556–563.

- Gao Z, Pan X, Shao J, Jiang X, Su Z, Jin K, Ye J. 2022. Automatic interpretation and clinical evaluation for fundus fluorescein angiography images of diabetic retinopathy patients by deep learning. *British Journal of Ophthalmology* 107(12):1852–1858 DOI 10.1136/bjo-2022-321472.
- Gevaert O, Smet FD, Timmerman D, Moreau Y, Moor BD. 2006. Predicting the prognosis of breast cancer by integrating clinical and microarray data with Bayesian networks. *Bioinformatics* 22(14):e184–e190.
- Gonzalez-Angulo AM, Morales-Vasquez F, Hortobagyi GN. 2007. Overview of resistance to systemic therapy in patients with breast cancer. *Breast Cancer Chemosensitivity* 608:1–22 DOI 10.1007/978-0-387-74039-3\_1.
- Gonzales-Martinez R, van Dongen DM. 2023. Deep learning algorithms for early detection of breast cancer: a comparative study with traditional machine learning. *Informatics in Medicine Unlocked* 41:101317 DOI 10.1016/j.imu.2023.101317.
- Hamidinekoo A, Denton E, Rampun A, Honnor K, Zwigelaar R. 2018. Deep learning in mammography and breast histology, an overview and future trends. *Medical Image Analysis* 47:45–67 DOI 10.1016/j.media.2018.03.006.
- He B, Dai C, Lang J, Bing P, Tian G, Wang B, Yang J. 2020a. A machine learning framework to trace tumor tissue-of-origin of 13 types of cancer based on DNA somatic mutation. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* 1866(11):165916.
- He B, Lu Q, Lang J, Yu H, Peng C, Bing P, Li S, Zhou Q, Liang Y, Tian G. 2020b. A new method for CTC images recognition based on machine learning. *Frontiers in Bioengineering and Biotechnology* 8:897 DOI 10.3389/fbioe.2020.00897.
- He B, Zhang Y, Zhou Z, Wang B, Liang Y, Lang J, Lin H, Bing P, Yu L, Sun D. 2020c. A neural network framework for predicting the tissue-of-origin of 15 common cancer types based on RNA-Seq data. *Frontiers in Bioengineering and Biotechnology* 8:737 DOI 10.3389/fbioe.2020.00737.
- Hossam A, Harb HM, Abd El Kader HM. 2018. Automatic image segmentation method for breast cancer analysis using thermography. *Journal of Engineering Sciences* 46(1):12–32.
- Houssein EH, Emam MM, Ali AA, Suganthan PN. 2021. Deep and machine learning techniques for medical imaging-based breast cancer: a comprehensive review. *Expert Systems with Applications* 167:114161 DOI 10.1016/j.eswa.2020.114161.
- Jiang Z-R, Yang L-H, Jin L-Z, Yi L-M, Bing P-P, Zhou J, Yang J-S. 2022. Identification of novel cuproptosis-related lncRNA signatures to predict the prognosis and immune microenvironment of breast cancer patients. *Frontiers in Oncology* 12:988680 DOI 10.3389/fonc.2022.988680.
- Joo S, Yang YS, Moon WK, Kim HC. 2004. Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple sonographic features. *IEEE Transactions on Medical Imaging* 23(10):1292–1300 DOI 10.1109/TMI.2004.834617.
- Kiyan T, Yildirim T. 2004. Breast cancer diagnosis using statistical neural networks. *IU-Journal of Electrical & Electronics Engineering* 4(2):1149–1153.

- Kousalya K, Saranya T. 2021.** Improved the detection and classification of breast cancer using hyper parameter tuning. *Materials Today: Proceedings* **81**:547–552 DOI [10.1016/j.matpr.2021.03.707](https://doi.org/10.1016/j.matpr.2021.03.707).
- Li C, Lin L, Zhang L, Xu R, Chen X, Ji J, Li Y. 2021.** Long noncoding RNA p21 enhances autophagy to alleviate endothelial progenitor cells damage and promote endothelial repair in hypertension through SESN2/AMPK/TSC2 pathway. *Pharmacological Research* **173**:105920 DOI [10.1016/j.phrs.2021.105920](https://doi.org/10.1016/j.phrs.2021.105920).
- Liu A-A, Zhai Y, Xu N, Nie W, Li W, Zhang Y. 2021.** Region-aware image captioning via interaction learning. *IEEE Transactions on Circuits and Systems for Video Technology* **32**(6):3685–3696.
- Liu Y, Zhang Y, Tan Z, Wang J, Hu Y, Sun J, Bao M, Huang P, Ge M, Chai YJ. 2022.** Lysyl oxidase promotes anaplastic thyroid carcinoma cell proliferation and metastasis mediated via BMP1. *Gland Surgery* **11**(1):245 DOI [10.21037/gs-21-908](https://doi.org/10.21037/gs-21-908).
- Lu S, Liu S, Hou P, Yang B, Liu M, Yin L, Zheng W. 2023a.** Soft tissue feature tracking based on deepmatching network. *Computer Modeling in Engineering & Sciences* **2023** **136**(1):363–379 DOI [10.32604/cmescs.2023.025217](https://doi.org/10.32604/cmescs.2023.025217).
- Lu S, Yang B, Xiao Y, Liu S, Liu M, Yin L, Zheng W. 2023b.** Iterative reconstruction of low-dose CT based on differential sparse. *Biomedical Signal Processing and Control* **79**:104204 DOI [10.1016/j.bspc.2022.104204](https://doi.org/10.1016/j.bspc.2022.104204).
- Martin LR, Williams SL, Haskard KB, DiMatteo MR. 2005.** The challenge of patient adherence. *Therapeutics and Clinical Risk Management* **1**(3):189–199.
- Martinez RG, van Dongen D-M. 2023.** Deep learning algorithms for early detection of breast cancer: a comparative study with traditional machine learning. *Informatics in Medicine Unlocked* **41**:101317 DOI [10.1016/j.imu.2023.101317](https://doi.org/10.1016/j.imu.2023.101317).
- Naji MA, El Filali S, Aarika K, Benlahmar EH, Abdelouhahid RA, Debauche O. 2021.** Machine learning algorithms for breast cancer prediction and diagnosis. *Procedia Computer Science* **191**:487–492 DOI [10.1016/j.procs.2021.07.062](https://doi.org/10.1016/j.procs.2021.07.062).
- Qi X, Zhang L, Chen Y, Pi Y, Chen Y, Lv Q, Yi Z. 2019.** Automated diagnosis of breast ultrasonography images using deep neural networks. *Medical Image Analysis* **52**:185–198 DOI [10.1016/j.media.2018.12.006](https://doi.org/10.1016/j.media.2018.12.006).
- Rajinikanth V, Satapathy SC, Fernandes SL, Nachiappan S. 2017.** Entropy based segmentation of tumor from brain MR images—a study with teaching learning based optimization. *Pattern Recognition Letters* **94**:87–95 DOI [10.1016/j.patrec.2017.05.028](https://doi.org/10.1016/j.patrec.2017.05.028).
- Rakha EA, Reis-Filho JS, Baehner F, Dabbs DJ, Decker T, Eusebi V, Fox SB, Ichihara S, Jacquemier J, Lakhani SR. 2010.** Breast cancer prognostic classification in the molecular era: the role of histological grade. *Breast Cancer Research* **12**(4):207 DOI [10.1186/bcr2607](https://doi.org/10.1186/bcr2607).
- Rivenbark AG, O'Connor SM, Coleman WB. 2013.** Molecular and cellular heterogeneity in breast cancer: challenges for personalized medicine. *The American Journal of Pathology* **183**(4):1113–1124 DOI [10.1016/j.ajpath.2013.08.002](https://doi.org/10.1016/j.ajpath.2013.08.002).
- Sharifrazi D, Alizadehsani R, Roshanzamir M, Joloudari JH, Shoeibi A, Jafari M, Hussain S, Sani ZA, Hasanzadeh F, Khozeimeh F. 2021.** Fusion of convolution neural network, support vector machine and Sobel filter for accurate detection of



- COVID-19 patients using X-ray images. *Biomedical Signal Processing and Control* 68:102622 DOI 10.1016/j.bspc.2021.102622.
- Shen X, Du S-C, Sun Y-N, Sun PZ, Law R, Wu EQ. 2023.** Advance scheduling for chronic care under online or offline revisit uncertainty. *IEEE Transactions on Automation Science and Engineering* Epub ahead of print 2023 4 September 2023 DOI 10.1109/TASE.2023.3310116.
- Smith RA, Andrews KS, Brooks D, Fedewa SA, Manassaram-Baptiste D, Saslow D, Brawley OW, Wender RC. 2017.** Cancer screening in the United States, 2017: a review of current American Cancer Society guidelines and current issues in cancer screening. *CA: A Cancer Journal for Clinicians* 67(2):100–121 DOI 10.3322/caac.21392.
- Song S, Jia H, Ma J. 2019.** A chaotic electromagnetic field optimization algorithm based on fuzzy entropy for multilevel thresholding color image segmentation. *Entropy* 21(4):398 DOI 10.3390/e21040398.
- Stone P, Lund S. 2007.** Predicting prognosis in patients with advanced cancer. *Annals of Oncology* 18(6):971–976 DOI 10.1093/annonc/mdl343.
- Sun Y, Goodison S, Li J, Liu L, Farmerie W. 2007.** Improved breast cancer prognosis through the combination of clinical and genetic markers. *Bioinformatics* 23(1):30–37 DOI 10.1093/bioinformatics/btl543.
- Sun L, Zhang M, Wang B, Tiwari P. 2023.** Few-shot class-incremental learning for medical time series classification. *IEEE Journal of Biomedical and Health Informatics* Epub ahead of print 2023 22 February DOI 10.1109/JBHI.2023.3247861.
- Sun Y-S, Zhao Z, Yang Z-N, Xu F, Lu H-J, Zhu Z-Y, Shi W, Jiang J, Yao P-P, Zhu H-P. 2017.** Risk factors and preventions of breast cancer. *International Journal of Biological Sciences* 13(11):1387 DOI 10.7150/ijbs.21635.
- Tang L, Xiang Q, Xiang J, Li J. 2022.** The haplotypes GCA and ACA in ESR1 gene are associated with the susceptibility of recurrent spontaneous abortion (RSA) in Chinese Han: a case-control study and meta-analysis. *Medicine* 101(21):e29168 DOI 10.1097/MD.00000000000029168.
- Tiwari M, Bharuka R, Shah P, Lokare R. 2020.** Breast cancer prediction using deep learning and machine learning techniques. Available at <http://dx.doi.org/10.2139/ssrn.3558786>.
- The Early Breast Cancer Trialists' Collaborative Group (EBCTCG). 1990.** Treatment of early breast cancer, vol. 1. In: *Worldwide evidence, 1985-1990*.
- Wang Y, Zhai W, Li J, Liu H, Li C, Li J. 2023.** Friction behavior of biodegradable electrospun polyester nanofibrous membranes. *Tribology International* 188:108891 DOI 10.1016/j.triboint.2023.108891.
- Wehle H-D. 2017.** Machine learning, deep learning and AI: what's the difference. In: *Data Scientist Innovation Day*. Available at [https://www.researchgate.net/publication/318900216\\_Machine\\_Learning\\_Deep\\_Learning\\_and\\_AI\\_What's\\_the\\_Difference](https://www.researchgate.net/publication/318900216_Machine_Learning_Deep_Learning_and_AI_What's_the_Difference).
- Wen L, Cheng F, Zhou Y, Yin C. 2015.** MiR-26a enhances the sensitivity of gastric cancer cells to cisplatin by targeting NRAS and E2F2. *Saudi Journal of Gastroenterology* 21(5):313 DOI 10.4103/1319-3767.166206.

- Xie X, Wang X, Liang Y, Yang J, Wu Y, Li L, Sun X, Bing P, He B, Tian G. 2021.** Evaluating cancer-related biomarkers based on pathological images: a systematic review. *Frontiers in Oncology* **11**:763527 DOI [10.3389/fonc.2021.763527](https://doi.org/10.3389/fonc.2021.763527).
- Xu X, Zhang Y, Zou L, Wang M, Li A. 2012.** A gene signature for breast cancer prognosis using support vector machine. In: *2012 5th International conference on biomedical engineering and informatics*. Piscataway: IEEE, 928–931.
- Zeng Q, Bie B, Guo Q, Yuan Y, Han Q, Han X, Chen M, Zhang X, Yang Y, Liu M. 2020.** Hyperpolarized Xe NMR signal advancement by metal-organic framework entrapment in aqueous solution. *Proceedings of the National Academy of Sciences of the United States of America* **117**(30):17558–17563 DOI [10.1073/pnas.2004121117](https://doi.org/10.1073/pnas.2004121117).
- Zhang Z, Hong W-C. 2019.** Electric load forecasting by complete ensemble empirical mode decomposition adaptive noise and support vector regression with quantum-based dragonfly algorithm. *Nonlinear Dynamics* **98**:1107–1136 DOI [10.1007/s11071-019-05252-7](https://doi.org/10.1007/s11071-019-05252-7).
- Zhang J, Shen Q, Ma Y, Liu L, Jia W, Chen L, Xie J. 2022a.** Calcium homeostasis in Parkinson’s disease: from pathology to treatment. *Neuroscience Bulletin* **38**(10):1267–1270 DOI [10.1007/s12264-022-00899-6](https://doi.org/10.1007/s12264-022-00899-6).
- Zhang Z, Wang L, Zheng W, Yin L, Hu R, Yang B. 2022b.** Endoscope image mosaic based on pyramid ORB. *Biomedical Signal Processing and Control* **71**:103261 DOI [10.1016/j.bspc.2021.103261](https://doi.org/10.1016/j.bspc.2021.103261).
- Zhou Y, Xu J, Liu Q, Li C, Liu Z, Wang M, Zheng H, Wang S. 2018.** A radiomics approach with CNN for shear-wave elastography breast tumor classification. *IEEE Transactions on Biomedical Engineering* **65**(9):1935–1942 DOI [10.1109/TBME.2018.2844188](https://doi.org/10.1109/TBME.2018.2844188).
- Zhu Y, Huang R, Wu Z, Song S, Cheng L, Zhu R. 2021.** Deep learning-based predictive identification of neural stem cell differentiation. *Nature Communications* **12**(1):2614 DOI [10.1038/s41467-021-22758-0](https://doi.org/10.1038/s41467-021-22758-0).
- Zhuang Y, Chen S, Jiang N, Hu H. 2022.** An effective WSENet-based similarity retrieval method of large lung CT image databases. *KSII Transactions on Internet & Information Systems* **16**(7).
- Zhuang Y, Jiang N, Xu Y. 2022.** Progressive distributed and parallel similarity retrieval of large CT image sequences in mobile telemedicine networks. *Wireless Communications and Mobile Computing* **2022**:1–13.
- Zuluaga-Gomez J, Al Masry Z, Benaggoune K, Meraghni S, Zerhouni N. 2021.** A CNN-based methodology for breast cancer diagnosis using thermal images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* **9**(2):131–145.