scientific reports

Check for updates

OPEN

Fusing global context with multiscale context for enhanced breast cancer classification

Niful Islam[1], Khan Md Hasib[2], M. F. Mridha[3]✉, Sultan Alfarhood[4], Mejdl Safran[4]✉ & M. K. Bhuyan[5]

Breast cancer is the second most common type of cancer among women. Prompt detection of breast cancer can impede its advancement to more advanced phases, thereby elevating the probability of favorable treatment consequences. Histopathological images are commonly used for breast cancer classification due to their detailed cellular information. Existing diagnostic approaches rely on Convolutional Neural Networks (CNNs) which are limited to local context resulting in a lower classification accuracy. Therefore, we present a fusion model composed of a Vision Transformer (ViT) and custom Atrous Spatial Pyramid Pooling (ASPP) network with an attention mechanism for effectively classifying breast cancer from histopathological images. ViT enables the model to attain global features, while the ASPP network accommodates multiscale features. Fusing the features derived from the models resulted in a robust breast cancer classifier. With the help of five-stage image preprocessing technique, the proposed model achieved 100% accuracy in classifying breast cancer on the BreakHis dataset at 100X and 400X magnification factors. On 40X and 200X magnifications, the model achieved 99.25% and 98.26% classification accuracy respectively. With a commendable classification efficacy on histopathological images, the model can be considered a dependable option for proficient breast cancer classification.

Keywords Breast Cancer, Classification, CNN, Vision Transformer, Transfer learning

Breast cancer is one of the main causes of death in women between the ages of 20 to 59[1]. At almost 12% of all cancer cases, it is one of the most common malignancies among women and is expected to increase by 3.1% each year[2,3]. Ninety percent of patients receiving an early diagnosis of breast cancer survive five years longer[4]. Traditional diagnostic approaches are not only time-consuming but also prone to errors, emphasizing the importance of a computer vision-based diagnostic approach characterized by rapid and consistent classification performance. There are several approaches to breast cancer image acquisition including ultrasound, CT scan, MRI, histopathology, mammography, and thermography. Among them, histopathology is the most reliable one since it provides tissue-level information[5]. Precise categorization of breast cancer from histopathological images can greatly influence the prognosis of patients and assist doctors in creating rational treatment plans.

Traditional image classification techniques predominantly rely on Convolutional Neural Networks (CNNs), a subset of Artificial Neural Networks (ANNs)[6]. CNNs are engineered to excel in tasks such as image classification and object recognition. Convolutional layers, composed of learnable filters or kernels, are the fundamental building blocks of this architecture. These filters efficiently identify and encode complex patterns and features at various hierarchical levels and sizes by convolving over the input images[7,8]. These networks have revolutionized various fields, particularly in medical imaging, with their remarkable performance. Due to their high accuracy, they are commonly used for disease classification, segmentation, and object detection. Their ability to process large volumes of images at a very convenient time makes CNN a useful tool in medical image processing[9].

Over the past few years, CNNs have been extensively used in processing histopathological images[10]. However, due to their limited ability in accommodating multiscale and global contexts, they often lead to lower classification accuracy. CNNs have fixed-size convolutional filters with limited receptive fields, which hinders their ability to grasp context at different scales[11]. Moreover, CNNs frequently show sensitivity to changes in

[1]Department of Computer Science and Engineering, United International University, Dhaka 1212, Bangladesh. [2]Department of Computer Science and Engineering, Bangladesh University of Business and Technology, Mirpur, Dhaka 1216, Bangladesh. [3]Department of Computer Science, American International University - Bangladesh, Dhaka 1229, Bangladesh. [4]Department of Computer Science, College of Computer and Information Sciences, King Saud University, P.O.Box 51178, Riyadh 11543, Saudi Arabia. [5]Department of Electronics and Electrical Engineering, Indian Institute of Technology Guwahati, Assam 781039, India. ✉email: firoz.mridha@aiub.edu; mejdl@ksu.edu.sa

the resolution of the input images, which might impact performance when handling scale variations. Since medical images exhibit variations in resolution and texture due to differences in capturing devices, CNNs fail to generalize in this situation.

Atrous Spatial Pyramid Pooling, or ASPP for short, appears as a tactical way to deal with the limitation of multiscale context in CNNs. By adding numerous dilation rates to dilated convolutions, also known as atrous convolutions, ASPP greatly expands the network's spatial sense without adding superfluous parameters[12,13]. While the ASPP module addresses the inability to capture multiscale context, accommodating global context has remained a challenge[14]. This inability becomes more visible when processing histopathological images since analyzing tissue patterns and structures at various scales is necessary for accurate classification[15]. Vision Transformers (ViT)[16] emerged as a solution to this problem by integrating self-attention mechanism. The core concept of ViT is to treat images as a series of patches, which is a fundamentally different approach from that of CNN. This new approach eliminates the requirement for handcrafted spatial hierarchies and hierarchical features, allowing ViT to successfully capture both local and global features using self-attention.

In this paper, we have presented a fusion for effectively classifying histopathological breast cancer images. The proposed model is composed of two streams, namely the global stream and the multiscale stream. The global stream is composed of ViT that captures the global context through self-attention mechanism. On the other hand, the multiscale stream accommodates a custom ASPP model, composed of atrous convolutions and attention mechanism for capturing multiscale contextual information. The features extracted from the stream are fused to construct the final classifier. Moreover, the proposed system also incorporates a five stage image preprocessing technique for enhancing the image quality yielding a high classification accuracy. We have tested the model on the BreakHis dataset and the performance indicates superior results over existing methods. In summary, this article presents the following major contributions.

- This research article presents a novel architecture composed of ViT and a custom ASPP network for effectively classifying breast cancer images. While ViT allows the model to capture global features; the custom ASPP network provides multiscale information. Conjoining these two enables the model to distinguish between benign and malignant patterns at varying magnification levels by capturing both fine-grained information and the general context of the tissue. The proposed system integrates a series of image preprocessing techniques resulting in enhanced image quality.
- A comparison with the existing studies is made which shows the superior performance of the model.The rest of the article is structured as follows. Section "Related Works" presents a brief overview of the existing methods for classifying breast cancer images. Section "Proposed method" holds the description of the proposed method followed by the results on Section "Results". Finally, the article terminates in Section "Conclusion".

Related Works

Classifying breast cancer, especially from histopathological images, has always been a challenging task. Researchers have presented various methods to address this issue. According to Zhuo et al[17]., CNN-based transfer learning models are most commonly used for breast cancer classification. Khan et al[18]. presented a fusion model that is composed of two streams. Each stream was composed of three CNN architectures (ResNet, DenseNet, and EffecientNet) which resulted in a total of six feature extractors. Although the model achieved noteworthy performance, it is extremely resource-consuming due to its architecture. Another resource-hungry architecture was presented by Ijaz et al[19]. where the authors altered the VGG architectures (VGG16 and VGG19) by integrating Convolutional Block Attention Module (CBAM) module and presented a fusion model. Features extracted from each block of the VGG networks are passed through a global average pooling and a CBAM block which are fused later on. The outputs of the two modified VGG networks are merged and passed to a feed forward layer for final classification. While the VGG networks were known for their heavy architectures, the research also did not integrate transfer learning, which made the model extremely difficult to train. Khan et al[20]. incorporated three pretrained feature extractors named DenseNet201, NasNetMobile, and VGG16 for constricting a fusion model. The features from the three feature extractors were concatenated and forwarded to a shallow multi-layer perceptron (MLP) head for classification. This model achieved an outstanding accuracy of 99% in differentiating Benign and Malignant types of breast cancers. Wakili et al[21]. developed a classifier named DenTnet, made of DenseNet and transfer learning, which achieves an excellent classification accuracy of 99.28% on the BreakHis dataset. Although the proposed solution is evaluated on four benchmark datasets, the lack of interpretability of the model's decision-making process raises a concern about the model's generalization ability.

Since pretrained models are generally trained on ImageNet which differs from biomedical images, fine-tuning the top layers of the state-of-the-art feature extractors make a significant difference in the classification performance. Ashurov et al[22]. experimented with four pretrained CNN models named Xception, VGG16, ResNet50, MobileNet, and DenseNet121 for classifying breast cancers. They also integrated CBAM after the pretrained feature extractor to enhance the performance. According to the experiment, Xception with CBAM outperformed the three classifiers. While squeeze and excitation blocks enhance feature maps adaptively with channel attention, dual squeeze and excitation (DSE) blocks allow the model to focus on both channel and spatial features[23]. The article, however, lacks a comparative analysis of the existing solutions to prove the model's superiority over the existing ones. Sarker et al[24]. integrated DSE networks in EfficientNetV2 architecture for classifying breast cancers. The incorporation of DSE largely improves the model's performance on 100X magnification dataset. While on the other magnifications, the module improved the performance by a small margin. Additionally, the study lacks an ablation study, leaving a gap in the contribution of individual components of the model. Abbasniya et al[25]. leveraged InceptionResNetV2 to enhance the classification performance. The features from InceptionResNetV2 were then passed to an ensemble of classifiers made of three boosting algorithms namely CatBoost, XGBoost, and LightGBM. The algorithm returns the class label

with the highest probability by calculating the soft voting of the classifiers. Although the model achieves an acceptable performance on the breakhis dataset, a comparative analysis suggests the model fails to outperform the existing solutions on 100X magnification. Joshi et al[26]. integrated three convolution and pooling layers after the base Xception backbone for accelerating the classifier's performance. Although the model achieves an acceptable accuracy of 93.33% in classifying two types of cancer cells, no comparison is presented indicating the improvement in performance with the fine-tuning. While CNNs produced a noteworthy performance in many research works, according to Kode and Barkana[27], CNNs resulted in overfitting in the histopathological breast cancer images. They experimented with a custom CNN model with 3 convolution blocks and a pertrained VGG16 architecture. The research showed that the custom shallow CNN model tended to produce more biased results than VGG16, even with 20% dropout. Therefore, they presented a knowledge-based feature extraction module that extracted geometrical, directional, and intensity-based features leveraging mathematical equations. The features extracted from this module are then passed through seven classifiers for final classification. Among the seven classifiers, all the classifiers except the decision tree produced an acceptable accuracy on the BreakHis dataset. Nevertheless, the article only considered images at 400X magnification, leaving the efficiency of the model unknown for other magnifications.

Feature extraction with deep learning techniques (typically CNN) and classification with machine learning algorithms is an established method that achieves very high classification performance. The deep learning methods allow the model to extract complex features while the machine learning model makes the model more interpretable and robust[28]. Sharmin et al[29]. employed ResNet50V2 for feature extraction and Light Boosting Classifier (LGB) for classification. The research, nevertheless, was conducted only on one dataset with 162 samples which may lead to a misleading conclusion. Liew et al[30]. employed DenseNet201 as a feature extractor and eXtreme Gradient Boosting (XGBoost) as the classifier for differentiating various types of breast cancers. The model achieved an accuracy of 97% in both binary class and multi-class classification. The study makes no differentiation of the histopathological images based on magnification levels. Jasti et al[31]. employed AlexNet to extract useful features which are passed to relief algorithm. This feature selection algorithm retained the most important features while discarding the less important ones. The filtered features were finally passed to four machine learning classifiers for inference. Among the four machine learning classifiers (SVM, KNN, random forest, and Naïve Bayes), SVM performed the best. Shen et al[32]. employed ResNet18, a lightweight feature extractor that efficiently draws out useful features from X-ray images. The features are then refined with DSLPSO algorithm, a feature selection method for the reduction of redundant and unnecessary features. Finally, the images are classified with SVM.

Recently transformer based approaches have gained attention in classifying various medical images. Originating from natural language processing, this self-attention based architecture is currently dominating the field of computer vision. Ayana et al[33]. investigated three transformer architectures namely vision transformer, swin transformer, and pyramid vision transformer for classifying breast cancer from mammogram images. According to the experiment, the transformers achieved a very high classification performance with a perfect area under curve of 1.0. Although transformers have shown remarkable performance in a wide range of image recognition tasks, the architecture is extremely difficult to train with less number of images. Data-efficient Image Transformer (DeiT), a variation of the vision transformer, was presented to address this issue. Chen et al[34]. leverages DeiT for breast cancer classification. They fine-tuned the model on a custom dataset. However, the proposed system achieved relatively low classification performance compared to the existing literature. Since vision transformers integrate self-attention mechanism which consumes quadratic time, some transformers are relatively slow for high-resolution histopathological images. Mehta et al[35]. presented a transformer based breast cancer classifier named HATNet to address this issue. Moreover, the proposed classifier achieved state-of-the-art performance in classifying breast cancer from histopathological images.

A summary of some existing works discussed in this section is presented in Table 1. The table illustrates a lack of high-performing image classifiers that achieve high classification performance in a wide range of histopathological images. Therefore, this article presents a novel classifier that achieves a very high classification accuracy in four magnification levels.

Proposed method

The proposed system can be divided into two main parts. The first part involves data preprocessing where the quality of the input image is enhanced. Subsequently, we construct a robust image classifier for identifying breast cancer images. Figure 1 presents an overview of the proposed system. The following sections illustrate these steps in detail.

Dataset description

The dataset used for this experiment is the well-known BreakHis dataset[36]. This is composed of a total of 7909 histopathological images acquired on 82 patients. P& D Laboratory in Brazil acquired the images from January 2014 to December 2014. The dataset has mainly two types of images (Benign and Malignant) of four magnification levels. Each image in this dataset is of size 752×582. A sample of the dataset is presented in Figure 2. Due to its relatively large number of samples, the dataset is commonly used as the benchmark dataset in breast cancer classification[37]. Table 2 presents a description of the dataset.

Data preprocessing

The histopathological images of this dataset are processed through a four-stage image preprocessing pipeline that includes image denoising, contrast enhancement, unsharp masking, resizing, and rescaling. Figure 3 illustrates the process in brief. Before the preprocessing stage, the dataset has been randomly divided into a ratio of 80:20 for training and testing respectively. Since the method lacks ablation study and hyper-parameter tuning, we have

Paper	Method	Contribution	Limitation
Khan et al[18].	Multi-scale feature fusion	Has relatively high accuracy.	Very heavy architecture makes the model slow.
Ijaz et al[19].	Two stream VGG architecture	Presents a novel VGG architecture.	Achieves relatively low performance on 40X magnification and high training complexity.
Khan et al[20].	DenseNet201, NasNetMobile and VGG16	Achieves an accuracy of 99% on BreakHis dataset.	No differentiation is made based on magnification levels.
Ashurov et al[22].	Xception with CBAM	Achieves very high accuracy.	No comparison with the existing methods is made.
Abbasniya et al[25].	InceptionResNetv2 with ensemble of three bosting classifiers	An extensive comparison of various state-of-the-art methods is conducted.	Relatively low performance on 100X magnification.
Joshi et al[26].	Xception with additional layers	Lightweight CNN architecture presented.	Only 40X magnification level is considered and result analysis lacks details.
Kode and Barkana[27]	knowledge-based feature extraction	Presents a comparison between CNN models and knowledge-based methods.	Only 400X magnification level is considered.
Sharmin et al[29].	ResNet50 with LGB	A leightweight model is employed.	Experimented only one dataset with 162 samples.
Liew et al[30].	DenseNet201 with XGBoost	Relatively high accuracy in both binary and multiclass.	No differentiation is made based on magnification levels.
Jasti et al[31].	AlexNet, relief algorithm with SVM	Presents a method for breast cancer classification from mammography images.	No comparison with the existing methods is made.
Chen et al[34].	Two-view DeiT	Investigates the result of transformers in breast cancer detection.	Relatively low performance.

Table 1. Comparative analysis of existing works.

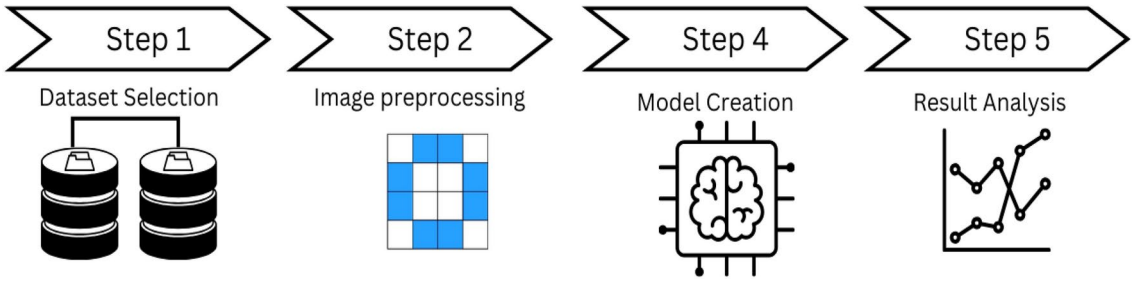

Fig. 1. Overview of the proposed system.

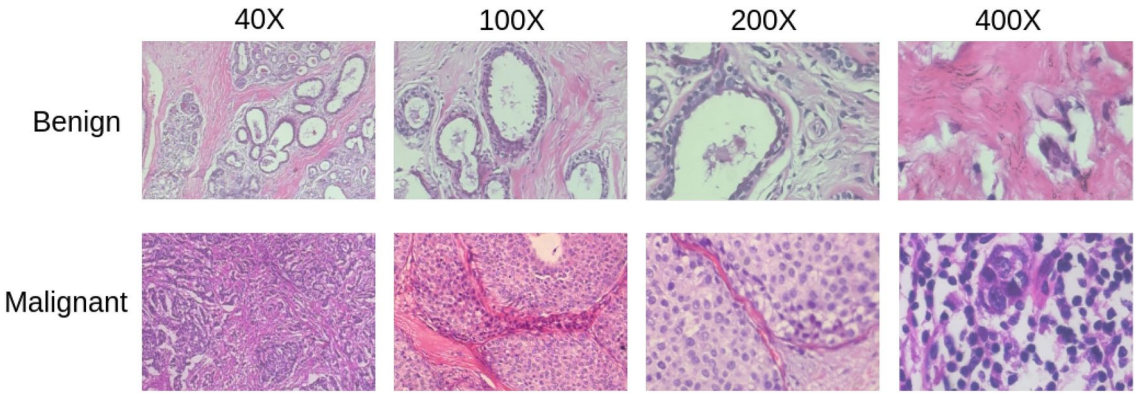

Fig. 2. Sample of the BreakHis dataset.

not kept a set for validation data. We have refrained from allocating a set for validation data due to the absence of an ablation study and hyperparameter tuning.

Image denoising

Images of histopathological breast cancer typically include noise, which might reduce the precision of the diagnostic procedures. Although noise reduction filters are frequently used to improve image quality, conventional low-pass filters have a tendency to unintentionally remove important diagnostic features, which results in information loss. This problem has led to the development of the bilateral filter, a useful tool in the field of medical image processing. The bilateral filter excels at both noise reduction and the preservation of subtle

Magnification	Benign	Malignant	Total
40X	652	1,370	1,995
100X	644	1,437	2,081
200X	623	1,390	2,013
400X	588	1,232	1,820
Total Images	2,480	5,429	7,909

Table 2. BreakHis dataset description.

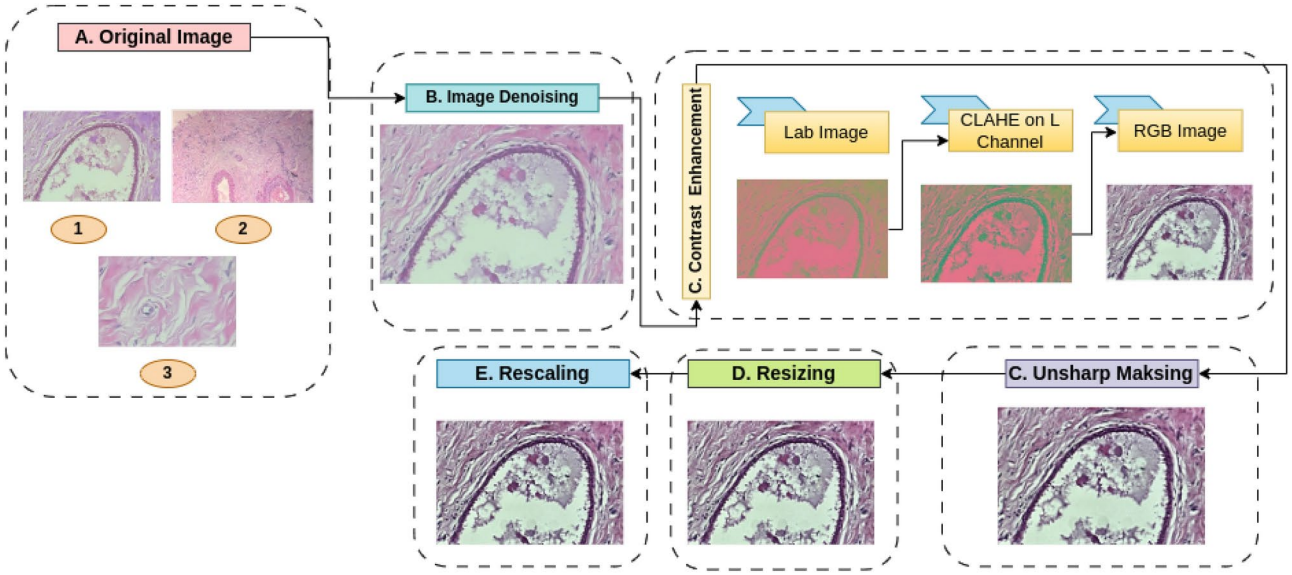

Fig. 3. Image preprocessing overview.

image gradients[38]. It functions by taking into account both intensity similarity and spatial proximity. It uses two main components: a range kernel (usually a Gaussian function) that distributes weights based on the similarity of pixel intensities and a spatial kernel (also a Gaussian function) that provides weights based on pixel proximity. Bilateral filtering carries out weighted averaging of pixel values within a limited neighborhood due to this dual-kernel technique. Pixels with similar intensity and spatial proximity are given larger weights, making the filter extremely useful in preserving sharp transitions in intensity, such as edges and boundaries, while simultaneously lowering noise. Additionally, it provides the freedom to modify variables like kernel widths to precisely control the level of noise reduction and edge preservation. Let I_q represent the intensity value of the image I at q position. Thus the filtered pixel, $BF[I]_p$, can be constructed using Equation (1).

$$BF[I]_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) \cdot G_{\sigma_r}(|I_p - I_q|) \cdot I_q \tag{1}$$

The equation illustrates the filtered image can be generated by considering the weighted sum of neighboring pixel intensities. The weights are determined by the spatial Gaussian kernel, G_{σ_s} and range Gaussian kernel, G_{σ_r} . S denotes the spatial neighborhood of pixel p. Neighboring pixel values are now efficiently blended to create the filtered image. In this experiment, the diameter of the pixel neighborhood is selected 3, and the standard deviation of the Gaussian function for the range kernel and the spatial kernel is set to 75 and 100 respectively. While the higher values of the standard deviations allow effective noise reduction, a smaller pixel neighborhood allows the function to prevent oversmoothing the edges.

Contrast enhancement

The Lab color space is considered to be more appropriate for histopathology image analysis than the RGB color space[10]. Lab color space divides an image into three components: L (lightness), a (green to red), and b (blue to yellow). The image's luminance or intensity information is precisely represented by the L channel, whereas the a and b channels provide the chromatic information. In histopathological images, the diagnostic information is primarily in the luminance details, such as tissue structures and cellular features. It is easier to concentrate on improving the crucial diagnostic details when the luminance and chromatic information are separated. By separating the luminance and chromatic information, we can focus on enhancing the necessary diagnostic

details. Therefore the RGB image is converted to Lab image first for further processing. Since the L channel typically contains the majority of the image's structural and textural information, the Contrast Limited Adaptive Histogram Equalization (CLAHE) is only applied to the L channel.

CLAHE is an image processing approach that improves the contrast and visibility of features in images while reducing noise over-amplification[39]. It is especially useful in applications such as medical imaging where it is necessary to highlight minute details without adding too much noise. It normalizes local intensity variations within a specified neighborhood that limits the contrast enhancement to a local region. If the intensity value after equalization exceeds a certain limit, the value is clipped. In this experiment, an 8×8 neighborhood is considered with a clipping limit of 2.0 which illustrates that if the intensity value is increased by a factor of 2.0, this will be scaled down to meet the limit. Following the CLAHE on the L channel of the Lab image, the image is again converted into an RGB image to fit the pretrained image classifier.

Unsharp masking

Unsharp making is a widely used image sharpening technique. By focusing on the edges, it improves the appearance of the image. This method is divided into four steps. In the first step, the input image I is convolved with a low pass filter f, typically a Gaussian filter. As shown in Equation (2), the convolution operation of input image I and low pass filter f results in a blurred image I_{b}. In our experiment, the low pass filter is a Gaussian filter with a standard deviation of 4.0. Since histopathological images contain subtle details, a higher standard deviation proves beneficial in generating a more effective mask.

$$I_{\mathrm{b}} = I * f \tag{2}$$

In the subsequent step, the blurred image is subtracted from the original image, yielding a masked image that predominantly highlights the edges and fine details. Equation (3) illustrates this process in detail. This process involves iterating through every pixel in x and y dimensions of the image and taking the difference of the blurred image, I_{b}, from the original image, I, to retrieve the masked image I_{m}.

$$I_{\mathrm{m}} = \sum_{i=1}^{x} \sum_{j=1}^{y} I_{(}i, j) - I_{\mathrm{b}}(i, j) \tag{3}$$

The mask is then put through a scalar multiplication with a constant k, which is set to 2 in our experiment. This procedure further emphasizes the edges' visibility. Equation (4) provides a mathematical equation of this procedure that illustrates a simple multiplication of every pixel of the masked image,I_{m}, with a constant K.

$$I_{\mathrm{m}} = I_{\mathrm{m}} \times k \tag{4}$$

Finally, as illustrated in Equation (5), the enhanced mask is summed up with the original image to get the shaped image. In this equation, i and j resemble an arbitrary position on the spatial dimension of the image in x and y coordinates respectively. Therefore, the output image, I_{o}, is obtained by summing the pixel values of the mask, I_{m}, and the input image, I.

$$I_{\mathrm{o}} = \sum_{i=1}^{x} \sum_{j=1}^{y} I_{\mathrm{m}}(i, j) + I(i, j) \tag{5}$$

Resizing

The original histopathological images of size 752×562 are resized to a standard size of 224×224 using nearest neighbor interpolation. This process effectively reduces the computational complexity. For comparison, without resizing, the model would consume more than 8 times the computation. Moreover, resizing images allows them to be more compatible with the pretrained image classifiers.

Rescaling

Lastly, the images are rescaled, with the original range of intensity values being 0 to 255 modified to a standard range of 0 to 1. This rescaling process is essential for maintaining data consistency, and improving interoperability with different machine and deep learning models[40]. It is an essential stage in the preprocessing pipeline that optimizes the images for a variety of computer vision and image analysis applications.

Preprocessing effects

The changes in the image after each preprocessing stage except rescaling are presented in Figure 4. Since rescaling does not result in any visual changes in the image, this step is excluded. The figure holds two images of Malignant and Benign type cancer on 40X and 100X magnifications respectively. Due to the shorter size of the presented images, the impact of image denoising is less visible. However, this method efficiently eliminates subtle noises that are more visible at higher magnification levels. The figure illustrates significant changes are seen after the contrast enhancement stage. The preprocessed image provides a better and more accurate depiction by capturing minute details.

To further understand the impact of the preprocessing steps, we have conducted an ablation study by eliminating three major preprocessing steps (i.e., image denoising, contrast enhancement, and unsharp masking). Table 3 holds the classification accuracy of the proposed method after deleting a specific preprocessing step.

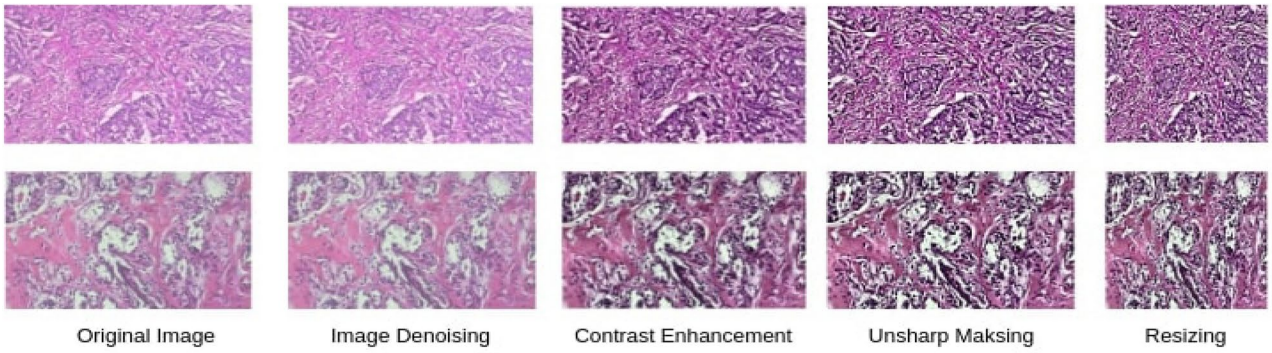

Fig. 4. Effects of image preprocessing on histopathological images.

Magnification	Image Denoising	Contrast Enhancement	Unsharp Masking
40X	0.9913	0.982	0.9897
100X	0.9895	0.9864	0.9874
200X	0.9823	0.9824	0.9807
400X	0.9603	0.9757	0.9757

Table 3. Classification accuracy of the model after eliminating preprocessing steps.

The analysis shows the highest impact of image preprocessing is observed at 400X magnification. The specified preprocessing methods have led to an average increase in classification accuracy of 1.18%. Image denoising has resulted in an accuracy increase of 1.29%, while contrast enhancement and unsharp making have improved the performance by 1.22% and 1.118% respectively.

Model creation

The proposed model is composed of two streams (i.e., the global stream and the multiscale stream). The global stream utilizes a vision transformer that enables the model global context of the image through multi-headed self-attention, while the multiscale stream employs Atrous Spatial Pyramid Pooling (ASPP) layers that return multiscale features with the help of dilated kernels. Each stream extracts features independently. The features are merged before making the final classification. The subsequent sections illustrate this process in detail.

Global stream

The global stream incorporates ViT-base-32, a vision transformer with 86 Million parameters. Vision transformers operate by dividing an image into fixed size non-overlapping segments named patches which is of size 32×32 in ViT-base-32. Following that, the patches are unrolled to convert into 1D vectors. The vectors are then multiplied to a learnable weight matrix and added position embedding. The position-aware patches are now passed through a transformer block which consists of normalization, multi-headed self attention, and multi-layer perceptions. The multi-headed self-attention mechanism allows every patch to attend to every other patch, enabling global context awareness. The ViT feature extractor is initially trained on ImageNet dataset. To enable the model to classify breast cancer images, additional layers have been integrated.

Following the ViT feature extractor, a batch normalization layer is incorporated to speed up the training process. Subsequently, a fully connected (dense) layer with 512 neurons is embedded followed by a 30% dropout. Similarly, two fully connected layers are integrated having 256 and 128 neurons respectively with another 30% dropout in between. Since no data augmentation is used, the dropout layers prevent the stream from overfitting. All the layers in the fully connected layers are activated using Scaled Exponential Linear Unit (SeLu). SeLu, an adaption of the widely used activation function Rectified Linear Unit (ReLu), is developed to address the dying ReLu problem in deep learning[41]. Figure 5 holds a diagrammatic overview of the stream.

Multiscale stream

As presented in Figure 6, the multiscale stream encompasses a custom model made of Atrous Spatial Pyramid Pooling (ASPP) network and Convolutional Block Attention Module (CBAM). ASPP module employs dilated kernels to capture multiscale features. A dilated kernel performs convolution operation for a larger receptive field by introducing gaps between kernel elements. Equation 6 illustrates the process of dilated convolution. Let I denote the input image and i and j be an arbitrary position at X and Y axes of the image respectively. For performing the convolution operation, a kernel, W, is selected of size $K \times L$. Now, if a dilation rate d is applied, each step of the convolution operation will skip $d-1$ pixels along both axes, effectively enlarging the receptive field of the kernel. For d=1, the equation functions as a regular convolution.

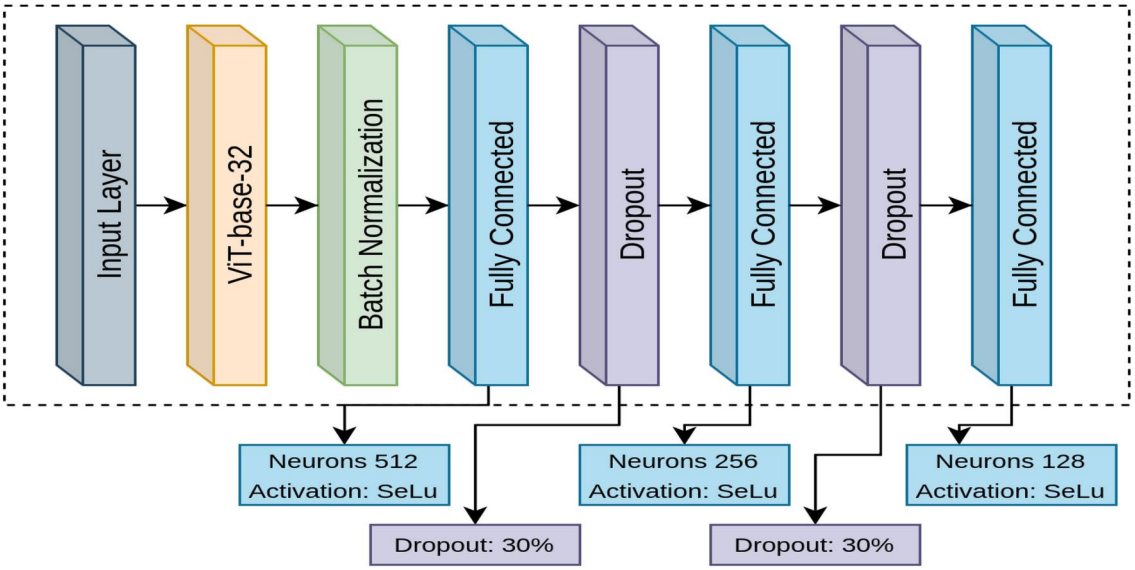

Fig. 5. Breakdown of the global stream used in the fusion model.

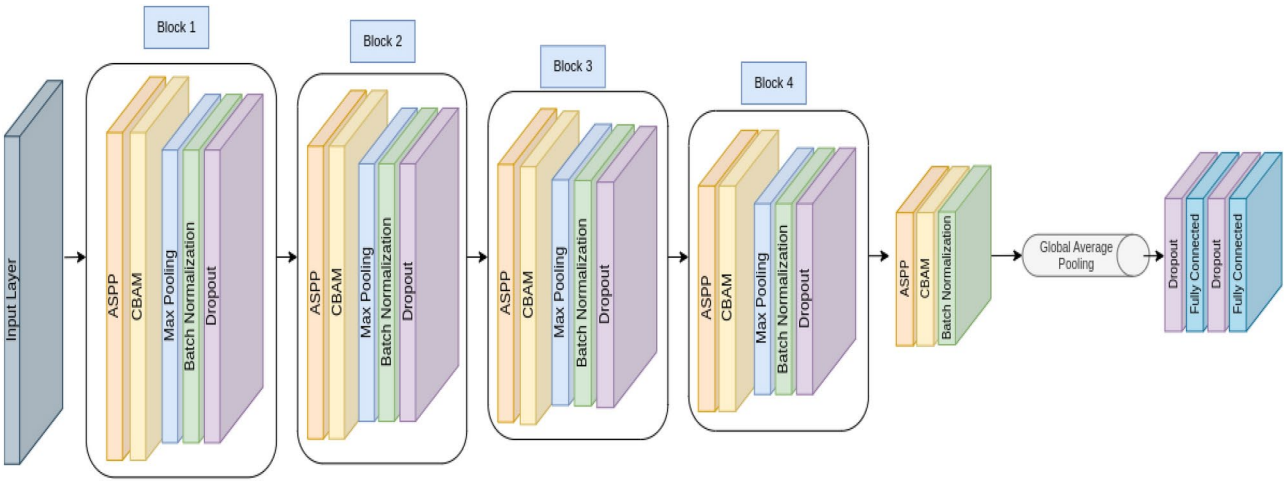

Fig. 6. Breakdown of the multiscale stream used in the fusion model.

$$I_o[i,j] = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} W[k,l] \cdot I[i + d \cdot k, j + d \cdot l] \tag{6}$$

The ASPP module, presented in Figure 7, involves four 3×3 convolutional layers with dilation rates of 1, 5, 7, and 9 respectively. Dilation rate 1 allows the model to capture fine gradients while other layers allow the model to attain global features.

Since some of the features and regions are more important than others, the features are passed through CBAM block for prioritizing. CBAM is a neural network module that dynamically focuses on informative regions within an image to improve the expressive abilities of CNNs. The channel attention mechanism and the spatial attention mechanism are the two main parts that are used to accomplish this. The goal of channel attention is to suppress less informative channels and highlight relevant ones. In order to extract complementary information from the feature maps, this module integrates average and maximum pooling. While average pooling provides a smoother representation of the overall channel, max pooling highlights the most noticeable aspects of each channel. By combining both of these pooling methods, CBAM attends to both local and global features, boosting the model's ability to focus on relevant information in feature maps. The refined features from the channel attention module are then passed to the spatial attention module for prioritizing the spatial regions in the feature maps. This is accomplished through the use of a convolution operation along the spatial dimensions, followed by sigmoid

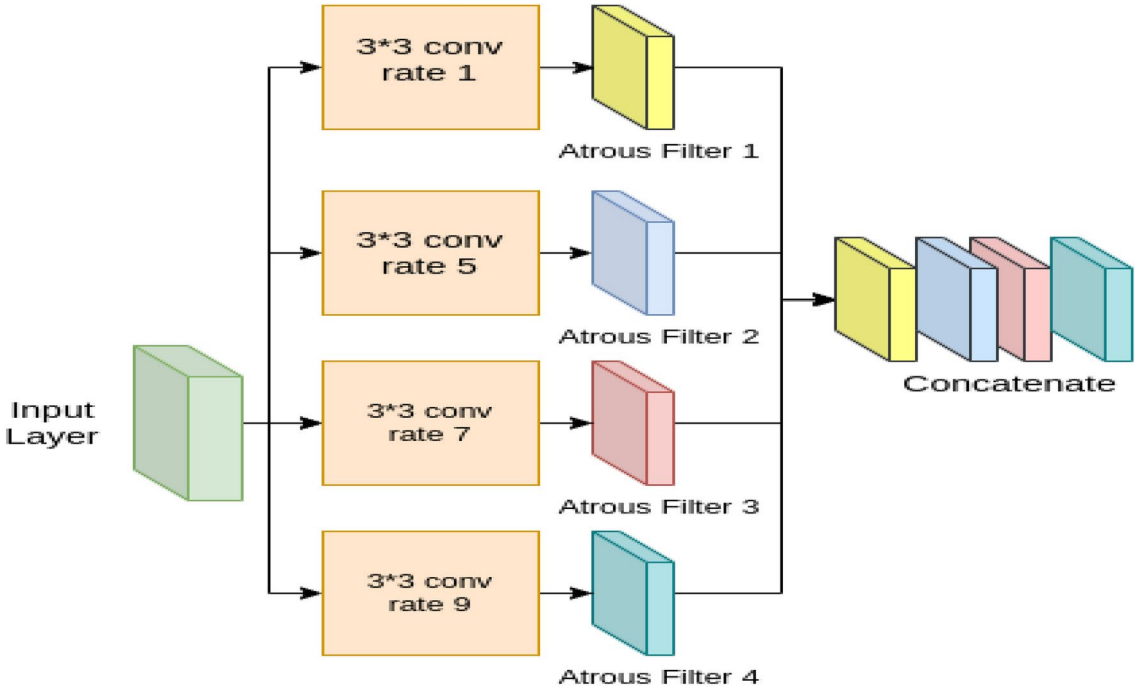

Fig. 7. Breakdown of the atrous spatial pyramid pooling model.

activation. The resulting spatial attention map emphasizes key spatial regions while suppressing others that are less important. Subsequently, element-wise multiplication is used to merge the spatial attention map with the original feature map, creating a modified feature map that highlights key spatial regions.

Following the CBAM block, a 2×2 max pooling block is integrated to reduce the dimensionality. Subsequently, a batch normalization layer is employed to make the training process of this large model faster. The last layer of each block employs a dropout layer to prevent overfitting. The first block has a 20% dropout rate while the others have a 10% random dropout. The multiscale stream incorporates four blocks of feature extractor. After the original feature extractor, additional ASPP and CBAM blocks are added for higher performance supported by batch normalization for faster convergence.

The feature extractor returns 2D feature maps which are converted to 1D feature vectors using the global average pooling layer. The subsequent layers incorporate two dropout layers with 30% random dropout rate and two fully connected layers with 256 and 128 neurons respectively. The fully connected layers are activated through SeLu.

Fusion process

The features extracted from the global streams and the multiscale stream are fused before the final classification. The features are then passed through four fully connected layers. The first three layers have 128, 64, and 32 neurons respectively, and are activated through the SeLu activation function. The final fully connected layer has two neurons with softmax activation function. The final model also incorporates two dropout layers after the first and second fully connected layer with a 30% dropout rate. Figure 8 illustrates the proposed model in brief. The model outputs probabilities of an image being benign or malignant type breast cancer. The type having the higher probability is returned as the inference. The model is trained with categorical-crossentropy loss function and Adamax optimizer. The training process lasted for 100 epochs.

Results

This section describes the results obtained from the experiment along with a thorough comparison with the existing works.

Experimental setup

We have conducted this experiment on Kaggle, a data science platform. We have leveraged Python programming language (version 3.7) along with six Python libraries. The libraries are Numpy (version 1.16.4), Pandas (version 2.0.2), OS(version 3.7), Matplotlib (version 3.0.3), SkLearn (version 1.0.2), and TensorFlow (version 2.1.0). The training process is sped up using GPU P100.

Evaluation metrics

For evaluating the model's performance, we have employed five evaluation matrics namely accuracy, precision, recall, f1-score, and Matthews Correlation Coefficient (MCC). Accuracy is a fundamental measure of the model's overall correctness in producing predictions, whereas precision focuses on the proportion of true

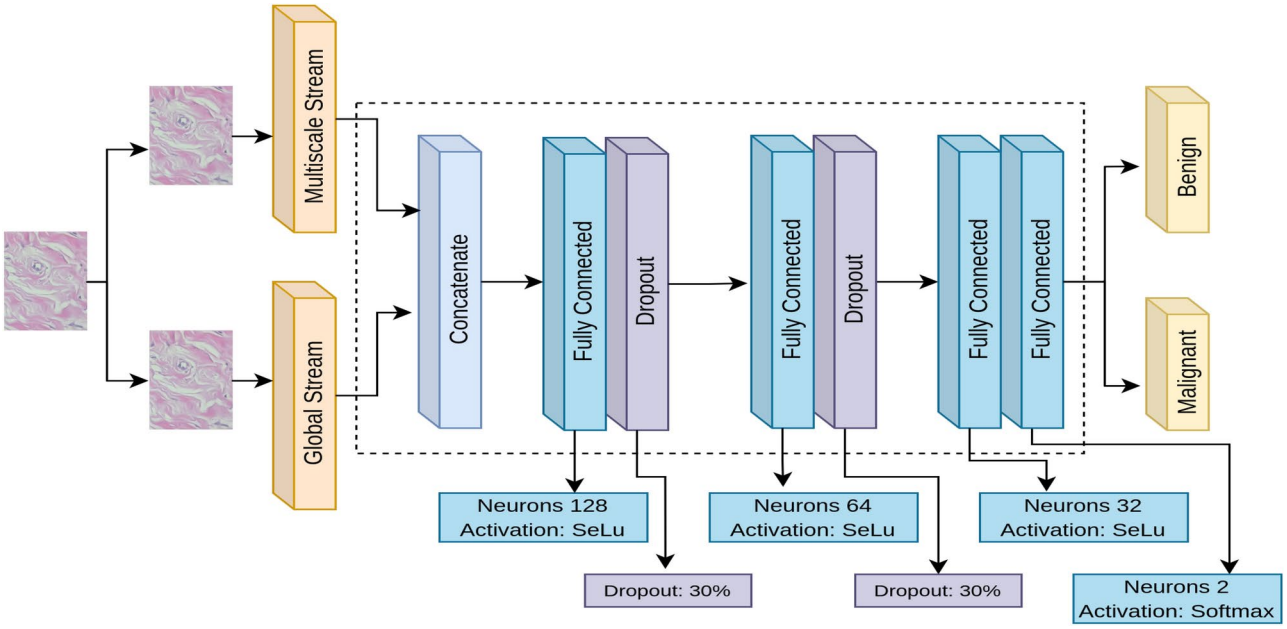

Fig. 8. The proposed fusion model for breast cancer classification.

Magnification	Accuracy	Precision	Recall	F1-score	MCC
40X	0.9474	0.9438	0.9390	0.9413	0.8828
100X	0.9209	0.8926	0.9241	0.9060	0.8161
200X	0.9007	0.8655	0.9031	0.8808	0.7677
400X	0.9313	0.9175	0.9234	0.9204	0.8409

Table 4. Classification performance of the global stream on different magnification levels.

positive predictions among all positive predictions, evaluating the model's capacity to minimize false positives. Recall, sometimes referred to as sensitivity, evaluates the model's ability to accurately identify every positive event, whereas the F1-Score balances recall and accuracy by accounting for erroneous positives and false negatives. Nonetheless, MCC stands out as a particularly robust metric for binary classification. MCC considers all elements of the confusion matrix, making it the most useful metric in binary classification[42]. The equations below exhibit the mathematical approach to calculating these metrics.

$$.Accuracy = \frac{TP + TN}{TP + PP + TN + FN} \tag{7}$$

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

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{10}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{11}$$

To further understand the model's performance, we have presented the confusion matrix and Receiver Operating Characteristic (ROC) curve.

Result analysis

We have evaluated our model on the BreakHis dataset on four magnification levels. For a detailed analysis of the result, the breakdown of the performance of the global and multiscale stream is also presented in Table 4 and 5 respectively. According to the analysis, the global stream produces higher performance than the multiscale

Magnification	Accuracy	Precision	Recall	F1-score	MCC
40X	0.9198	0.8988	0.9165	0.9068	0.8151
100X	0.6259	0.7307	0.6906	0.6207	0.4194
200X	0.9007	0.8660	0.9062	0.8819	0.7711
400X	0.8709	0.8806	0.8611	0.8664	0.7415

Table 5. Classification performance of the multiscale stream on different magnification levels.

Magnification	Accuracy	Precision	Recall	F1-score	MCC
40X	0.9925	0.9943	0.9891	0.9916	0.9834
100X	1.0	1.0	1.0	1.0	1.0
200X	0.9826	0.9758	0.9828	0.9792	0.9585
400X	1.0	1.0	1.0	1.0	1.0

Table 6. Classification performance of the model on different magnification levels.

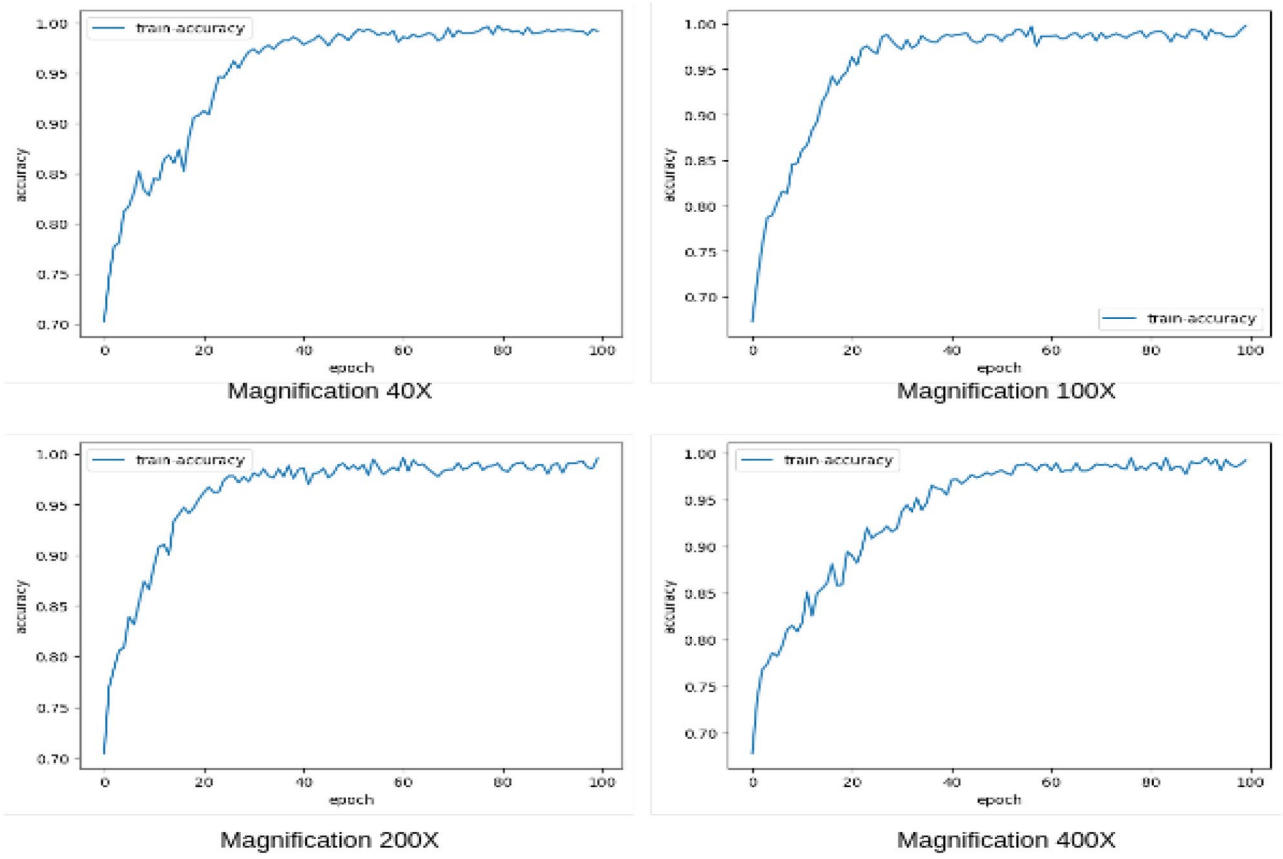

Fig. 9. Training accuracy over the number of epochs.

stream. While both streams produced satisfactory results in the majority of the cases, the multiscale stream overfitted on 100X magnification level. However, fusing the streams yields a perfect classification performance.

Table 6 holds the results of the fusion model on five evaluation metrics. The results illustrate the model achieves 100% classification accuracy in classifying Benign and Malignant types of breast cancers on 100X and 400X magnifications. Remarkably, the model achieves a significant accuracy of 99.25% and 98.26% in 40X and 200X magnifications respectively. Since the dataset is slightly imbalanced, other metrics such as high MCC score exhibit the model's discriminative ability. To measure the convergence of the proposed algorithm, Figure 9 illustrates the training accuracy across the number of epochs. The figure depicts that the fusion model can reach a near-perfect training accuracy within very few epochs of training, particularly on the 100X and 200X magnification datasets. The convergence graph exhibits an opportunity for additional investigation in

determining the optimal number of epochs, which could lead to a significant reduction in resource consumption since the proposed solution consumes over six hours of GPU resources.

For further understanding of the model's prediction, confusion metrics are presented in Figure 10. The figure illustrates three Malignant type cancers are misclassified as Benign on 40X magnification. On 200X magnification, however, two Malignant are misclassified as Benign, and five Benign are mispredicted as Malignant. Other than the mentioned ones, all the breast cancers are classified correctly.

The ROC curve presents a diagrammatic representation of the trade-off between the true positive rate and the false positive rate. The ROC curve of the model on four datasets is presented in Figure 11. The curves demonstrate the model's reliability with a perfect value of 1.0 in all the datasets except 200X magnification. On 200X magnification, however, the model achieves a micro-average ROC of 0.97 which is a noteworthy performance.

We have also presented the model's attention heatmap obtained from Grad-CAM[43]. Grad-CAM is a commonly used technique for visualizing the deep neural network's attention to specific regions in an input image. Figure 12 presents the attention heatmap of the model on the 40X magnification factor. The figure illustrates the model predominantly focuses on the detailed structural regions. Through the highlighting of these informative regions, the model effectively discriminates between Benign and Malignant patterns in a range of histopathology images.

Comparison with existing works

A comparison with the recent works of breast cancer classification on BreakHis images is presented in Table 7. The table holds the accuracy of the model on different magnification factors. The majority of the image classification works rely on CNNs. However, due to the limited ability to capture global and multiscale information, CNNs tend to produce relatively lower classification performance. Fusing CNNs with transformer architectures, however, enables the model to accommodate global context. Ukwuoma et al[44]. leveraged both CNN and transformer architecture and achieved 100% accuracy on 400X magnification. Joseph et al[45]. have employed handcrafted feature extractors which were classified using a feed-forward network[45]. This approach, however, resulted a poor performance. Although some researchers have employed machine learning classification heads instead of MLP after the CNN feature extractor, the limitation of global and multiscale features remains which is essential for histopathological images. Since all the related works do not use the same percentage of data for training, we have also mentioned the portion of the training dataset. While the majority of the works have kept a certain percentage for training, there has been also a method that randomly selected 800 images (400 from each class) due to the imbalanced dataset. According to the results analysis, the model is found to outperform the majority of the existing solutions for effective breast cancer classification.

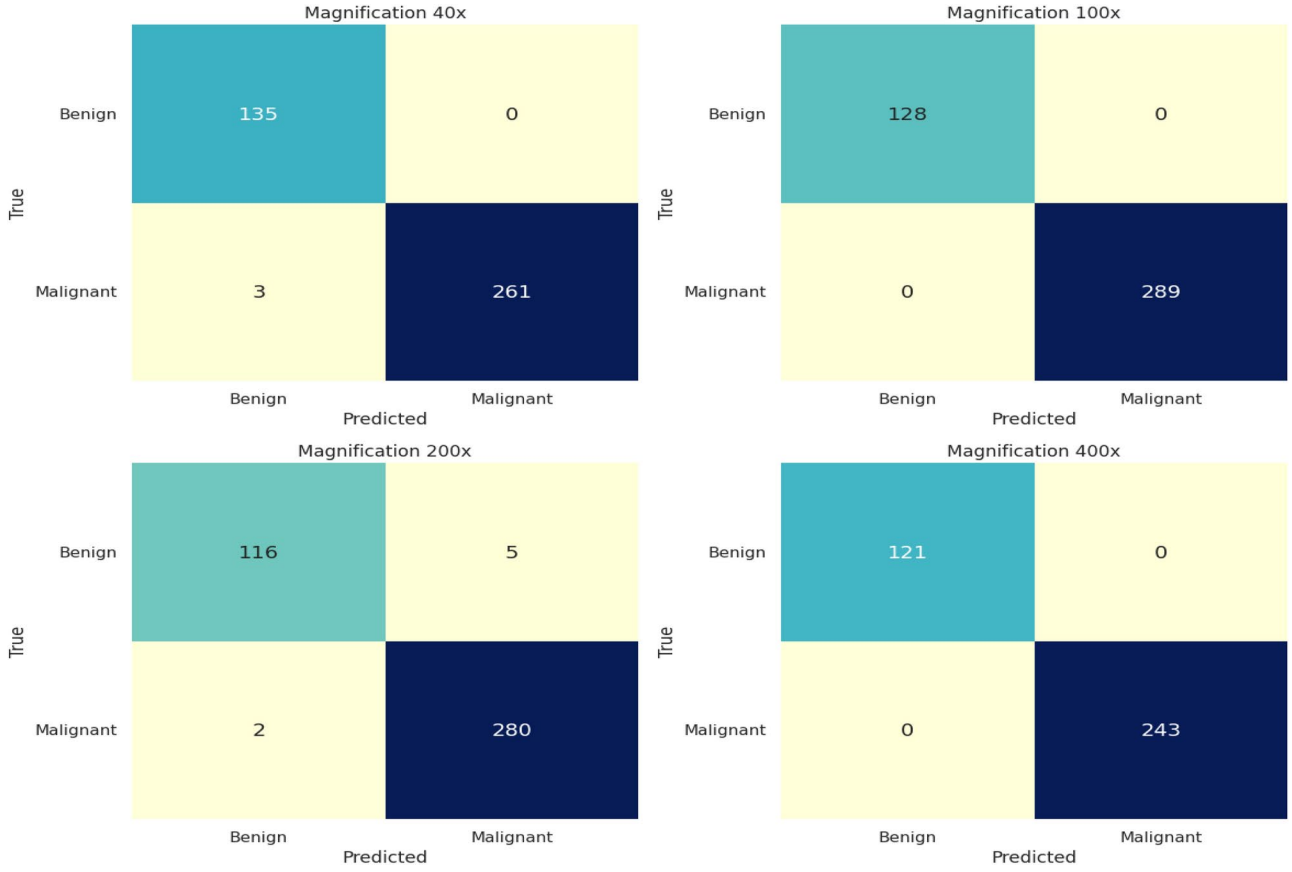

Fig. 10. Confusion matrix of the proposed model on different magnification levels.

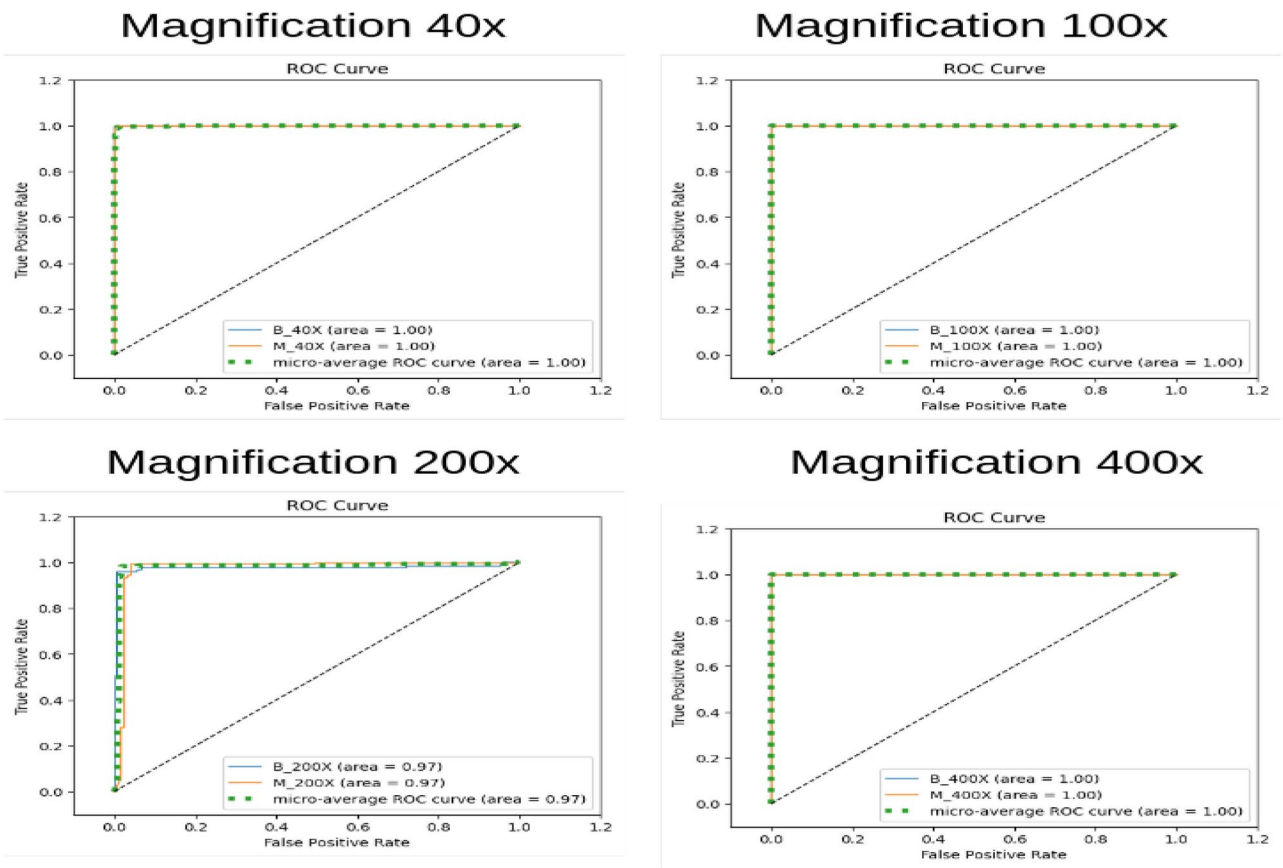

Fig. 11. ROC curve of the proposed model on different magnification levels.

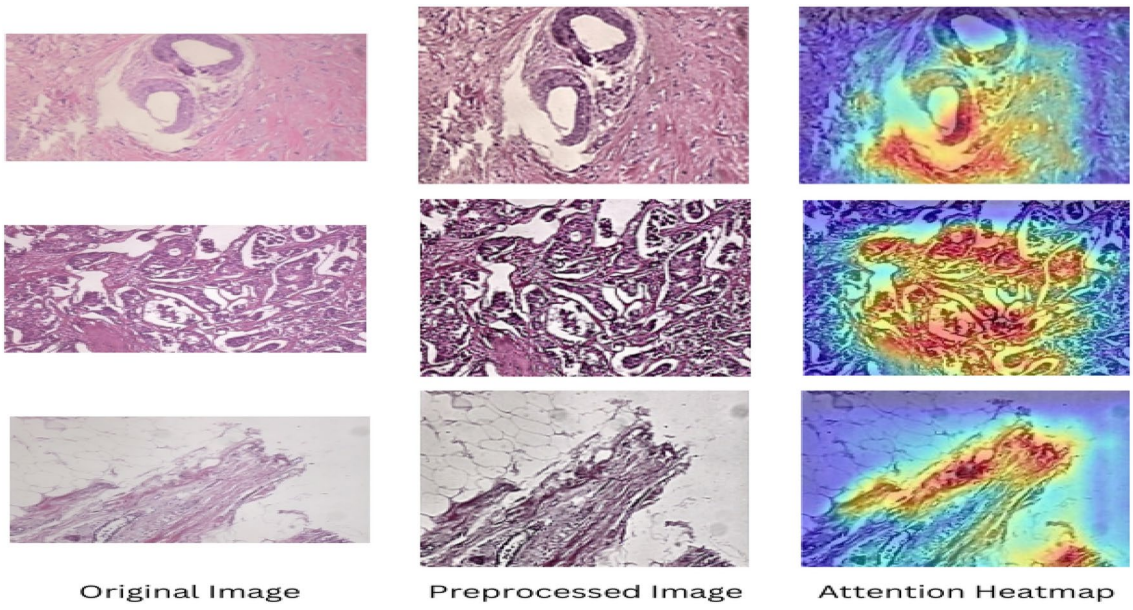

Fig. 12. Attention heatmap of the model on 40X magnification.

Method	Training data	40X	100X	200X	400X
InceptionResNetV2, ensemble of boosting[25]	70%	96.82	95.84	97.01	96.15
Handcrafted features, feed-forward layer[45]	80%	90.87	89.57	91.58	88.67
Fusion of seven CNN models with four classification head[46]	-	92.61	92.00	93.93	91.73
Handcrafted features, Xception[47]	800 images	96.25	96.25	95.74	94.11
VGG, CBAM[48]	70%	98.44	98.04	97.02	97.65
DenseNet201, VGG1, vision transformer[44]	80%	99.80	99.80	99.80	100.00
AlexNet, SVM[49]	70%	87.85	86.68	87.75	85.30
Custom CNN[50]	60%	88.59	87.21	87.21	90.86
ViT and ASPP fusion model	80%	99.25	100.00	98.26	100.00

Table 7. Comparison of classification accuracy on BreakHis dataset.

Discussion

The effectiveness of our proposed model is compared with the recent studies and the results suggest the superiority of this model over the existing ones. The higher performance can be attributed to two main factors. Firstly, the majority of the research work does not integrate image preprocessing except for data augmentation. In this research, however, we have employed five stage image preprocessing techniques which enhance the image quality making the classification process more robust. Although the system does not integrate data augmentation, the overfitting problem is solved through the dropout layer. Secondly, the model fuses global context and multiscale context for making the classification. The global context is critical in histopathological classification since it provides a larger perspective on the tissue or cell structures. Histopathological images frequently have complex patterns and details at various scales. It is challenging to comprehend the general structure and content of tissue without considering the larger context into account, which is essential for precise classification. Global context facilitates the identification of connections among various patches in the image, enabling a more thorough comprehension of the sample and enhancing the precision of classification tasks. Multiscale context, on the other hand, enables the analysis of structures and features at different levels of granularity. By accounting for the variation and complexity of histopathological samples, multiscale analysis ensures that both fine-grained and global properties are analyzed. The proposed system incorporated both global and multiscale contexts with the help of ViT and ASPP modules. By fusing the features extracted from the two streams, a robust classifier is presented that achieves 100% accuracy on 100X and 400X magnification factors. The proposed system, however, has two major limitations. Firstly, the system lacks a thorough ablation study. The study presents the impact of the hyperparameters which is unclear in this work. Moreover, due to the integration of two heavy streams, the fusion model becomes extremely heavy. Training such models are resource extensive. Furthermore, deploying a heavy classifier in a resource-constrained environment is challenging. Nevertheless, considering the noteworthy classification performance of the model, it can be an appreciable option for real-time breast cancer classification.

Conclusion

In this paper, we have presented a system for automatic classification of breast cancer from histopathological images. The system employs five stage image preprocessing technique for enhancing the image quality. The processed images are then passed to a fusion model composed of vision transformer and a custom Atrous Spatial Pyramid Pooling network. This network efficiently extracts useful global and multiscale features which are fused for final classification. The model achieves a perfect 100% accuracy in the BreakHis dataset on two magnification factors (100X and 400X). On 40X and 200X magnification, the model demonstrates a noteworthy performance of 99.25% and 98.26% accuracy respectively. Compared to the existing solutions, the model exhibits superior performance indicating its reliability in real-life breast cancer classification. The system, however, has certain limitations. Since no transfer learning is employed on the Atrous Spatial Pyramid Pooling network, the training process consumes a significant amount of resources. Moreover, no ablation studies have been performed, leaving the impact of the hyperparameters unknown. Addressing these constraints in future studies can improve diagnostic reliability.

Data availability

The BreakHis dataset[36] is publicly available and can be accessed from this link: https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/.

Code availability

The code for this experiment is available at: http://doi.org/10.5281/zenodo.11001346.

Received: 15 December 2023; Accepted: 30 October 2024
Published online: 09 November 2024

References

1. Shamai, G. et al. Deep learning-based image analysis predicts pd-l1 status from h &e-stained histopathology images in breast cancer. *Nature Communications* **13**, 6753 (2022).
2. Giaquinto, A. N. et al. Breast cancer statistics, 2022. *CA: a cancer journal for clinicians* **72**, 524–541 (2022).
3. De Schepper, M. et al. Results of a worldwide survey on the currently used histopathological diagnostic criteria for invasive lobular breast cancer. *Modern Pathology* **35**, 1812–1820 (2022).
4. Rahaman, M. M., Millar, E. K. & Meijering, E. Breast cancer histopathology image-based gene expression prediction using spatial transcriptomics data and deep learning. *Scientific Reports* **13**, https://doi.org/10.1038/s41598-023-40219-0 (2023).
5. Dar, R. A., Rasool, M., Assad, A. *et al.* Breast cancer detection using deep learning: Datasets, methods, and challenges ahead. *Computers in biology and medicine* 106073 (2022).
6. Balkenende, L., Teuwen, J. & Mann, R. M. Application of deep learning in breast cancer imaging. In *Seminars in Nuclear Medicine*, vol. 52, 584–596 (Elsevier, 2022).
7. Green, A. G. et al. A convolutional neural network highlights mutations relevant to antimicrobial resistance in mycobacterium tuberculosis. *Nature communications* **13**, 3817 (2022).
8. Pramanik, R., Dey, S., Malakar, S., Mirjalili, S. & Sarkar, R. Topsis aided ensemble of cnn models for screening covid-19 in chest x-ray images. *Scientific Reports* **12**, 15409 (2022).
9. Aggarwal, R. et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ digital medicine* **4**, 65 (2021).
10. Rashmi, R., Prasad, K. & Udupa, C. B. K. Breast histopathological image analysis using image processing techniques for diagnostic purposes: A methodological review. *Journal of Medical Systems* **46**, 1–24 (2022).
11. Sun, X., Zhang, Y., Chen, C., Xie, S. & Dong, J. High-order paired-aspp for deep semantic segmentation networks. *Information Sciences* **646**, 119364 (2023).
12. Liu, Y. et al. Image semantic segmentation approach based on deeplabv3 plus network with an attention mechanism. *Engineering Applications of Artificial Intelligence* **127**, 107260 (2023).
13. Qiu, Y. *et al.* A2sppnet: attentive atrous spatial pyramid pooling network for salient object detection. *IEEE Transactions on Multimedia* (2022).
14. Dierickx, P., Van Damme, A., Dupuis, N. & Delaby, O. Comparison between cnn, vit and cct for channel frequency response interpretation and application to g. fast. *IEEE Access* **11**, 24039–24052 (2023).
15. Xu, H. *et al.* Vision transformers for computational histopathology. *IEEE Reviews in Biomedical Engineering* (2023).
16. Dosovitskiy, A. *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020).
17. Zhou, X. et al. A comprehensive review for breast histopathology image analysis using classical and deep neural networks. *IEEE Access* **8**, 90931–90956 (2020).
18. Khan, H. U., Raza, B., Waheed, A. & Shah, H. Msf-model: Multi-scale feature fusion-based domain adaptive model for breast cancer classification of histopathology images. *IEEE Access* **10**, 122530–122547 (2022).
19. Ijaz, A. et al. Modality specific cbam-vggnet model for the classification of breast histopathology images via transfer learning. *IEEE Access* **11**, 15750–15762 (2023).
20. Khan, S. I., Shahrior, A., Karim, R., Hasan, M. & Rahman, A. Multinet: A deep neural network approach for detecting breast cancer through multi-scale feature fusion. *Journal of King Saud University-Computer and Information Sciences* **34**, 6217–6228 (2022).
21. Wakili, M. A. *et al.* Classification of breast cancer histopathological images using densenet and transfer learning. *Computational Intelligence and Neuroscience* **2022** (2022).
22. Ashurov, A. et al. Improved breast cancer classification through combining transfer learning and attention mechanism. *Life* **13**, 1945 (2023).
23. Chen, X. et al. Ct-free attenuation correction for dedicated cardiac spect using a 3d dual squeeze-and-excitation residual dense network. *Journal of Nuclear Cardiology* **29**, 2235–2250 (2022).
24. Sarker, M. M. K. et al. Efficient breast cancer classification network with dual squeeze and excitation in histopathological images. *Diagnostics* **13**, 103 (2022).
25. Abbasniya, M. R., Sheikholeslamzadeh, S. A., Nasiri, H. & Emami, S. Classification of breast tumors based on histopathology images using deep features and ensemble of gradient boosting methods. *Computers and Electrical Engineering* **103**, 108382 (2022).
26. Joshi, S. A. et al. Enhanced pre-trained xception model transfer learned for breast cancer detection. *Computation* **11**, 59 (2023).
27. Kode, H. & Barkana, B. D. Deep learning-and expert knowledge-based feature extraction and performance evaluation in breast histopathology images. *Cancers* **15**, 3075 (2023).
28. Li, Y., Feng, X., Liu, Y. & Han, X. Apple quality identification and classification by image processing based on convolutional neural networks. *Scientific Reports* **11**, 16618 (2021).
29. Sharmin, S., Ahammad, T., Talukder, M. A. & Ghose, P. A hybrid dependable deep feature extraction and ensemble-based machine learning approach for breast cancer detection. *IEEE Access* (2023).
30. Liew, X. Y., Hameed, N. & Clos, J. An investigation of xgboost-based algorithm for breast cancer classification. *Machine Learning with Applications* **6**, 100154 (2021).
31. Jasti, V. D. P. et al. Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis. *Security and communication networks* **2022**, 1–7 (2022).
32. Shen, C., Zhang, K. & Tang, J. A covid-19 detection algorithm using deep features and discrete social learning particle swarm optimization for edge computing devices. *ACM Transactions on Internet Technology (TOIT)* **22**, 1–17 (2021).
33. Ayana, G. et al. Vision-transformer-based transfer learning for mammogram classification. *Diagnostics* **13**, 178 (2023).
34. Chen, X. et al. Transformers improve breast cancer diagnosis from unregistered multi-view mammograms. *Diagnostics* **12**, 1549 (2022).
35. Mehta, S. et al. End-to-end diagnosis of breast biopsy images with transformers. *Medical image analysis* **79**, 102466 (2022).
36. Spanhol, F. A., Oliveira, L. S., Petitjean, C. & Heutte, L. A dataset for breast cancer histopathological image classification. *Ieee transactions on biomedical engineering* **63**, 1455–1462 (2015).
37. Chanchal, A. K., Lal, S., Kumar, R., Kwak, J. T. & Kini, J. A novel dataset and efficient deep learning framework for automated grading of renal cell carcinoma from kidney histopathology images. *Scientific Reports* **13**, 5728 (2023).
38. Asokan, A. & Anitha, J. Adaptive cuckoo search based optimal bilateral filtering for denoising of satellite images. *ISA transactions* **100**, 308–321 (2020).
39. Ha, E.-G., Jeon, K. J., Kim, Y. H., Kim, J.-Y. & Han, S.-S. Automatic detection of mesiodens on panoramic radiographs using artificial intelligence. *Scientific reports* **11**, 23061 (2021).
40. Isaksson, L. J. et al. Effects of mri image normalization techniques in prostate cancer radiomics. *Physica Medica* **71**, 7–13 (2020).
41. Jahan, I., Ahmed, M. F., Ali, M. O. & Jang, Y. M. Self-gated rectified linear unit for performance improvement of deep neural networks. *ICT Express* **9**, 320–325 (2023).
42. Hicks, S. A. et al. On evaluation metrics for medical applications of artificial intelligence. *Scientific reports* **12**, 5979 (2022).
43. Selvaraju, R. R. *et al.* Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, 618–626 (2017).
44. Ukwuoma, C. C. et al. Multi-classification of breast cancer lesions in histopathological images using deep_pachi: Multiple self-attention head. *Diagnostics* **12**, 1152 (2022).

45. Joseph, A. A., Abdullahi, M., Junaidu, S. B., Ibrahim, H. H. & Chiroma, H. Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer). *Intelligent Systems with Applications* **14**, 200066 (2022).
46. Zerouaoui, H. & Idri, A. Deep hybrid architectures for binary classification of medical breast cancer images. *Biomedical Signal Processing and Control* **71**, 103226 (2022).
47. Sharma, S. & Kumar, S. The xception model: A potential feature extractor in breast cancer histology images classification. *ICT Express* **8**, 101–108 (2022).
48. Liang, Y. *& Meng, Z* (An interpretable dual-attention network for imbalanced breast cancer classification. IEEE Access, Brea-net, 2023).
49. Li, X., Li, H., Cui, W., Cai, Z. & Jia, M. Classification on digital pathological images of breast cancer based on deep features of different levels. *Mathematical Problems in Engineering* **2021**, 1–13 (2021).
50. Ashtaiwi, A. Optimal histopathological magnification factors for deep learning-based breast cancer prediction. *Applied System Innovation* **5**, 87 (2022).

Acknowledgements

The authors extend their appreciation to King Saud University for funding this research through Researchers Supporting Project Number (RSPD2024R1027), King Saud University, Riyadh, Saudi Arabia.

Author contributions

Conceptualization, N.I.; methodology, N.I.; software, N.I.; validation, K.M.H., M.F.M.; formal analysis, M.F.M.; investigation, M.F.M.; data curation, N.I.; writing-original draft preparation, N.I., M.F.M., K.M.H., S.A., M.S.; writing-review and editing, N.I., M.F.M., K.M.H.; visualization, N.I.; supervision, M.F.M., K.M.H., S.A., M.S., M.K.B.; project administration, M.F.M., K.M.H.; funding acquisition, S.A., M.S.; All authors have read and agreed to the published version of the manuscript.

Funding

This research is funded by the Researchers Supporting Project Number (RSPD2024R1027), King Saud University, Riyadh, Saudi Arabia.

Declarations

Competing interests

The authors declare no competing interests.

Ethical compliance

The study does not involve human or animal subjects, and all data used were obtained from publicly available sources.

Additional information

Correspondence and requests for materials should be addressed to M.F.M. or M.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.