Engineering Applications of Artificial Intelligence 148 (2025) 110439

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

A patch-based deep learning framework with 5-B network for breast cancer multi-classification using histopathological images

Jehoiada Jackson [a,*], Linda E. Jackson [b], Chiagoziem C. Ukwuoma [c,d,e], Maame D. Kissi [b], Ariyo Oluwasanmi [a], Qin Zhiguang [a]

[a] School of Information and Software Engineering, University of Electronic Science and Technology of China, Sichuan P.R., China
[b] School of Medicine and Dentistry, University of Ghana, Legon Boundary, Accra, Ghana
[c] College of Nuclear Technology and Automation Engineering, Chengdu University of Technology, Sichuan P.R., 610059, China
[d] Sichuan Engineering Technology Research Center for Industrial Internet Intelligent Monitoring and Application, Chengdu University of Technology, Sichuan P.R., 610059, China
[e] Oxford Brooks University, Sino-British Collaborative Education, Chengdu University of Technology, Sichuan P.R., 610059, China

ARTICLE INFO

Keywords:
Deep learning
Vision transformers
Breast cancer
ConvMixer
Multi-classification

ABSTRACT

Despite the fact that convolutional networks have long been the preferred architecture for investigating breast cancer (BC), new research has revealed that transformer-based architectures perform better in specific circumstances. Currently, Vision Transformer (ViT) has shown to be the most successful transformer-based architecture in vision tasks by employing patch encoding. Drawing inspiration from the patch-based modeling, this study proposed a Patch-Based deep learning network with direct operation on patches as input and the segregation of blending of spatial and channel parameters without maintaining the same size and resolution across the structure as well as simply employing conventional convolutions to carry out the mixing phases. This study further introduced a novel 5-B Network at the end of the pointwise convolutional layer for tiny feature extraction. The 5-BNet comprises five branches that process information concurrently. The primary distinction among these branches lies in the size of their convolutional kernels. In order to capture high-level image features, the 5-BNet gradually reduces the filter size of each convolution layer in every concurrent branch. Based on the proposed model, an end-to-end training for breast cancer (BC) multi-classification using the publicly available BreakHis (Benign class and Malignant Class) is carried out. In addition, an Eight-class multi-classification set by incorporating the benign four classes and the malignant four classes to further evaluate the robustness of the proposed model in multi-classification tasks. Also, the proposed model visualized internal composition to demonstrate how it understood the different patterns of the input images is illustrated. The proposed model obtained 98.1 ± 1.0% accuracy for eight classes and 98.01 ± 1.0% for four classes on all magnifications. The experimental results show that the proposed method achieves the highest breast classification accuracy when compared to cutting-edge models.

1. Introduction

Breast Cancer (BC) is a widespread and serious disease primarily affecting middle-aged women globally, ranking as the second leading cause of cancer-related deaths worldwide after lung cancer (Bray et al., 2018). According to the World Health Organization (WHO), out of 1, 350,000 recorded BC cases annually, 460,000 result in fatalities (Wang et al., 2020a). In 2019, the United States alone reported 268,600 BC cases (Ahmad, 2019) (Man et al., 2020). The disease, primarily driven by abnormal cell growth in middle-aged women, poses a significant health risk, with WHO predicting an 8% chance of women experiencing BC in their lifetime (Ahmad et al., 2022). The breast's complex anatomy, consisting of blood vessels, tendons, ligaments, milk ducts, lacrimal glands, and lymph ducts, provides various avenues for cancer development, necessitating accurate detection and classification for proper diagnosis (Mansha et al., 2022) (Honkanen et al., 2021). Imaging methods, including magnetic resonance imaging (MRI), diagnostic mammography (X-rays), thermography, and ultrasound, play a crucial role in BC analysis (Houssein et al., 2021)– (Arhatari et al., 2021). The differentiation between benign and malignant tumors, which vary in

* Corresponding author.
E-mail address: kofijackson@uestc.edu.cn (J. Jackson).

https://doi.org/10.1016/j.engappai.2025.110439
Received 24 December 2022; Received in revised form 1 November 2024; Accepted 25 February 2025
Available online 3 March 2025

size, shape, and position, presents a formidable challenge in identifying and localizing cancer cells in breast images (Houssein et al., 2021)–(Ukwuoma et al., 2022a).

Rationale: Early detection and accurate diagnosis of breast cancer (BC) are essential for effective treatment and improved survival outcomes. Traditional methods, including screening tests, self-exams, and timely specialist consultations, play a critical role in slowing disease progression and lowering mortality rates. Radiologists typically rely on clinical techniques such as mammography and biopsy to identify early signs of cancer (Choudhary et al., 2021) (Tsapatsaris et al., 2021). However, while various image processing techniques have been developed to enhance these methods (Ukwuoma et al., 2022b) (El Agouri et al., 2022) (Iqbal et al., 2022), they still largely depend on manual interpretation, making them less feasible in resource-limited or remote areas. This underscores a need for an automated, highly accurate system for BC detection that can operate effectively even in settings with limited access to specialized expertise.

Motivation: Developing an efficient and automated system for BC detection is crucial to enhance both diagnostic efficiency and accuracy, especially in regions where medical resources are scarce and rapid intervention is challenging. Advances in computer-aided diagnostics (CAD) and artificial intelligence (AI), particularly through machine learning (ML) and deep learning (DL), offer promising avenues for improving BC detection. However, despite these advancements, existing approaches face limitations. For instance, while convolutional neural networks (CNNs) are commonly used for feature extraction and classification, they can struggle with accurately capturing complex image patterns, which may result in incorrect evaluations. Moreover, CNN-based methods risk missing critical factors that influence classification outcomes, potentially affecting the reliability of BC diagnoses.

The proposed approach aims to address these gaps by leveraging advanced DL techniques tailored for BC detection, optimizing feature extraction to capture intricate patterns in medical images and enhance classification accuracy. By overcoming the limitations of current CNN architectures, this approach seeks to improve detection reliability across diverse clinical settings, ultimately contributing to more effective BC diagnosis and treatment.

New research has revealed that transformer-based architectures, particularly the Vision Transformer (ViT), may perform better in specific circumstances than convolutional networks in vision tasks. A unique approach that sheds insight into image processing and deep learning for the particular classification of BC from histopathological images was proposed by Thomas et al. (2022). Modern CNN performs worse than the vision transformer, which has a 96% accuracy rate in identifying BC cells. A novel deep learning model named Multi-scale Dual Residual Recurrent Network (MTRRE-Net) was put forward by Chattopadhyay et al. (2022a) for the classification of BC using histopathology images. Using a dual residual block in conjunction with a recurrent network, this model tackles the vanishing gradient issue in deep networks and offers a unique way to improve performance even in deep topologies. ViTs uses patch encoding, which combines discrete portions of the image into a single data feature, to change to bigger dimensions owing to the exponential latency of the self-attention stacks in Transformers. Sequel to the wonderful performance of these models, this study proposed a patch-based model that has some characteristics with the ViT and the ConvMixer, including their direct operation on patches as input and the segregation of blending of spatial and channel parameters, although this study did not maintain the same size and resolution across the structure and employs conventional convolutions to carry out the mixing phases. The proposed approach starts with a patch embedding layer and then applies a fully convolutional block repeatedly with varying feature dimensions. The patch embeddings' spatial organization was kept constant. Convolution with c_{in} input channels, h output channels, kernel size p, and stride p may be used to construct patch embeddings with a patch size of p and an embedding dimension of h. The convolutions used in the proposed model include depthwise convolution, separable convolution, and pointwise convolution. This study further introduced a novel 5-B Network with five branches that process information concurrently, each with a distinct input patch to extract the most discriminative top-level characteristics, to tackle the multi-classification problem of BC. The primary distinction among these branches lies in the size of their convolutional kernels. In order to capture high-level image features, the proposed model gradually reduces the filter size of each convolution layer in every concurrent branch. This study summarizes the major contributions as follows:

- ❖ This study proposed a simple but effective patch-based convolutional network to address the multi-classification challenge of breast cancer by extracting discriminative top-level characteristics.
- ❖ In addition to separating "channel-wise blending" from "spatial blending" of information, the proposed model acts directly on patches with varying resolution and size representation throughout all levels without losing useful features at subsequent layers.
- ❖ Furthermore, a 5-B network comprising five branches that process information concurrently for top-level feature characteristics of the input image along the patch-extracted features is proposed. The primary distinction among these branches lies in the size of their convolutional kernels. In order to capture high-level image features, the proposed model gradually reduces the filter size of each convolution layer in every concurrent branch.
- ❖ The 5-B network is significant within the broader context of deep learning in medical imaging as it addresses the unique challenge of multi-classification in complex medical datasets by efficiently capturing spatial and channel-wise information. By processing multiple image patches concurrently, this model preserves essential detail across varying resolutions, a common issue in traditional single-path networks. The multi-branch architecture allows for a comprehensive feature extraction approach, enhancing the model's ability to classify diverse breast cancer subtypes.
- ❖ This study went on to give a thorough analysis of the BreakHis dataset via Four-class and Eight-class multi-classification. A very effective deep-learning strategy for the multiclassification of breast cancer disorders that performs well in terms of the used evaluation metrics based on the complete experimental evaluation is described in this paper.

The rest of this study is structured as follows: section 2 presents recent studies, section 3 describes our methodology and material and section 4 presents a comprehensive analysis of results. This study finally concludes in section 5.

2. Related works

The least reliable way to find any growth in a certain part of the breast is to examine the breast. Mammography has evolved as a replacement and is utilized extensively in the medical industry. However, there is a substantial danger of false positives when depending solely on mammograms, which frequently result in needless biopsies and surgeries (Nassif et al., 2022). The medical industry over the past years has contributed to the detection and classification of BC. This section reviews the various works carried out on BC analysis based on deep learning approaches. Iqbal et al. (2022) evaluate public datasets about breast cancer diagnosis and the current deep-learning techniques for breast cancer detection to highlight the difficulties and future work for deep learning-based diagnosis. Additionally, publicly accessible code repositories are introduced. Naik et al. (2020) developed a multiple instances learning-based deep neural network that determines ERS from H&E-stained whole slide images (WSI), achieving an area under the curve (AUC) of 0.92 for sensitivity and specificity. This showed that machine learning models can determine molecular marker status, as assessed by hormone receptors, directly from cellular morphology.

Convolutional neural network models have shown to be promising in

computer vision tasks, especially medical imaging. A convolutional neural network (CNN) was employed by Madduri et al. (Chattopadhyay et al., 2022a). to extract features for BC classification before using a support vector machine (SVM) classifier on the retrieved features. Abdel-Zaher et al. (Abdel-Zaher and Eldeib, 2016) and Han et al. (2017), identified labeling problems of mitotic and non-mitotic nuclei in histopathology images which Mathew et al. (2021) tried to address this by providing a two-stage deep CNN technique in a three-stage classification procedure that includes patch-level classification, a refinement model, and heatmap-based postprocessing. To classify breast tumors from ultrasound images into the "Normal," "Benign," and "Malignant" categories, Karthik et al. (2022a) proposed a novel configuration of a Stacking Ensemble with custom Convolutional Neural Network architectures, yielding accuracy, f1-score, precision, and recall of 92.15%, 92.21%, 92.26%, and 92.17%, respectively. Liu et al.' 's (Liu et al., 2022a) innovative deep feature creation approach improves the ability to identify breast cancer from BUS (breast ultrasound) pictures. Using a grid-based methodology, they use 16 pre-trained Convolutional Neural Network (CNN) models as feature generators. The input image is split up into rows and columns, and the trained models process each area separately. The final feature vector is formed by the grid-based deep feature generator, which determines the error values for each generator and chooses the top three feature vectors. To produce deep features, Kaplan et al. (2022) presented a complex deep model that combines transfer learning from three pre-trained networks with a special pyramid triple deep feature generator (PTDFG). Bilinear interpolation is applied to the input image to create a four-level pyramid. These pyramid levels are then used by the pre-trained networks for downstream feature creation and neighborhood component analysis. These results are then sent into a support vector machine classifier, which uses a ten-fold cross--validation strategy, for automated classification.

Wang et al. (2020b), improved the early detection of BC by combining DL and ML frameworks to create a classification framework for histology images. They created a multi-network feature extraction model utilizing pre-trained deep convolutional neural networks (DCNNs), a feature dimension reduction method, and an ensemble support vector machine (E-SVM). Khan et al. (Ibrokhimov et al., 2020) used a transfer learning (TL) method to develop a new DL-based method for detecting and predicting BC. It shows that the proposed method surpasses the vast majority of prior DL methods in terms of detecting and predicting BC in cytology images. Moreover (Liu et al., 2019), described a strategy for BC classification that employed an information-gain-directed simulated annealing genetic algorithm wrapper (IGSAGAW) for feature selection. The IG method was used to rank the features, and the cost-search support vector machine (CSSVM) learning method was utilized to find the five best characteristics. Al-Haija et al. (Abu Al-Haija and Manasra, 2020) categorized histopathology microscope images and presented a reliable and comprehensive computational BC detection system based on the ResNet-50 CNN. The proposed model trains and classifies the BreakHis dataset using a TL technique of the powerful ResNet-50 CNN pre-trained on ImageNet. Their simulation results demonstrated that the proposed model outperformed other models trained on the same dataset with extraordinary classification accuracy. Soumik et al. (2021) utilize a TL technique based on DL models to categorize BC. The technique recorded an accuracy of 0.995, 0.989, 0.989, and 0.985 were recorded for magnification factors of 40x, 100x, 200x, and 400x. Mahmoud et al. (2021) classify breast tumors using deep CNN with an attention mechanism. Additionally, Boumaraf et al. (2020) proposed a new TL-based mechanism for magnification-dependent and independent BCs. To properly diagnose tumors in the breast, Jayanthi et al. (Jayandhi and Leena Jasmine, 2022) utilize a very deep convolutional network (VGG) and SVM on a public tumor dataset. They optimize the performance accuracy of their model to 0.98.

The method of combining two or more CNN architectures for performance gain has attracted wide attention. This method referred to as an ensemble approach allows researchers to combine the prediction of several models into one prediction either by voting, averaging, etc. An ensemble of two CNN architectures combined with Channel and Spatial attention was proposed by Karthik et al. (2022b). Parallel feature extraction from the histopathological pictures is achieved by two robust proprietary deep architectures: CSAResnet and DAMCNN. Ultimately, group education is used to enhance performance even further. On the BreakHis dataset, the suggested framework was able to obtain a 99.55% classification accuracy. To analyze histopathology images, Shankar et al. (2022) provide a unique chaotic sparrow search method combined with a deep transfer learning-enabled breast cancer classification (CSSADTL-BCC) model. The CSSADTL-BCC model generated feature vectors by using a MixNet-based feature extraction model and Gaussian filtering to remove noise. The chaotic sparrow search algorithm (CSSA) helps with hyperparameter tuning while the stacked gated recurrent unit (SGRU) method is used for classification.

Attention mechanism-based frameworks especially Transformers have shown impressive performance for vision tasks in which medical imaging is not left out. Researchers such as Brancati et al. (2021) used CNN with a reduction path and a learning path to examine gigapixel histopathology images. Specifically, the gigapixel image is compressed in the compression path by employing a residual network dedicated to the feature extraction of each patch into which the image has been partitioned. Attention modules are applied to the grid-based feature map in the learning path to locate regions of interest, which are subsequently employed for the final full slide analysis while taking into consideration the spatial correlations of nearby patch features. Togaçar et al. (Tougaçar et al., 2020) introduced a unique DL model called BreastNet that is based on CNN for the classification of BC. The hyper-column method attention component and the residual block are the two components that make up the model. When applied to the BreakHis dataset, the model achieves 98.80% accuracy. Karthik et al. (2022b) proposed a new algorithm to address the time-consuming manual tumor delineation in breast histopathology analysis. They introduced an ensemble of two custom CNN architectures with Channel and Spatial attention mechanisms. The algorithm successfully achieves a notable classification accuracy of 99.55% on the BreakHis dataset, effectively improving the efficiency of breast cancer classification. Chattopadhyay et al. (2022b) to best detect breast tumors. Their proposed model incorporates a channel attention mechanism, which improves the model's capacity to learn the intricate patterns of images. This mechanism was inspired by the bottleneck unit of the ShuffleNet design. Despite being trained on a very tiny dataset, the model's tightly connected blocks handle both the overfitting and the vanishing gradient problems. A unique approach that sheds insight on image processing and deep learning for the particular classification of breast cancer from histological images was proposed by Thomas et al. (2022). Modern convolution neural networks perform worse than the vision transformer, which has a 96% accuracy rate in identifying breast cancer cells. Recent studies continue to substantiate the benefits of transformer-based architectures and patch-based techniques in medical imaging. Xu et al. (Xu and Li, 2022) proposed a patch-based vision transformer model specifically tailored for histopathological image analysis, showing improved accuracy over traditional CNNs in BC classification tasks.Wang and Zhou (2022) presented a hybrid deep learning approach combining CNNs and transformers for multi-scale feature extraction, achieving a remarkable 97.8% accuracy in multi-class BC classification.Patel et al. (2022) developed a self-supervised transformer model, which leverages unlabeled histopathology images to improve BC classification accuracy by a substantial margin, demonstrating a 96.7% accuracy on diverse datasets. Huang et al. (2022) utilized a vision transformer with cross-attention mechanisms to effectively capture inter-patch relationships, resulting in a 98% classification accuracy on BC datasets, highlighting the potential of transformers for complex image analyses. Yang et al. (2022) introduced a transformer-based model for multi-modal BC classification, incorporating both image and patient metadata, achieving state-of-the-art

performance with a 98.3% accuracy. Rahman and Khan (2022) developed a patch-based transformer model with an adaptive feature aggregation technique, showing significant improvements in BC detection accuracy, especially for challenging cases with mixed tissue characteristics.

Research Gap: Convolutional neural networks (CNNs) have completely changed medical imaging, especially when it comes to the field of histopathological image processing for the diagnosis of breast cancer. While CNN models have shown encouraging results, more powerful alternatives in the form of transformer-based techniques more precisely, patch-based techniques have emerged as a result of recent developments in deep learning. Furthermore, the shift to multiclass situations in the BC domain which involves eight different BC classes remains a somewhat uncharted area. The majority of research that has been done so far has focused on binary or four-class classification, which has left a clear vacuum in the literature when it comes to the use of patch-based transformer techniques for multiclass breast cancer classification. In order to highlight the lack of focus on the subtleties of multiclass classification in breast cancer research, this study proposes a novel Patch-Based Deep Learning Framework for multiclass breast cancer classification that makes use of a 5-B network. Our methodology not only leverages transformer-based designs to handle the complexities of a multiclass classification situation involving eight different classes but also expands the capabilities of CNNs beyond what they can do today.

3. Materials and methods

The authors of the "Patches Are All You Need" article (Dosovitskiy et al., 2021) developed the concept of patch training for an all-convolutional network and presented compelling results. Their design, called ConvMixer, borrows concepts from more recent isotropic architectures like ViT and MLP-Mixer (Vaswani et al., 2023), including residual connections, utilizing the same depth and magnification throughout the network's various layers. Drawing inspiration from patch-based models and residual connections, this study developed a simple patch-based deep learning network with varying convolution feature sizes and resolutions, as well as a novel 5-B Network comprising five branches that process information concurrently for top-level feature characteristics of the input image along the patch-extracted features, is proposed. The primary distinction among these branches lies in the size of their convolutional kernels. In order to capture high-level image features, the proposed model gradually reduces the filter size of each convolution layer in every concurrent branch Fig. 1 depicts the flow chart of the work done in this research.

3.1. Proposed model

The proposed architecture starts with a patch embedding layer and then applies a straightforward fully-convolutional block repeatedly with diverse feature sizes. The patch embeddings' spatial organization is kept constant as seen in Fig. 2. Convolution with c_{in} (input channels), h (output channels), kernel size p, and stride p is used to create patch embeddings with a patch size of p and an embedding dimension of h;

$$z_0 = BN(\sigma)\{Conv_{c_{in}\to h}(X, stride = p, kernal_{size} = p)\} \quad (1)$$

Convolution is performed in three different ways in the proposed model; depthwise (i.e., grouped convolution with groups equal to the number of channels, h), separable convolution, and pointwise convolution (i.e., kernel size 1×1). For depthwise convolution, the proposed model performs best with extremely large kernel sizes. An activation and post-activation BatchNorm come after each convolution. After 64 features, the Depthwise convolution is followed by the separable convolution. Following a ReLU activation and post-activation BatchNorm, each separable convolution;

$$z'_l = BN(\sigma\{ConvDeptwise(z_{l-1})\}) \quad (2)$$

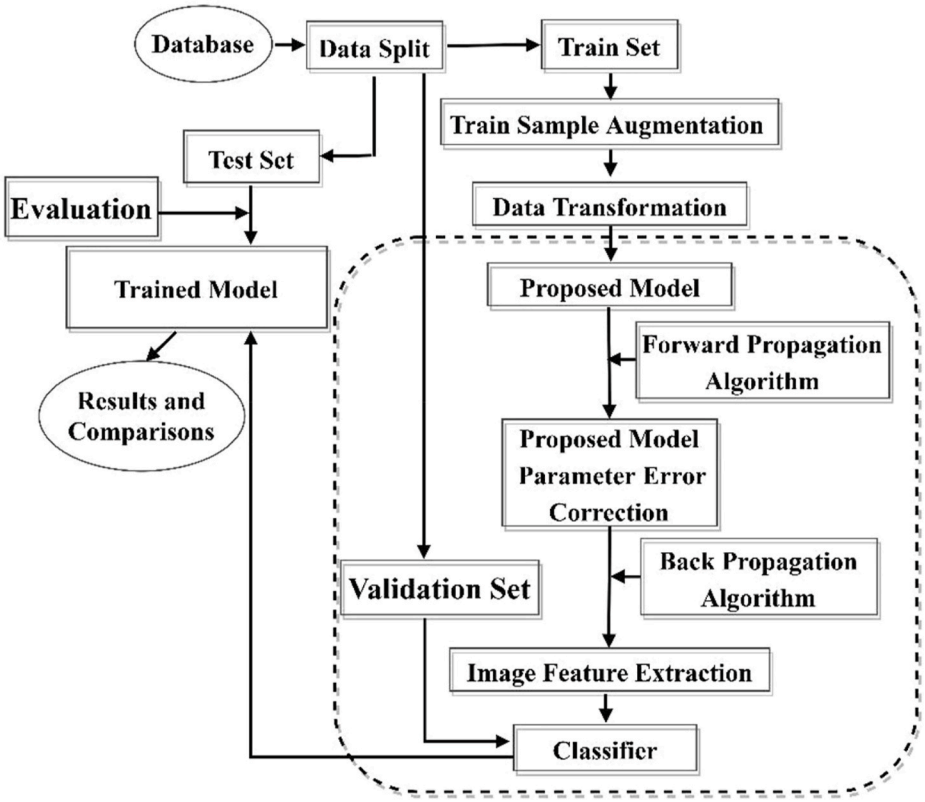

Fig. 1. The proposed model features an extraction flowchart.

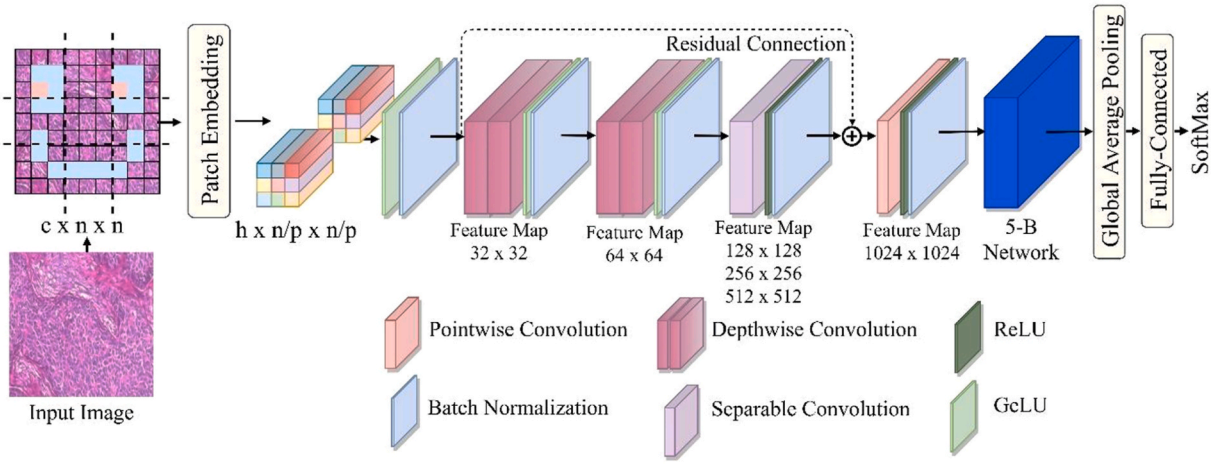

Fig. 2. Proposed Architecture. The proposed model applies the copies of a straightforward fully-convolutional block that consists of large-kernel depthwise convolution for 32 and 64 depth, separable convolution from 128 to 512 depth followed by pointwise convolution and the novel 5-B network, global pooling, and a straightforward linear classifier to preserve locality.

$$z_l^{ii} = z_l^{i} + \{ConvSeperable(z_{l-1})\} + z_{l-1} \tag{3}$$

$$z_l^{iii} = BN\left(\sigma\left\{ConvPointwise\left(z_l^{ii}\right)\right\}\right) \tag{4}$$

$$z_{l-1} = BN\left(\left\{5B\_Network\left(z_l^{iii}\right)\right\}\right) \tag{5}$$

The output of the pointwise convolution is the input of the 5-B Network as shown in Fig. 3, which has five branches running concurrently. Each branch was composed of three layers: batch normalization, a convolution with distinct filter sizes, and an activation layer. Each convolution layer's filter size was gradually decreased from 13x13 to 5x5 in every concurrent branch of the 5-B Network in order to extract the high-level input characteristics. Equation (6), presents the mathematical definition of five concurrent branches.

$$5B_i = \left(\left(\left(fil_{2,2}^{i\times j} * \left(pointwise_{output} \cdot Weight\right) + bais\right)\sigma_{ReLUi}\right)B_{norm}\right) \tag{6}$$

The parameter $5B_i$ = the number of the branch, i = one to five branches, $fil_{2,2}^{i\times j}$ = the size of the filter with the stride of (Wang et al., 2020a; Wang et al., 2020a, 2020a)," and " ixj" denotes the filer sizes of five concurrent branches as $5x5$, $7x7$, $9x9$, $11x11$, *and* $13x13$, respectively. Each branch *i's* activation function is represented by σ_{ReLUi} and its batch normalization layer is represented by B_{norm}. After five branches were processed concurrently, a new layer was added to continue the feature learning procedure. A flattening operation is performed after many iterations of this block to obtain a feature vector of size h, after which it is sent to a SoftMax classifier after a dropout of 0.5 layers is passed.

Like (Dosovitskiy et al., 2021) and others, the proposed architecture is built on the concept of blending. This study specifically selected pointwise convolution for mixing channel locations and depthwise and separable convolution for mixing spatial locations. The ability of MLPs and self-attention to combine distant geographical places, or to have an arbitrary broad receptive field, is an important insight from prior research. Therefore, to combine far-off spatial locations, we employed convolutions with an exceptionally high kernel size. Self-attention and MLPs are technically more adaptable, permitting wide activation functions and information behavior, but convolution's intuitive bias is well-suited to vision tasks and results in great data throughput. By employing such a routine technique, this study exposes how the patch representation differs from the typical spire, gradually downsampling the architecture of convolutional networks. Three factors determine how the proposed model is initialized: (1) the size of the patch embeddings, (2) the patch size p, which determines the model's intrinsic granularity, and (3) the convolutional layers (Depthwise Convolution and the Separable convolution) kernel size k. Internal resolution is

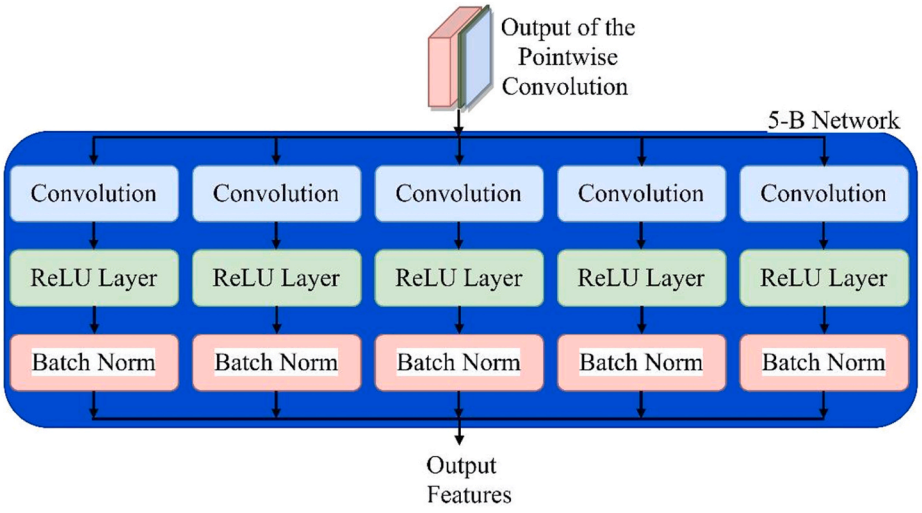

Fig. 3. Illustration of the implemented Novel 5-B Network running concurrently.

defined as the input vector size n divided by the patch size p; however, the Model accepts variable-sized inputs. We demonstrate the convolutional blocks used in our model in Fig. 4.

Relevance to Feature Extraction: The multi-branch structure of the 5-B Network is particularly beneficial for feature extraction in medical imaging because it addresses the need for a varied and extensive feature set to capture subtle differences in medical images. In traditional networks, a single kernel size may not be sufficient to capture both small details and larger contextual information. By having branches with progressively smaller kernel sizes, the 5-B Network achieves a hierarchical feature extraction process.

- **Large Kernels (e.g., 13x13):** Capture broad, coarse features, useful for understanding the overall structure of the image.
- **Small Kernels (e.g., 5x5):** Capture fine details, which are crucial for identifying subtle patterns and irregularities that may indicate disease.

The concurrent processing within the 5-B Network allows for more comprehensive feature extraction, leading to improved classification accuracy. After the concurrent branches process the features, they are aggregated and passed through additional layers, including a SoftMax classifier, to produce the final output.

Training Setup: For the size [32–64], the activation block used is the GeLu activation function while using the ReLU activation function for size [128–1024]. For the augmentation techniques used via the Keras Image data generator pipeline, the input images are first rescaled to 1./255, rotation range of 40 with horizontal flip and shuffle set to True for the train set and only rescaled to 1./255 and shuffle = False for the test and validation set. The LR_reduce using Keras.callbacks.ReduceLROnPlateau with monitor = 'val_loss', factor = 0.3, epsilon = 0.0001, patience = 5, verbose = 1, mode = 'min' was used. For the callback, the Keras.callbacks.EarlyStopping is used with monitor = 'val_loss', mode = 'min', min_delta = 0, patience = 10, restore_best_weights = True. A batch size of 8 and an image size of 224 are used. The classifier activation is set to SoftMax (multi-class). During model compilation, optimizer = Keras.optimizers.Adam (learning_rate = 0.001, clipvalue = 0.2), loss = 'categorical_crossentropy', metrics = ['accuracy'] (see Table 1).

3.2. Dataset and dataset preprocessing

BreaKHis, the biggest dataset of BC histopathology images currently accessible, was published by Man, R et al. (Man et al., 2020). The information was contributed by the Pathological Anatomy and Cytopathology (P&D) Lab in Brazil. The diagnosis of 82 people produced microscopic images of normal and cancerous tissue at various magnifications. There are 7909 total images, 2480 BI, and 5429 MI. The generated microscopic images have four distinct magnifications: 40, 100, 200, and 400. Fig. 4 displays a visual depiction of the BreaKHis dataset. It displays the binary separation of each class into subclasses that are benign and malignant. The benign categories include adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenoma (TA); the malignant classifications include ductal carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC), and papillary cancer (PC). The data split for the BreaKHis (Man et al., 2020) data is shown in Table 2. Fig. 5 shows some of the employed dataset samples.

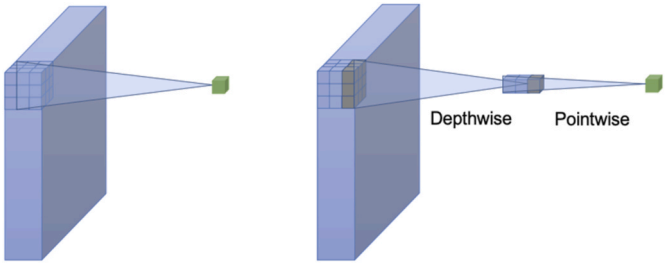

Fig. 4. Illustration of the Convolution blocks used in the implementation. The proposed model made use of the Depthwise Convolution on depths 32, 64, and. For depth in 128– 512, the Separable convolution is used. The pointwise convolution was used only at depth 1024.

First, the used dataset was divided into the train, validation, and test sets in the proportions of 70:20:10. Data augmentation was applied to 1002 samples of only the train set using the Python Augmentor package. The Augmentor Python package was used to upsample the train sets and

Table 1
Basic Keras Implementation of the Proposed Model Excluding the Dataset Preprocessing and the Final Model Hyperparameters Settings Presented in this Paper.

```
def gelu(x):
return 0.5*x*(1+tf.tanh(np.sqrt(2/np.pi)*(x+0.044715*tf.pow(x, 3))))
def activation_block(x):
x = layers.Activation("gelu")(x)
return layers.BatchNormalization()(x)
def trunk (x, patch_size: int):
x = layers.Conv2D(32, kernel_size = patch_size, strides = patch_size)(x)
return activation_block(x)
def conv_block(x, kernel_size: int):
filters = [32, 64, 128, 256, 512]
# Depthwise convolution.
for filter in filters[:2]
x0 = x
x = layers.DepthwiseConv2D(kernel_size = kernel_size, padding = "same")(x)
x = layers.Add()([activation_block(x), x0]) # Residual.
for filter in filters[3:]
x = layers.Activation("relu")(x)
x = layers.SeparableConv2D(filter, 3, padding = "same")(x)
x = layers.BatchNormalization()(x)
x = layers.MaxPooling2D(3, strides = 2, padding = "same")(x)
# Project residual residual = layers.Conv2D(filter, 1, padding = "same")
(activation_block(x))
x = layers.add([x, residual]) # Add back residual
x = layers.SeparableConv2D(1024, 3, padding = "same")(x)
x = layers.BatchNormalization()(x)
x = layers.Activation("relu")(x)
return x
def Model(image_sizee = 224, kernel_size = 5, patch_size = 2, num_classes =
8):
inputs = keras.Input((image_size, image_size, 3))
# Extract patch embeddings. x = conv_trunk(inputs, patch_size)
# Conv blocks. x = conv_block(x, kernel_size)
# 5-B Network x = Five_B_Network(x, kernel_size)
# Classification block. x = layers.GlobalAvgPool2D()(x)
activation = "softmax"
units = num_classes
x = layers.Dropout(0.5)(x)
outputs = layers.Dense(units, activation = activation)(x)
return keras.Model(inputs, outputs)
my_model = Model()
```

Table 2
An explanation of the BreakHis (Man et al., 2020) dataset, including the initial partition, splitting, and sample counts for each split.

Class	Sub_Class	Magnification				Total	Nos_Patients
		40x	100x	200x	400x		
Benign	Adenosis	114	113	111	106	444	24
	Fibroadenoma	253	260	264	237	1014	
	Phyllodes_tumor	109	121	108	115	453	
	Tubular_adenoma	149	150	140	130	569	
Malignant	Ductal_carcinoma	864	903	896	788	3451	58
	Lobular_carcinoma	156	170	163	137	626	
	Mucinous_carcinoma	205	222	196	169	792	
	Papillary_carcinoma	145	142	135	138	560	
Total		1995	2081	2013	1820	7909	82
Augmented-Train							
Benign	Adenosis	1002	1002	1002	1002	4008	–
	Fibroadenoma	1002	1002	1002	1002	4008	
	Phyllodes_tumor	1002	1002	1002	1002	4008	
	Tubular_adenoma	1002	1002	1002	1002	4008	
Malignant	Ductal_carcinoma	1002	1002	1002	1002	4008	–
	Lobular_carcinoma	1002	1002	1002	1002	4008	
	Mucinous_carcinoma	1002	1002	1002	1002	4008	
	Papillary_carcinoma	1002	1002	1002	1002	4008	
Test							
Benign	Adenosis	11	11	11	10	43	–
	Fibroadenoma	25	25	26	23	99	
	Phyllodes_tumor	10	12	10	11	43	
	Tubular_adenoma	14	14	13	12	53	
Malignant	Ductal_carcinoma	86	90	89	78	343	–
	Lobular_carcinoma	15	16	16	13	60	
	Mucinous_carcinoma	20	22	19	16	77	
	Papillary_carcinoma	14	14	13	13	54	
Validation							
Benign	Adenosis	98	98	98	98	392	–
	Fibroadenoma	98	98	98	98	392	
	Phyllodes_tumor	98	98	98	98	392	
	Tubular_adenoma	98	98	98	98	392	
Malignant	Ductal_carcinoma	98	98	98	98	392	–
	Lobular_carcinoma	98	98	98	98	392	
	Mucinous_carcinoma	98	98	98	98	392	
	Papillary_carcinoma	98	98	98	98	392	

apply the following transformations: crop, random rotation, random flipping, skew, shear, zoom, etc.

3.3. Evaluation metrics

Several assessment metrics were used to gauge how robust the proposed model was. The metrics include;

$$Accuracy = \frac{TP + TN}{(TP + TN) + (FP + FN)} * 100 \tag{7}$$

$$Precision = \frac{TP}{TP + FP} * 100 \tag{8}$$

$$Specificity = \frac{TN}{N} * 100 = \frac{TN}{TN + FP} * 100 \tag{9}$$

$$Sensitivity = \frac{TP}{P} * 100 = \frac{TP}{TP + FN} * 100 \tag{10}$$

$$F_1\ score = \left(\frac{SEN^{-1} + PRC^{-1}}{2}\right)^{-1} = \frac{2 * TP}{2 * TP + FP + FN} \tag{11}$$

The probability curve created by plotting at various threshold levels is referred to as the ROC (Receiver Operating Characteristic).

4. Results

This section presents the classification results of the proposed model alongside the result discussion and comparison with the state-of-the-art results. In addition, the performance of the proposed model on the different classes is also discussed in this section.

4.1. Classification results

The proposed model is implemented and evaluated in two multiclass scenarios; first, the class scenario which includes the evaluation of the benign class and the malignant classes, and second the Eight-class multiclassification which is the main objective of the proposed model. The evaluation metrics include Accuracy denoted as AC, Sensitivity

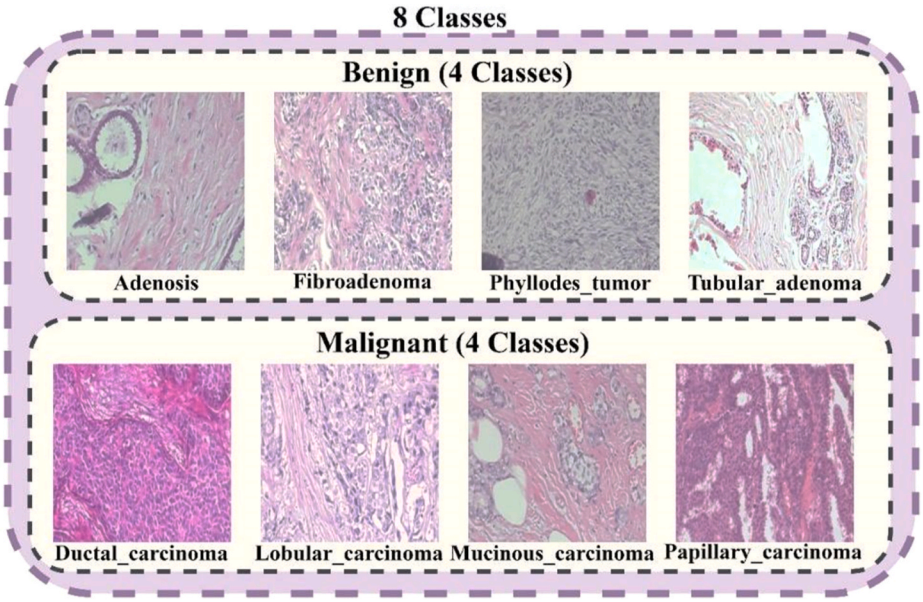

Fig. 5. A BreakHis (Man et al., 2020) pictorial representation. The inner dash lines show the multiclass (4 classes) in two categories: Benign and malignant, whereas the outside dash lines show the multiclass (8 classes).

denoted as SE, Specificity denoted as SP, Precision denoted as PR, F1_Score denoted as F1_S, and area under ROC Curve dented as AUC.

Table 3 summarizes the results of all the experiments carried out in this Paper. The four classes scenario as mentioned earlier is performed in two types on all magnifications. For the Benign class, the benign class yielded generally high performance with accuracy ranging from 0.958 to 0.991, high sensitivity (0.933–0.983) indicating good identification of true benign cases, and high specificity (0.945–0.994) suggesting accurate identification of true negative cases. Precision ranges from 0.933 to 0.990, showing the accuracy of positive predictions. For the Malignant, the performance is good with accuracy ranging from 0.923 to 0.970. Sensitivity ranging from 0.845 to 0.941, indicating the model's ability to identify true malignant cases. Specificity ranges from 0.926 to 0.980, showing accurate identification of true negative cases while precision ranges from 0.845 to 0.941, representing the accuracy of positive predictions.

The eight-class experiment is also recorded in Table 3. The model recorded an accuracy ranging from 0.972 to 0.981 across different magnifications indicating an overall correctness of the model predictions. Sensitivity (SE) varies between 0.886 and 0.923 which measures the ability of the model to correctly identify positive instances (True Positives) out of the total actual positives. The Specificity (SP) shows a range from 0.984 to 0.989 which measures the ability of the model to correctly identify negative instances (True Negatives) out of the total actual negatives. Precision (PR) ranges from 0.886 to 0.923 and represents the accuracy of positive predictions made by the model. F1 Score (F1_S) varies between 0.886 and 0.923. The F1 score is the harmonic mean of precision and sensitivity, providing a balanced measure of a model's performance. Area Under the Curve (AUC) ranges from 0.924 to 0.951 which is a measure of the model's ability to distinguish between classes. A higher AUC suggests better performance.

The model generally performs well across different magnifications and classes. Sensitivity and specificity are balanced, indicating the model's ability to handle both positive and negative cases effectively. The AUC values are consistently high, indicating good discrimination between classes. The results show that the model is robust, but further analysis, such as confusion matrices or class-wise performance, would confirm the robustness and provide more insights into the specific class performance of the proposed model in the subsequent tables and figures.

Fig. 6 illustrates the confusion matrix of the proposed model on the multiclass four-class scenario. For the benign class, the performance of the model at 100x magnification shows that the model correctly classified all the samples except at the Phyllodes_tumor class where it misclassified only one sample as Fibroadenoma. The confusion matrix at 400x Malignant shows that the model classified 76 Ductal_carcinoma correctly while misclassifying 2 samples as Mucinous_carcinoma and Papillary_carcinoma. Classified 11 Lobular_carcinoma samples correctly while misclassifying 2 samples as Ductal_carcinoma. Classified 14 Mucinous_carcinoma samples while misclassifying 2 samples as Ductal_carcinoma and lastly classifying all the Papillary_carcinoma samples correctly. Fig. 7 shows the eight classes' confusion matrix. The results show that the proposed model is robust enough to accurately differentiate the eight classes of the BreakHis samples (see Fig. 8).

4.2. Class performance of the proposed model

Table 4 illustrates the class performance of the proposed model. Based on the Benign - Class Performance, the Ad (Adenosis) class achieves high accuracy across all magnifications (0.983–1.00), excellent AUC values (0.955–1.00) indicating strong discriminatory ability, high precision (1.00) suggesting accurate positive predictions and balanced sensitivity and specificity. For the Fi (Fibroadenoma) class, the model recorded good accuracy (0.933–0.984), and AUC values range from 0.920 to 0.986, indicating good discrimination, and high specificity (1.00) suggesting accurate identification of true negatives. F1 score and sensitivity vary, suggesting potential challenges in certain magnifications. On the other hand, for the Phy (Phyllodes Tumor), the proposed model achieves perfect accuracy (1.00) across magnifications, consistently high AUC values (1.00) indicate excellent discrimination, high precision (1.00) and sensitivity suggesting accurate positive predictions. Meanwhile, for the Tu (Tubular Adenoma) class, the proposed model had an overall good accuracy (0.917–1.00), AUC values range from 0.946 to 1.00, showing good discriminatory ability, precision is lower for some magnifications, indicating potential false positives.

Based on the Malignant - Class Performance, the Dc (Ductal Carcinoma) class achieved good accuracy (0.948–0.992), AUC values range from 0.946 to 0.960, indicating good discriminatory ability and high specificity (0.939–0.992) suggesting accurate identification of true negatives. for the Lc (Lobular Carcinoma) class, the proposed model resulted in an accuracy range of 0.873 to 0.983. The AUC values vary,

Table 3
Performance of the proposed model. AC = accuracy, SE=Sensitivity, SP= specificity, PR=Precision, F1_S=F1_Score, Ad = Adenosis, Fi = Fibroadenoma, Phy = Phyllodes_Tumor, Tu = Tubular_Adenoma, Dc = Ductal_Carcinoma, Lc = Lobular_Carcinoma, Mc = Mucinous_Carcinoma and Pc = Papillary_Carcinoma.

k	40 × Magnification						100 × Magnification						200 × Magnification						400 × Magnification					
	AC	SE	SP	PR	F1_S	AUC	AC	SE	SP	PR	F1_S	AUC	AC	SE	SP	PR	F1_S	AUC	AC	SE	SP	PR	F1_S	AUC
	Eight Classes																							
	0.981	0.923	0.989	0.923	0.923	0.951	0.974	0.897	0.985	0.897	0.897	0.938	0.975	0.917	0.985	0.898	0.898	0.951	0.972	0.886	0.984	0.886	0.886	0.924
Four	Benign																							
Classes	0.958	0.937	0.972	0.934	0.928	0.955	**0.991**	0.983	0.994	0.990	0.984	0.986	0.967	0.933	0.978	0.955	0.933	0.945	0.973	0.946	0.983	0.946	0.946	0.967
	Malignant																							
	0.970	0.941	0.980	0.941	0.941	0.952	0.923	0.845	0.948	0.845	0.845	0.872	0.953	0.905	0.968	0.905	0.905	0.926	**0.975**	0.950	0.983	0.950	0.950	0.947
Four	Benign - Receiver Operating Characteristics (ROC)																							
Classes	Mi-A	Ma-A	Ad	Fi	Phy	Tu	Mi-A	Ma-A	Ad	Fi	Phy	Tu	Mi-A	Ma-A	Ad	Fi	Phy	Tu	Mi-A	Ma-A	Ad	Fi	Phy	Tu
	0.94	0.96	0.95	0.92	1.00	0.95	0.99	0.99	1.00	0.99	0.96	1.00	0.96	0.95	0.95	0.94	0.90	0.99	0.96	0.96	0.94	0.98	0.93	1.00
	Benign - Precision-Recall Curve (AP)																							
	Mi-A	Ad	Fi	Phy	Tu	Mi-A	Ad	Fi	Phy	Tu	Mi-A	Ad	Fi	Phy	Tu	Mi-A	Ad	Fi	Phy	Tu				
	0.86	0.93	0.91	1.00	0.74	0.97	1.00	0.96	0.93	1.00	0.89	0.93	0.88	0.83	0.93	0.91	0.83	0.97	0.78	1.00				
	Malignant - Receiver Operating Characteristics (ROC)																							
	Mi-A	Ma-A	Dc	Lc	Mc	Pc	Mi-A	Ma-A	Dc	Lc	Mc	Pc	Mi-A	Ma-A	Dc	Lc	Mc	Pc	Mi-A	Ma-A	Dc	Lc	Mc	Pc
	0.96	0.95	0.95	0.95	0.95	0.96	0.90	0.87	0.87	0.79	0.93	0.89	0.94	0.93	0.93	0.83	0.96	0.99	0.97	0.95	0.94	0.92	0.93	1.00
	Malignant - Precision-Recall Curve (AP)																							
	Mi-A	Dc	Lc	Mc	Pc	Mi-A	Dc	Lc	Mc	Pc	Mi-A	Dc	Lc	Mc	Pc	Mi-A	Dc	Lc	Mc	Pc				
	0.90	0.95	0.78	0.87	0.87	0.75	0.88	0.43	0.69	0.69	0.84	0.94	0.54	0.82	0.81	0.92	0.94	0.86	0.83	0.93				
Eight	ROC	Mi-A	Ma-A	Ad	Dc	Fi	Lc	Mc	Pc	Phy	Ta	ROC	Mi-A	Ma-A	Ad	Dc	Fi	Lc	Mc	Pc	Phy	Ta		
Classes	40 × Magnification	0.96	0.95	0.95	0.96	0.94	0.89	0.95	0.96	1.00	0.91	100 × Magnification	0.94	0.93	1.00	0.94	0.95	0.91	0.97	0.89	0.88	0.93		
	200 × Magnification	0.94	0.95	1.00	0.94	0.87	0.93	0.94	0.99	0.93	1.00	400 × Magnification	0.94	0.93	0.90	0.92	0.95	0.84	0.87	1.00	0.95	0.99		
	PR/AP	Mi-A	Ad	Dc	Fi	Lc	Mc	Pc	Phy	Ta	PR/AP	Mi-A	Ad	Dc	Fi	Lc	Mc	Pc	Phy	Ta				
	40 × Magnification	0.86	0.83	0.94	0.86	0.70	0.86	0.87	1.00	0.62	100 × Magnification	0.82	1.00	0.92	0.82	0.54	0.81	0.68	0.76	0.87				
	200 × Magnification	0.82	1.00	0.92	0.67	0.78	0.73	0.87	0.55	1.00	400 × Magnification	0.80	0.72	0.87	0.85	0.59	0.62	0.93	0.76	0.86				

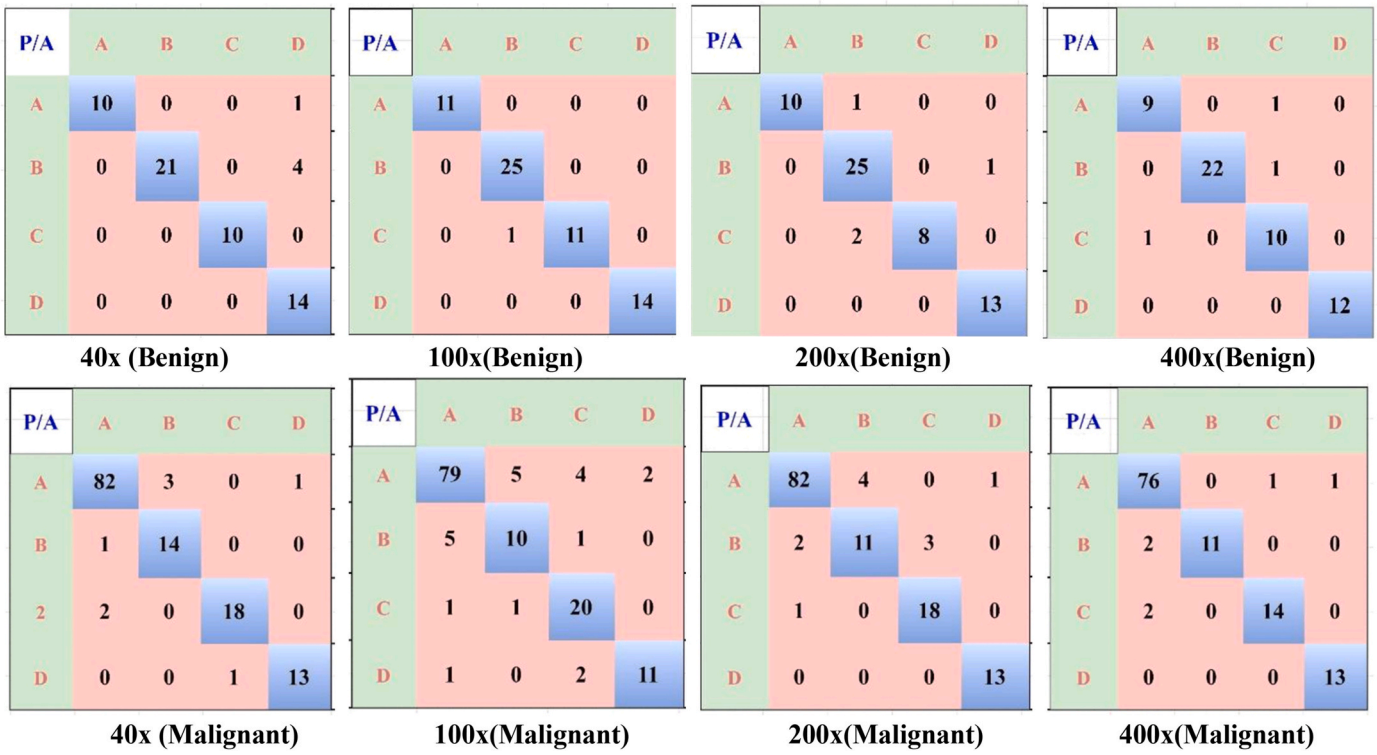

Fig. 6. Confusion matrix of the four classes experiments on all magnifications. P/A denotes predicted and actual. For the benign (Upper row), A = adenosis, B = fibroadenoma, C = Phyllodes_Tumor, d = Tubular_Adenoma whereas for the malignant (down row), A = Ductal_Carcinoma, B = Lobular_Carcinoma, C = Mucinous_Carcinoma, d = Papillary_Carcinoma.

with the lowest at 0.789 and the highest at 0.988. the proposed model may face challenges in correctly identifying lobular carcinoma, especially at lower magnifications. The Mc (Mucinous Carcinoma) class consistently has high accuracy (0.937–0.975). The AUC values range from 0.925 to 0.961, indicating good discriminatory ability as well as good specificity (0.941–0.990) and sensitivity. On the same hand, the Pc (Papillary Carcinoma) class had high accuracy (0.965–0.992), and AUC values ranging from 0.885 to 0.995, suggesting good discriminatory ability. High precision and specificity indicate accurate positive predictions and true negatives.

The Eight Classes - Magnification-wise Performance is also shown in Table 4. The 40 × Magnification had generally high accuracy, AUC, and specificity and an identified potential challenge in sensitivity for some classes. The 100 × Magnification on the other hand had high accuracy and AUC and challenges in sensitivity for certain classes. Meanwhile, the 200 × Magnification had a perfect accuracy for some classes, and AUC and specificity remained high. Finally, the 400 × Magnification had high accuracy, AUC, and specificity, identifying potential challenges in sensitivity for some classes.

Generally, the proposed model demonstrates strong performance across different magnifications. Challenges in sensitivity for certain classes are believed to be magnification-dependent. AUC values generally indicate good discriminatory ability. Precision and specificity are consistently high, suggesting accurate positive predictions and identification of true negatives by the proposed model.

4.3. Proposed model internal composition visualization

Fig. 9 depicts the conceptual composition of the proposed architecture (patch embeddings and the learned convolution filters). recall that each patch embedding number of channels is kept constant as that of the interim feature map. Several patches (Fig. 9(a)) exhibit various patterns. Some are highly distinct, while others are identical to others. Wider context sizes make these renderings more noticeable. We display the underlying convolution kernels to further comprehend the behaviors that a certain kernel is susceptible to in Fig. 9 (b). We observe that the proximity spans of the various filters in the kernel vary, and this trend probably changes concerning the length of training.

4.4. Result discussion

BC is a common and fatal condition that can be successfully treated with early identification and appropriate care. For this goal, several deep learning-based algorithms have been implemented. To get their best performance, these models must be trained on a substantial number of training examples. To help academics conduct their study and give answers, several BC databases have been created. However, these databases feature significant class sample imbalances, which are known to impair the performance of deep learning models. The research community has also proposed some methods for dealing with the data imbalance that forms the basis of our proposed end-to-end hybrid deep learning ensemble model. Although generative models are frequently used for data augmentation, in this study the data sample was generated and balanced using the Python Augmentor module.

Deep learning methods have made it considerably simpler to extract distinguishing BC traits from high-resolution images. Our review of the literature revealed that binary classification is frequently used in studies to categorize BC, where models are created to determine if a breast lump is benign or malignant. Furthermore, according to BreakHis data, researchers who used multi-class classification mainly focused on the four class classifications that correspond to the benign set's Adenosis, Fibroadenoma, Phyllodes Tumor, and Tubular Adenoma, or the malignant set's Ductal Carcinoma, Lobular Carcinoma, Mucinous Carcinoma, and Papillary Carcinoma. Due to its poor classification accuracy and distorted receptive field, the eight-class multi-class categorization of BC is rarely mentioned. The main goal of this work is to use several strategies to resolve this, including ensemble models, CNN models with different receptive fields, feature extraction and fusion, attention

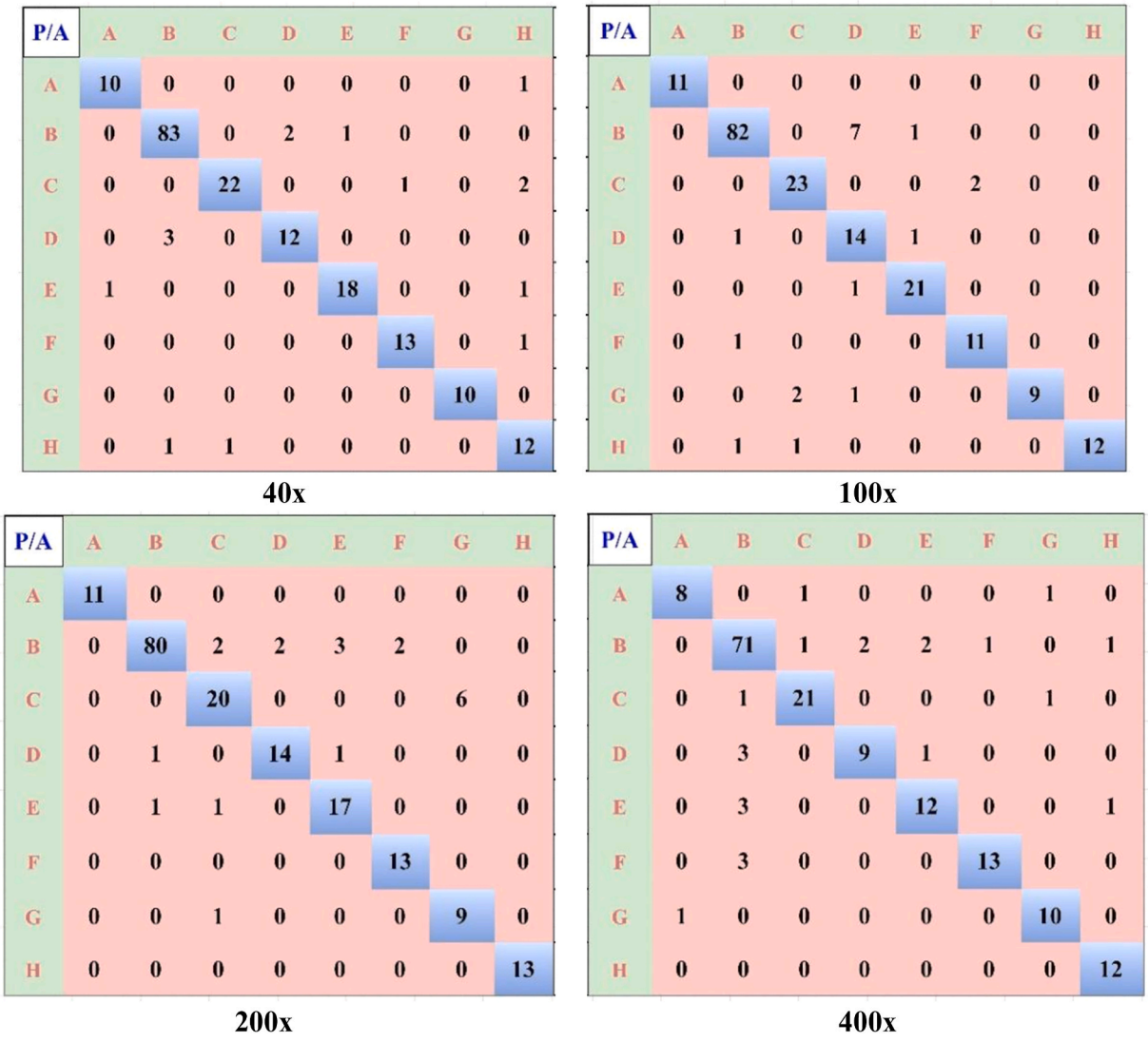

Fig. 7. Confusion matrix of the eight classes experiments on all magnifications. P/A denotes predicted and actual. A = adenosis, B = fibroadenoma, C = Phyllodes_Tumor, d = Tubular_Adenoma, e = Ductal_Carcinoma, F = Lobular_Carcinoma, g = Mucinous_Carcinoma and H = Papillary_Carcinoma.

processes, and more. This study suggests a two, four, and eight-class end-to-end hybrid ensemble model for the categorization of BC.

The effectiveness of the proposed model in the multi-class classification of BC demonstrates that it can be successfully used in the early detection and classification of BC from histopathology images and is a useful tool to aid healthcare professionals in the early and precise diagnosis of BC. We achieved an accuracy of 0.991, 0.983 (Sensitivity), 0.994 (specificity), 0.990 (precision), 0.984 (F1_score) and 0.986 (AUC) which is the model's best performance on the benign multiclass experiment while for the malignant class, it achieved an accuracy of 0.975, sensitivity of 0.950, specificity of 0.983, precision of 0.950, F1_score of 0.950, and AUC of 0.947. The behavior of the proposed model on the malignant class indicates that the proposed model needs lots of data for its optimal performance. The eight class multi-classification results are all promising with 0.97 as minimal performance in all magnifications using the employed evaluation metrics while recording the highest performance at the 40x magnification with an accuracy of 0.981, sensitivity of 0.923, specificity of 0.989, precision of 0.923, F1_score of 0.923 and AUC of 0.951. In general, using the confusion matrix, the performance of the model shows that the most difficult classes to be classified are Phyllodes_tumor, Mucinous_carcinoma, Papillary_carcinoma, and Ductal_carcinoma. However, we believe this is so due to the original number of samples of the mentioned classes.

4.5. Proposed model result validation using K-fold cross-validation

In this study, K-fold cross-validation methods are explored to evaluate the model's performance on unfamiliar data as shown in Table 5. The dataset is partitioned into K subsets in a K-fold fashion, with the standard K value of 10 employed in this study. The analysis, based on performance indicators, revealed the most promising outcome through cross-validation, facilitating the selection of the optimal model from various classifiers. Based on the Benign class, Table 5 indicates that for the 100 × Magnification, the proposed model had high accuracy (0.958 ± 0.016) with a narrow confidence interval, balanced sensitivity and specificity, Precision, F1-Score, and AUC are all high, indicating robust performance. For the 200 × Magnification, the proposed model had excellent accuracy (0.991 ± 0.010) with a narrow confidence interval, high sensitivity, specificity, precision, F1-Score, and AUC, Showing consistent and robust performance across metrics. Meanwhile, in the 300 × Magnification, the proposed model recorded a good accuracy (0.967 ± 0.014) with a narrow confidence interval, balanced sensitivity and specificity, Precision, F1-Score, and AUC at satisfactory levels whereas for the 400 × Magnification, this study recorded high accuracy (0.973 ± 0.157) with a wider confidence interval, indicating some variability, generally balanced sensitivity and specificity, precision, F1-Score, and AUC are reasonably high.

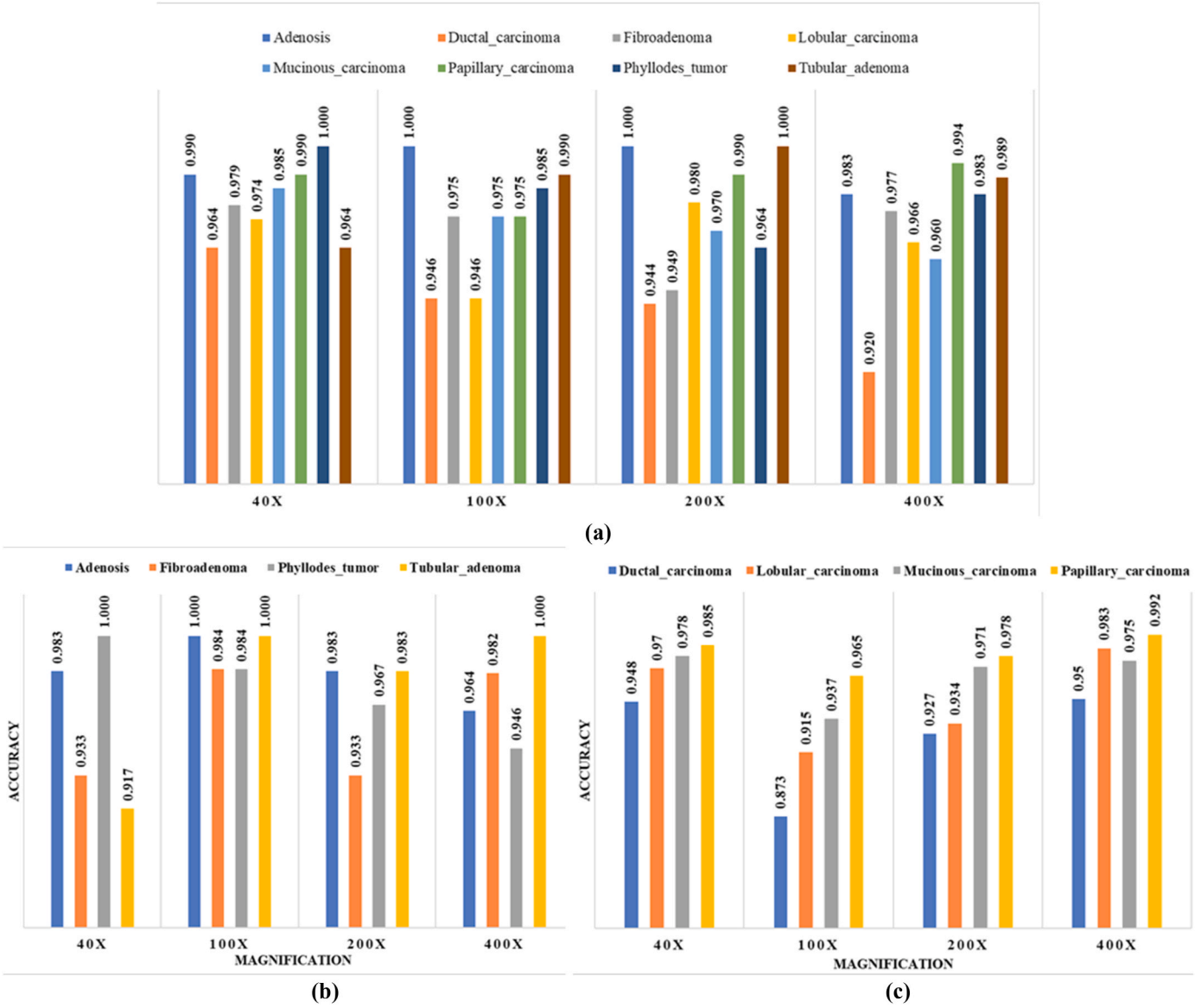

Fig. 8. Illustration of the model class performance (accuracy) in the two-classification scenario (eight classes, four (benign and malignant) classes). (A), represents the eight classes experiment, (B) and (C) represent the four classes experiment via benign and malignant.

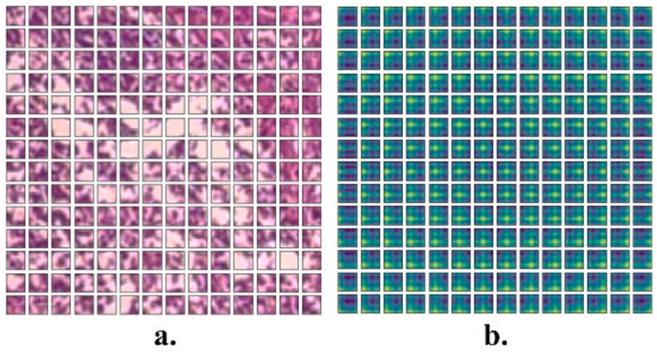

Fig. 9. Visualization of the proposed model internal composition. A. Depicts the generated patch embeddings whereas b. Depicts the learned features.

4.6. Comparison with the state-of-the-art models

For a fair comparison, the comparison of results is based on the networks that used the identical BreakHis dataset (Man et al., 2020), as shown in Table 6. Since the results and methods of newly published articles represent the state-of-the-art, we concentrated more on them. We noticed that the majority of the articles concentrated on a binary classification of Benign and Malignant, basing their decision on the sparse sample size and class imbalance of the BreakHis data (Man et al., 2020). To enhance the number of samples available for their tasks, researchers merged the various magnifications of the Benign and Malignant. Stanitsas et al. (2020) focused on the covariance area when classifying BC images into benign and malignant using a multiple learning instance and Support vector machine (SVM). Both the patient level and the image level were the subjects of their experiments, and the results showed an accuracy of 91.27% for the patient level and 92.00% for the image level. The study of Togacar et al. (Tougaçar et al., 2020) shows how the usage of the attention mechanism has been widely accepted in the medical community. With the CBAM integrated into the ResNet architecture, they were able to produce a promising result at

Table 4
Class performance of the proposed model using accuracy, sensitivity, specificity, precision, F1_Score in all magnifications. Ad = Adenosis, Fi = Fibroadenoma, Phy = Phyllodes_Tumor, Ta = Tubular_Adenoma, Dc = Ductal_Carcinoma, Lc = Lobular_Carcinoma, Mc = Mucinous_Carcinoma and Pc = Papillary_Carcinoma.

Metrics		40 × Magnification				100 × Magnification				200 × Magnification				400 × Magnification			
Four Classes	Benign – Class Performance																
	Metrics	Ad	Fi	Phy	Tu	Ad	Fi	Phy	Tu	Ad	Fi	Phy	Tu	Ad	Fi	Phy	Tu
	Accuracy	0.983	0.933	1.00	0.917	1.00	0.984	0.984	1.00	0.983	0.933	0.967	0.983	0.964	0.982	0.946	1.00
	AUC	0.955	0.920	1.00	0.946	1.00	0.986	0.958	1.00	0.955	0.937	0.900	0.989	0.939	0.978	0.932	1.00
	F1_Score	0.952	0.913	1.00	0.848	1.00	0.980	0.957	1.00	0.952	0.926	0.889	0.963	0.900	0.978	0.870	1.00
	Precision	1.00	1.00	1.00	0.737	1.00	0.961	1.00	1.00	1.00	0.893	1.00	0.929	0.900	1.00	0.833	1.00
	Specificity	1.00	1.00	1.00	0.891	1.00	0.973	1.00	1.00	1.00	0.912	1.00	0.979	0.978	1.00	0.956	1.00
	Sensitivity	0.909	0.84	1.00	1.00	1.00	0.973	0.917	1.00	0.909	0.961	0.800	1.00	0.900	0.957	0.909	1.00
	Malignant - Class Performance																
	Metrics	Dc	Lc	Mc	Pc	Dc	Lc	Mc	Pc	Dc	Lc	Mc	Pc	Dc	Lc	Mc	Pc
	Accuracy	0.948	0.970	0.978	0.985	0.873	0.915	0.937	0.965	0.927	0.934	0.971	0.978	0.950	0.983	0.975	0.992
	AUC	0.946	0.954	0.945	0.960	0.872	0.789	0.925	0.885	0.929	0.827	0.961	0.988	0.940	0.923	0.933	0.995
	F1_Score	0.959	0.875	0.923	0.929	0.898	0.625	0.816	0.815	0.943	0.710	0.900	0.897	0.962	0.917	0.903	0.963
	Precision	0.965	0.824	0.947	0.929	0.918	0.625	0.741	0.846	0.965	0.733	0.857	0.813	0.950	1.00	0.933	0.929
	Specificity	0.939	0.975	0.991	0.992	0.865	0.952	0.941	0.984	0.938	0.967	0.975	0.976	0.905	1.00	0.990	0.990
	Sensitivity	0.953	0.933	0.900	0.929	0.878	0.625	0.909	0.786	0.921	0.688	0.948	1.00	0.974	0.846	0.87	1.00

Eight Classes	Magnification	Metrics	Ad	Dc	Fi	Lc	Mc	Pc	Phy	Ta	Magnification	Ad	Dc	Fi	Lc	Mc	Pc	Phy	Ta
	40 ×	Accuracy	0.990	0.964	0.979	0.974	0.985	0.990	1.00	0.964	100 ×	1.00	0.946	0.975	0.946	0.975	0.975	0.985	0.990
		AUC	0.952	0.964	0.937	0.894	0.947	0.962	1.00	0.915		1.00	0.942	0.952	0.914	0.966	0.888	0.875	0.929
		F1_Score	0.909	0.960	0.917	0.828	0.923	0.929	1.00	0.774		1.00	0.937	0.902	0.718	0.894	0.815	0.857	0.923
		Precision	0.909	0.954	0.957	0.857	0.947	0.929	1.00	0.706		1.00	0.965	0.885	0.609	0.840	0.846	1.00	1.00
		Specificity	0.995	0.963	0.994	0.989	0.994	0.994	1.00	0.972		1.00	0.974	0.983	0.952	0.978	0.989	1.00	1.00
		Sensitivity	0.909	0.965	0.880	0.880	0.900	0.929	1.00	0.857		1.00	0.911	0.920	0.875	0.955	0.786	0.750	0.857
	200 ×	Accuracy	1.00	0.944	0.949	0.980	0.970	0.990	0.964	1.00	400 ×	0.983	0.920	0.977	0.966	0.960	0.994	0.983	0.989
		AUC	1.00	0.940	0.873	0.932	0.936	0.995	0.934	1.00		0.897	0.919	0.950	0.840	0.866	0.997	0.948	0.994
		F1_Score	1.00	0.936	0.800	0.875	0.850	0.929	0.720	1.00		0.842	0.910	0.913	0.750	0.774	0.963	0.870	0.923
		Precision	1.00	0.976	0.833	0.875	0.810	0.867	0.600	1.00		0.888	0.910	0.913	0.818	0.800	0.929	0.833	0.857
		Specificity	1.00	0.981	0.977	0.989	0.978	0.989	0.968	1.00		0.993	0.929	0.987	0.988	0.981	0.994	0.988	0.988
		Sensitivity	1.00	0.899	0.769	0.875	0.895	1.00	0.900	1.00		0.800	0.910	0.913	0.692	0.750	1.00	0.909	1.0

Table 5
The average score and standard deviation for accuracy, sensitivity, specificity, precision, F1-Score, and AUC of the proposed models.

Magnification	Accuracy	Sensitivity	Specificity	Precision	F1-Score	AUC
Benign						
100 ×	0.958 ± 0.016	0.937 ± 0.025	0.972 ± 0.021	0.934 ± 0.023	0.928 ± 0.011	0.955 ± 0.015
200 ×	**0.991 ± 0.010**	0.983 ± 0.008	0.994 ± 0.006	0.990 ± 0.009	0.984 ± 0.011	0.986 ± 0.003
300 ×	0.967 ± 0.014	0.933 ± 0.010	0.978 ± 0.017	0.955 ± 0.018	0.933 ± 0.015	0.945 ± 0.010
400 ×	0.973 ± 0.157	0.946 ± 0.018	0.983 ± 0.022	0.946 ± 0.018	0.946 ± 0.02	0.967 ± 0.017
Malignant						
100 ×	0.970 ± 0.016	0.941 ± 0.025	0.980 ± 0.021	0.941 ± 0.023	0.941 ± 0.011	0.952 ± 0.015
200 ×	0.923 ± 0.011	0.845 ± 0.012	0.948 ± 0.025	0.845 ± 0.250	0.845 ± 0.025	0.872 ± 0.011
300 ×	0.953 ± 0.017	0.905 ± 0.031	0.968 ± 0.028	0.905 ± 0.029	0.905 ± 0.040	0.926 ± 0.021
400 ×	**0.975 ± 0.011**	0.950 ± 0.019	0.983 ± 0.022	0.950 ± 0.008	0.950 ± 0.010	0.947 ± 0.005
Eight Classes						
100 ×	**0.981 ± 0.010**	0.923 ± 0.008	0.989 ± 0.006	0.923 ± 0.009	0.923 ± 0.011	0.951 ± 0.003
200 ×	0.974 ± 0.014	0.897 ± 0.010	0.985 ± 0.017	0.897 ± 0.018	0.897 ± 0.015	0.938 ± 0.010
300 ×	0.975 ± 0.011	0.917 ± 0.019	0.985 ± 0.022	0.898 ± 0.008	0.898 ± 0.010	0.951 ± 0.005
400 ×	0.972 ± 0.009	0.886 ± 0.008	0.984 ± 0.007	0.886 ± 0.006	0.886 ± 0.002	0.924 ± 0.0001

For the 100 × magnification, the proposed model had high accuracy (0.970 ± 0.016) with a narrow confidence interval, balanced sensitivity and specificity, precision, F1-Score, and AUC indicating robust performance whereas, for the 200 × magnification, it recorded good accuracy (0.923 ± 0.011) with a narrow confidence interval, lower sensitivity and specificity, precision, F1-Score, and AUC suggest moderate performance. For the 300 × magnification, the proposed model yielded a high accuracy (0.953 ± 0.017) with a narrow confidence interval, balanced sensitivity, specificity, precision, F1-Score, and AUC at satisfactory levels. For the 400 × magnification, very high accuracy (0.975 ± 0.011) with a narrow confidence interval, balanced sensitivity and specificity, precision, F1-Score, and AUC are consistently high. For the Eight classes in the 100 × magnification, very high accuracy (0.981 ± 0.010) with a narrow confidence interval, balanced sensitivity, specificity, precision, F1-Score, and AUC indicate strong performance. For the 200 × magnification, the proposed model had high accuracy (0.974 ± 0.014) with a narrow confidence interval, lower sensitivity and specificity compared to 100 × magnification. The precision, F1-Score, and AUC suggest satisfactory performance. The 300 × magnification, the proposed model has high accuracy (0.975 ± 0.011) with a narrow confidence interval, balanced sensitivity, specificity, precision, F1-Score, and AUC indicating robust performance. However, for the 400 × Magnification, the model had high accuracy (0.972 ± 0.009) with a narrow confidence interval, balanced sensitivity and specificity, precision, F1-Score, and AUC at satisfactory levels. The model generally performs well across different magnifications for both benign and malignant classes. Confidence intervals, especially for some metrics at 400 × magnification, suggest potential variability in performance.

Table 6
Summary of the state-of-the-art results against the proposed model.

Ref	Approach	Class	Magnification				
			Merge	40x	100x	200x	400x
Stanitsas et al. (Stanitsas et al., 2020)	Region covariance, multiple instance learning, and SVM	Binary	91.27 patient level 92.00 image level	–	–	–	–
Togacar et al. (Tougaçar et al., 2020)	Attention (CBAM) Based ResNet	Binary	98.80	97.99	97.84	98.51	95.88
Chattopadhyay et al. (Chattopadhyay et al., 2022b)	DRDA-Net7 model	Binary	–	96.10	96.03	96.08	96.02
Sharma and Kumar et al. (Sharma and Kumar, 2022)	Xception model and SVM classifier	Binary	–	96.25	96.25	95.74	94.11
M. Liu et al. (Liu et al., 2022b)	AE + Siamese Network	Binary	–	97.3	96.1	97.8	96.7
M. Saini et al. (Saini and Susan, 2022)	VGGIN-Net	Four	–	96.21	97.44	96.22	93.49
	Finetuned VGGIN-Net		–	97.56	96.89	97.49	94.21
S. Sharma et al. (Sharma and Kumar, 2022)	VGG16 + Linear SVM (Patch Based)	Four	–	93.97	92.92	91.23	91.79
	VGG16 + Linear SVM (Patient Based)		–	93.25	91.87	91.5	92.31
M. J. Umer et al. (Umer et al., 2022)	ResNet50 + ESKNN	Eight	80.10	–	–	–	–
	ResNet50 + ESD		69.40	–	–	–	–
	6B-Net + ESKNN		82.03	–	–	–	–
	6B-Net + ESD		70.43	–	–	–	–
	Fused + ESKNN		87.90	–	–	–	–
	Fused + ESD		74.30	–	–	–	–
	F-Selected + ESKNN		90.10	–	–	–	–
	F-Selected + ESD		75.60	–	–	–	–
A. Ameh et al. (Joseph et al., 2022)	HF + DNN (100) augmented	Eight	–	97.24	96.62	95.44	96.29
	HF + DNN (400) augmented		–	97.89	97.6	96.10	96.84
	HF + DNN (400)		–	90.87	89.57	91.58	88.67
	HF + DNN (1000)		–	90.01	88.35	92.95	89.54
S. Sharma et al. (Sharma and Mehra, 2020)	CNN + Pooling Strategy	Eight	–	80.76	76.58	79.90	74.21
S. K. Asare (Asare et al., 2020)	Self-Training and Self-Paced Learning	Eight	–	94.57	94.25	95.1	94.47
S. Boumaraf et al. (Boumaraf et al., 2020)	Transfer Learning (ResNet)	Eight	92.15	94.49	93.27	91.29	89.56
Ours	Proposed	Four	Benign	95.8	99.1	96.7	97.3
	Model		Malignant	97.0	92.3	95.3	97.5
		Eight		98.1	97.4	97.5	97.2

various magnifications, with accuracy rates of 97.99% at 40x, 97.84% at 100x, 98.51% at 200x, and 95.88% at 400x, respectively. The residual dual-shuffle attention network, or DRDA-Net, is a dual-shuffle attention-guided deep learning model that enhances the model's capability to detect complex patterns in images. It was proposed by Chattopadhyay et al. (2022b). The ShuffleNet bottleneck unit served as its

inspiration. For binary classification, Sharma and Kumar et al. (Sharma and Kumar, 2022) utilized the Xception model and SVM classifier, and M. Liu et al. (2022b) used the Autoencoder and Siamese Network. Although the reported findings are encouraging, the one by M. Liu et al. (Huang et al., 2022) showed more dependability with an accuracy of 97.3% for magnifications of 40x, 100x, 200x, and 400x, respectively.

The BreakHis was used by M. Saini et al. (Saini and Susan, 2022) and S. Sharma et al. (Sharma and Kumar, 2022) to classify BC using the VGG16. The VGGIN-Net was specifically proposed by M. Saini et al. (Saini and Susan, 2022) in two scenarios: fine-tuned and non-fine-tuned. The layers of the VGG16 pre-trained model up to the block 4 pool layer were concatenated and frozen to create the VGGIN-Net, which also included a non-linear mapping of basic Inception block units (at the higher level). The classification accuracy was 97.56% for 40x, 96.89% for 100x, 97.49% for 200x, 94.21% for 400x magnification using the finetuned approach, and 96.21% for 40x, 97. In contrast, S. Sharma et al. (Chattopadhyay et al., 2022b) patch-based technique utilized a linear SVM as a classifier and VGG16 as a feature extractor. However, in contrast to M. Saini et al. (Saini and Susan, 2022), their findings were subpar.

The multi-classification of the Eight classes is tackled by few researchers due to the poor accuracy of models. However, eminent writers have developed more accurate models that are more effective. M. J. Umer et al. [58] recommended the CNN model and the ResNet-50 model for deep feature extraction. Following feature extraction, the feature selection methods of PSO and ACS were used to accomplish horizontal feature merging at both of the selected feature vectors. An EBS feature selection technique was utilized before the classification process started. The use of entropy-based choosing features speeds up the training of ML classifiers. M. J. Umer et al. (2022) also suggested a novel 6BNet with six continuous nodes and varied receptive fields. The 6B-intended Net's role of learning high-level characteristics not found in prior studies is carried out using six different receptive field widths. The F-selected + ESKNN strategy obtained the best accuracy of 90.10% out of all the approaches listed in the study of M. J. Umer et al. (2022). The BreakHis dataset's eight classes of multiclassification were tackled by A. Ameh et al. (Joseph et al., 2022) using four distinct techniques across all magnifications. The techniques used are HF + DNN (100) augmented, HF + DNN (400) augmented, and HF + DNN (400). (1000). DNN stands for deep neural network, while HF refers to handmade features. The four techniques all depend on fusing a DNN classifier with manually crafted traits. According to the data, A. Ameh et al. (Joseph et al., 2022) believe that the Haralick texture, Hu moment, and colored histogram with the DNN classifier are more reliable and effective than other approaches. Additionally, data augmentation techniques are typically used to improve the classification's accuracy by correctly changing the variables. The accuracy of the HF + DNN (400) enhanced strategy was the greatest of the four, with values of 97.89% for magnifications of 40x, 100x, 200x, and 400x. While S. K. Asare (Asare et al., 2020) suggested the use of Self-Training and Self-Paced Learning, S. Boumaraf et al. (2020) suggested the use of Transfer learning (ResNet), which is based on a block-wise fine-tuning strategy; in which they made the last two residual blocks of the model more domain-specific to the target data for the Eight classes multiclassification of the BreakHis data. S. Sharma et al. (Sharma and Mehra, 2020) study used images at different magnification levels to teach CNN how to recognize distinctive features. The overfitting problem is also solved using methods for improving the data. S. K. Asare's (Asare et al., 2020) contribution is based on a semi-supervised learning strategy that produces and selects pseudo-labeled samples while using self-training and identity learning for classifying BC histopathology images. To create and choose self-assured pseudo-labeled samples from both classes with high and low representation, the learning approach offers a special pseudo-label generation and selection mechanism. A class-balancing design that trivializes the class-wise optimism ratings is also recommended to prevent the network from overlooking data from underrepresented classes (difficult samples). This essentially solves the data imbalance issue. In conclusion, the results of all trials using our suggested model showed that it outperformed all previously described models (Multi-classification).

4.7. Processing time comparison

This study assessed the inference times of the proposed model based on its design components (Conventional ConvMixer, Patch-based, and 5-BNet), which are essential for real-world applications where prompt decision-making is necessary, in order to assess the computational complexity of the proposed model. Table 7 displays each component's operating time analysis for a clear comparison picture. With a reasonable inference time of 18 s and 157 milliseconds/step, the Conventional ConvMixer performs well. This performance suggests a method that strikes a compromise between speed and complexity, making it appropriate for situations where a certain amount of delay is tolerable but extensive analysis is necessary. The inference time of the Patch-based model with 5B-Network is 15 s and 141 milliseconds/step. The inference time decreases little as the number of components is reduced, indicating that the 5-BNet adds very little to the computational effort. The fastest inference time of 12 s and 142 milliseconds/step is observed when the 5-BNet is eliminated. This finding implies that the 5-B Network affects computational complexity as well and that eliminating it expedites the inference process. Its inclusion does, however, improve the model's functionality.

A detailed comparison of several models on the 8-class multi-classification is given in Table 8, along with each model's corresponding inference time expressed in seconds (s). It's interesting to see that the suggested models all show much shorter inference times; 18, 15, and 12 s, respectively. The suggested model's simplifications have resulted in quicker processing, which makes them appealing options for situations where speed is essential. It's crucial to remember that these simplifications might result in decreased performance and accuracy. As a result, choosing a model should take into account the particular needs of the application while balancing processing speed and overall inference time.

5. Conclusion and future works

Early identification and treatment of breast cancer are key to its management. Thanks to current technological advancements. In the world of medicine, deep learning models have been demonstrated to be promising. In this study, we proposed a patch-based model for multi-classifying breast cancer that dynamically blends the spatial and channel positions of patch embeddings. Patch embeddings enable simultaneous internal and external downsampling, resulting in a larger effective receptive field size and a simpler mixing of far-off spatial information. This study further introduced a novel 5-B Network at the end of the pointwise convolutional layer for tiny feature extraction. The 5-BNet comprises five branches that process information concurrently. The primary distinction among these branches lies in the size of their convolutional kernels. In order to capture high-level image features, the 5-BNet gradually reduces the filter size of each convolution layer in every concurrent branch. The BreaKHis public dataset, which includes 7909 images from 82 different individual reports is used for evaluation. The dataset is divided into the following three sections: 70:20:10 (Train,

Table 7
Running Time Analysis of Conventional ConvMixer, Patch-based, and 5-BNet models.

Model	Inference Time (s)	Time per Step (ms/steps)
Conventional ConvMixer	18	157
A patch-based model with 5B-Network	15	141
Removed 5B-Network	12	142

Table 8
Comparison of running time between the proposed model and existing models.

Model	Inference Time (s)
S. K. Asare (Asare et al., 2020)	45
A. Ameh et al. (Joseph et al., 2022)	20
S. Sharma et al. (Sharma and Mehra, 2020)	39
M. J. Umer et al. (Umer et al., 2022)	23
S. Boumaraf et al. (Boumaraf et al., 2020)	18
Ours	15

Validate, Test). Under various image magnification factors and the data augmentation technique, several matrices, including accuracy, F1 score, AUC, precision, sensitivity, and specificity, were taken into consideration. The proposed model outperforms state-of-the-art breast cancer models with an accuracy of 99.1% for benign multiclassification, 97.1% for malignant multiclassification, and 98.1% for eight-class classifications, even though neither our model nor our experiments were designed to maximize accuracy or speed, i.e., we did not search for good hyperparameters. We demonstrated that the extremely prevalent "homogeneous" model with a straightforward patch embedding stem is a potent deep learning framework.

Broader Impact on Breast Cancer Diagnostics: The findings of this study have significant implications for breast cancer diagnostics. Achieving high classification accuracy across multiple classes (benign and malignant subtypes) enhances the reliability of automated diagnostic tools, potentially reducing human error and providing consistent, objective assessments. The model's ability to process multiple magnification levels and identify subtle patterns across spatial scales could be particularly beneficial in early-stage detection, where distinguishing between benign and malignant tissues is critical. Furthermore, with its capability to handle varying magnifications and capture intricate spatial details, the model could be adapted to support pathologists in identifying challenging cases, improving diagnostic confidence and reducing time spent on routine evaluations. This approach aligns with the broader trend of incorporating AI-assisted diagnostic tools in medical imaging to enhance accuracy and efficiency.

Future Work: While the model demonstrates promising results, future enhancements could further refine its effectiveness. Utilizing larger patches, extending training periods, and applying additional normalization and hyperparameter tuning may enable the model to strike an optimal balance between accuracy, performance, and efficiency. Minor design modifications, such as incorporating bottlenecks or more robust filters, might trade off some complexity for increased speed. Additionally, refining the depthwise convolution layers could significantly boost performance. Expanding the model to incorporate patient metadata, like age or family history, would further enhance diagnostic accuracy, creating a more holistic tool for personalized breast cancer care.

In summary, this study advances the application of patch-based deep learning models in breast cancer diagnostics, providing a framework that is accurate, scalable, and potentially transformative for clinical use. The 5-B Network's multi-branch structure, coupled with effective patch embeddings, offers a valuable contribution to AI-based medical imaging with the potential to support and improve clinical workflows in pathology labs worldwide.

CRediT authorship contribution statement

Jehoiada Jackson: Writing – review & editing, Writing – original draft, Data curation, Conceptualization. **Linda E. Jackson:** Writing – original draft, Formal analysis. **Chiagoziem C. Ukwuoma:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Maame D. Kissi:** Validation. **Ariyo Oluwasanmi:** Supervision. **Qin Zhiguang:** Supervision, Funding acquisition.

Informed consent

Informed consent was obtained from all participants included in the study.

Code availability/availability of data

https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/(retrieved: November 12, 2022) provides access to the dataset utilized in this study. The TensorFlow/Keras code we utilized in our experiment is not currently accessible to the general public, but it will be when the study is published.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research is supported by the National Science Foundation of China (NSFC) under the project "Development of fetal heart-oriented heart sound echocardiography multimodal auxiliary diagnostic equipment" (62027827). We also acknowledge the Network and Data Security Key Laboratory of Sichuan for providing us with a good environment for the study.

Data availability

The data used a an open access data and we provided the download link.

References

Abdel-Zaher, A.M., Eldeib, A.M., 2016. Breast cancer classification using deep belief networks. Expert Syst. Appl. 46, 139–144. https://doi.org/10.1016/j.eswa.2015.10.015.

Abu Al-Haija, Q., Manasra, G., 2020. Development of breast cancer detection model using transfer learning of residual neural network (ResNet-50). Am. J. Sci. Eng. Technol. 1, 30–39. https://doi.org/10.15864/ajse.1304.

Ahmad, A., 2019. Breast cancer statistics: recent trends. Breast cancer metastasis drug Resist. challenges Prog. 1–7.

Ahmad, N., Asghar, S., Gilani, S., 2022. Transfer learning-assisted multi-resolution breast cancer histopathological images classification. Vis. Comput. 38. https://doi.org/10.1007/s00371-021-02153-y.

Arhatari, B.D., et al., 2021. X-ray phase-contrast computed tomography for full breast mastectomy imaging at the Australian Synchrotron. In: Developments in X-Ray Tomography XIII, vol. 11840, p. 131.

Asare, S.K., You, F., Nartey, O.T., 2020. A semisupervised learning scheme with self-paced learning for classifying breast cancer histopathological images. Comput. Intell. Neurosci. 2020. https://doi.org/10.1155/2020/8826568.

Boumaraf, S., Liu, X., Ferkous, Z., 2020. A new transfer learning based approach to magnification dependent and independent classification of breast cancer in histopathological images. Biomed. Signal Process Control 63, 1746–8094. https://doi.org/10.1016/j.bspc.2020.102192.

Brancati, N., De Pietro, G., Riccio, D., Frucci, M., 2021. Gigapixel histopathological image analysis using attention-based neural networks. IEEE Access 9, 87552–87562. https://doi.org/10.1109/access.2021.3086892.

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A., 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J. Clin. 68 (6), 394–424.

Chattopadhyay, S., Dey, A., Singh, P.K., Oliva, D., Cuevas, E., Sarkar, R., 2022a. MTRRE-Net: a deep learning model for detection of breast cancer from histopathological images. Comput. Biol. Med. 150, 106155.

Chattopadhyay, S., Dey, A., Singh, P., Sarkar, R., 2022b. DRDA-Net: dense residual dual-shuffle attention network for breast cancer classification using histopathological images. Comput. Biol. Med. 145, 105437. https://doi.org/10.1016/j.compbiomed.2022.105437.

Choudhary, T., Mishra, V., Goswami, A., Sarangapani, J., 2021. A transfer learning with structured filter pruning approach for improved breast cancer classification on point-of-care devices. Comput. Biol. Med. 134, 104432.
Dosovitskiy, A., et al., 2021. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale.
El Agouri, H., et al., 2022. Assessment of deep learning algorithms to predict histopathological diagnosis of breast cancer: first Moroccan prospective study on a private dataset. BMC Res. Notes 15 (1), 1–7.
Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K., Li, S., 2017. Breast cancer multi-classification from histopathological images with structured deep learning model. Sci. Rep. 7 (1). https://doi.org/10.1038/s41598-017-04075-z.
Honkanen, N., Mustonen, L., Kalso, E., Meretoja, T., Harno, H., 2021. Breast reconstruction after breast cancer surgery–persistent pain and quality of life 1–8 years after breast reconstruction. Scand. J. pain 21 (3), 522–529.
Houssein, E.H., Emam, M.M., Ali, A.A., Suganthan, P.N., 2021. Deep and machine learning techniques for medical imaging-based breast cancer: a comprehensive review. Expert Syst. Appl. 167, 114161. https://doi.org/10.1016/j.eswa.2020.114161.
Huang, L., Chen, T., Sun, X., 2022. Cross-attention vision transformer for capturing inter-patch relationships in breast cancer imaging. Comput. Med. Imag. Graph. 98, 101922. https://doi.org/10.1016/j.compmedimag.2022.101922.
Ibrokhimov, B., Hur, C., Kim, H., Kang, S., 2020. An optimized deep belief network model for accurate breast Cancer classification. IEIE Trans. Smart Process. Comput. 9 (4), 266–273.
Iqbal, M.S., Ahmad, W., Alizadehsani, R., Hussain, S., Rehman, R., 2022. Breast cancer dataset, classification and detection using deep learning. Healthcare 10 (12), 2395.
Jayandhi, S.M.J.G., Leena Jasmine, J.S., 2022. Mammogram learning system for breast cancer diagnosis using deep learning SVM. Comput. Syst. Sci. Eng. 40 (2), 491–503. https://doi.org/10.32604/csse.2022.016376.
Joseph, A.A., Abdullahi, M., Junaidu, S.B., Ibrahim, H.H., Chiroma, H., 2022. Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer). Intell. Syst. with Appl. 14, 200066. https://doi.org/10.1016/j.iswa.2022.200066.
Kaplan, E., et al., 2022. Automated BI-RADS classification of lesions using pyramid triple deep feature generator technique on breast ultrasound images. Med. Eng. Phys. 108, 103895.
Karthik, R., Menaka, R., Kathiresan, G.S., Anirudh, M., Nagharjun, M., 2022a. Gaussian dropout based stacked ensemble CNN for classification of breast tumor in ultrasound images. Irbm 43 (6), 715–733.
Karthik, R., Menaka, R., V Siddharth, M., 2022b. Classification of breast cancer from histopathology images using an ensemble of deep multiscale networks. Biocybern. Biomed. Eng. 42 (3), 963–976.
Liu, N., Qi, E.-S., Xu, M., Gao, B., Liu, G.-Q., 2019. A novel intelligent classification model for breast cancer diagnosis. Inf. Process. {&}amp$\mathsemicolon$ Manag. 56 (3), 609–623. https://doi.org/10.1016/j.ipm.2018.10.014.
Liu, H., et al., 2022a. Artificial intelligence-based breast cancer diagnosis using ultrasound images and grid-based deep feature generator. Int. J. Gen. Med. 2271–2282.
Liu, M., He, Y., Wu, M., Zeng, C., 2022b. Breast histopathological image classification method based on autoencoder and siamese framework. Information 13 (3), 107.
Mahmoud, H.A.H., Alharbi, A.H., Khafga, D.S., 2021. Breast cancer classification using deep convolution neural network with transfer learning. Intell. Autom. Soft Comput. 29 (3).
Man, R., Yang, P., Xu, B., 2020. Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks. IEEE access 8, 155362–155377.
Mansha, M., Yaqub, S., Latif, A.A., Wasim, M., Thomson, P.C., 2022. HER2/neu levels, premature termination of pregnancy and breast cancer risk: a hospital-based study from lahore, Pakistan. Pak. J. Zool. 54 (1).
Mathew, T., Kini, J.R., Rajan, J., 2021. Computational methods for automated mitosis detection in histopathology images: a review. Biocybern. Biomed. Eng. 41 (1), 64–82. https://doi.org/10.1016/J.BBE.2020.11.005.
Naik, N., et al., 2020. Deep learning-enabled breast cancer hormonal receptor status determination from base-level H&E stains. Nat. Commun. 11 (1), 5727.
Nassif, A., Abu Talib, M., Nasir, Q., Afadar, Y., Elgendy, O., 2022. Breast cancer detection using artificial intelligence techniques: a systematic literature review. Artif. Intell. Med. 127, 102276. https://doi.org/10.1016/j.artmed.2022.102276.
Patel, A., Gupta, R., Kumar, S., 2022. Self-supervised transformer model for improving breast cancer classification accuracy on unlabeled histopathology images. IEEE Access 10, 8765–8775. https://doi.org/10.1109/ACCESS.2022.3146875.
Rahman, M., Khan, A., 2022. Adaptive patch-based transformer model with feature aggregation for enhanced breast cancer detection. Expert Syst. Appl. 120, 114789. https://doi.org/10.1016/j.eswa.2022.114789.
Saini, M., Susan, S., 2022. VGGIN-net: deep transfer network for imbalanced breast cancer dataset. IEEE/ACM Trans. Comput. Biol. Bioinforma. https://doi.org/10.1109/TCBB.2022.3163277. PP.
Shankar, K., Dutta, A.K., Kumar, S., Joshi, G.P., Doo, I.C., 2022. Chaotic sparrow search algorithm with deep transfer learning enabled breast cancer classification on histopathological images. Cancers (Basel) 14 (11), 2770.
Sharma, S., Kumar, S., 2022. The Xception model: a potential feature extractor in breast cancer histology images classification. ICT Express 8 (1), 101–108.
Sharma, S., Mehra, R., 2020. Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology Images{\textemdash}a comparative insight. J. Digit. Imaging 33 (3), 632–654. https://doi.org/10.1007/s10278-019-00307-y.
Soumik, M.F.I., Hossain, D.M.A., Bin Aziz, A.Z., 2021. Improved Transfer Learning Based Deep Learning Model for Breast Cancer Histopathological Image Classification, pp. 1–4. https://doi.org/10.1109/ACMI53878.2021.9528263.
Stanitsas, P., Cherian, A., Morellas, V., Tejpaul, R., Papanikolopoulos, N., Truskinovsky, A., 2020. Image descriptors for weakly annotated histopathological breast cancer data. Front. Digit. Heal. 2. https://doi.org/10.3389/fdgth.2020.572671.
Thomas, A.M., Adithya, G., Arunselvan, A.S., Karthik, R., 2022. Detection of breast cancer from histopathological images using image processing and deep-learning. In: 2022 Third International Conference on Intelligent Computing Instrumentation and Control Technologies (ICICICT), pp. 1008–1015.
To\ugaçar, M., Özkurt, K.B., Ergen, B., Cömert, Z., 2020. {BreastNet}: a novel convolutional neural network model through histopathological images for the diagnosis of breast cancer. Phys. A Stat. Mech. its Appl. 545, 123592. https://doi.org/10.1016/j.physa.2019.123592.
Tsapatsaris, A., Babagbemi, K., Reichman, M., 2021. Barriers to breast cancer screening are worsened amidst COVID-19 pandemic: a review. Clin. Imaging 82. https://doi.org/10.1016/j.clinimag.2021.11.025.
Ukwuoma, C.C., et al., 2022a. Automated lung-related pneumonia and COVID-19 detection based on novel feature extraction framework and vision transformer approaches using chest X-ray images. Bioengineering 9 (11), 709.
Ukwuoma, C.C., et al., 2022b. Boosting breast cancer classification from microscopic images using attention mechanism. In: 2022 International Conference on Decision Aid Sciences and Applications (DASA), pp. 258–264.
Umer, M., Sharif, M., Kadry, S., Alharbi, A., 2022. Multi-class classification of breast cancer using 6B-Net with deep feature fusion and selection method. J. Pers. Med. 12. https://doi.org/10.3390/jpm12050683.
Vaswani, A., et al., 2023. Attention Is All You Need.
Wang, Y., Zhou, Q., 2022. Hybrid deep learning approach combining CNNs and transformers for multi-scale feature extraction in breast cancer classification. IEEE Trans. Med. Imaging 41 (5), 1234–1242. https://doi.org/10.1109/TMI.2022.3156789.
Wang, K., Franch-Expósito, S., Li, L., Xiang, T., Wu, J., Ren, G., 2020a. 34P Comprehensive clinical and molecular portraits of grade 3 ER+ HER-breast cancer. Ann. Oncol. 31, S27.
Wang, Y., et al., 2020b. Breast cancer image classification via multi-network features and dual-network orthogonal low-rank learning. IEEE Access (1). https://doi.org/10.1109/ACCESS.2020.2964276. PP.
Xu, J., Li, Z., 2022. Patch-based vision transformer model for histopathological image analysis in breast cancer classification. Med. Image Anal. 74, 102346. https://doi.org/10.1016/j.media.2022.102346.
Yang, H., Liu, Z., Fang, M., 2022. Multi-modal transformer-based model integrating image and metadata for breast cancer diagnosis. J. Digit. Imaging 35 (2), 245–254. https://doi.org/10.1007/s10278-022-00511-6.