



GLNET: global–local CNN’s-based informed model for detection of breast cancer categories from histopathological slides

Saif Ur Rehman Khan¹ · Ming Zhao¹ · Sohaib Asif¹ · Xuehan Chen¹ · Yusen Zhu²

Accepted: 18 October 2023 / Published online: 6 November 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

In computer vision, particularly in label categorization, attributing features such as color, shape, and tissue size to each category presents a formidable challenge. Dense features related to each category have been validated in recent studies and developed as a multi-label classification problem. Still, notable difficulties remain in (1) classifying attributes more extensively over different object categories, (2) correlating category vulnerability, (3) capturing features in one way, and (4) predicting category labels of a slide with a dense feature map. We have proposed a pre-trained ResNet101-based novel global–local convolution technique to resolve these issues. The proposed model has used ResNet101 as a backbone with additional convolutional, regularization, and dense layers. This technique has two methods to extract the most contributed histopathological slide features. The global descriptor has helped the model to identify the WSI global feature $W_{GF}^{(color, shape, tissue size)}$. In contrast, the local feature extractor has W_{LF} , which fuses the region of interest toward the slides category. After that, we combined W_{GL} features (WSI to $patch^{3 \times 4}$) as an extension of the informed model to learn dense features of multi-label breast cancer categories. After that, the GLNET model uses a fine-tuning mechanism with informed learned to different categories of faded dense layers. Generally, the global–local blocks make sense of the WSI global feature while gaining the object-of-interest characteristic. The proposed model has used the global–local feature composition for each category of breast cancer. Our proposed model has improved accuracy on two benchmarks and challenging *BreakHis* and *IC1AR2018-BachChallenge* datasets for multi-label cancer category prediction. The model results stated in different evaluation matrices verify that the proposed model gains 2% accuracy compared to the existing classifier. Finally, through a series of experiments, we have demonstrated that the proposed model significantly improves accuracy in training on histopathological slides characterized by their complex nature.

Keywords Deep learning · Breast cancer · ResNet101 · Histopathology slides · CNN

Extended author information available on the last page of the article

1 Introduction

Globally, cancer is the deadliest disease that causes death, around about 9.8 million people yearly. Cancer statistics are represented in global data. Cancer prediction retains multiple challenges in the field of medical diagnosis. Cell growth in humans related to cancer rapidly increases if these damaging cells are not diagnosed at the right time. It will be harmful to the patient. *LC*-Lung Cancer is the most common disease in humans and the stated death-dealing cancer type in humans. The American Cancer Society evaluates that an average of eight women have *BC*-breast cancer in the *USA*, which is confirmed after diagnosis [1].

In the field of medical discipline, *BC* prediction is usually carried through via image processing techniques such as the image of histopathology, mammography, ultrasound, and *MRI*-magnetic resonance imaging. However, different non-surgical ineffective ways to analyze and predict *BC* exist. Using the techniques mentioned above, it is only confirmed to give evidence of malignant or benign tissues once a biopsy is used to analyze the disease [2]. *BC* pathological images comprise the different levels of malignant tissues; these images have large volumes and carry essential information related to *BC*. For diagnosing pathology images, *ML*-machine learning used some pre-processing steps as feature selection; after that, they performed the classification task. The final diagnosis of *BC* has been confirmed by analysis of the pathological digital images. This is the standard way for predicting *BC* in women, and the most crucial thing in all this diagnosis is the accurate identification of the pathological images, as these images play the most critical role in the treatment.

An expert medical examiner analyzes the slides of pathological images to identify the malignant cells in the images. It requires time to diagnose confirmed malignant cells [2]. Moreover, the result of an accurate diagnosis of *BC* mainly depends on the medical examiner's experience. Inaccurate diagnosis by medical examiners causes problems due to different factors. Using a *CAD*-computer-aided diagnosis system and *DL*-deep learning models, which require less time to diagnose with high accuracy, and the result of diagnosis can be determined by other subjective factors. The *CAD* diagnose *BC* with an automated end-to-end classification of benign and malignant tissues that not only solves the irregularity in the efficiency of results but also carries diagnosis results more accurately to help the medical examiner analyze objectives [3].

Premature identification of any disease is the most helpful task for treatment. In the case of *BC*, early identification of malignant cells will play an essential role in pathology as it is an efficient and valuable approach to detecting many diseases. The medical examiner uses the *CAD* system to diagnose *BC* and make a final decision. However, even diagnosis issues arise from different pathologist opinions among medical experts. There are many challenges in auto-classifying images related to pathological images. The factors that cause a high-computational cost of pathology images are objectives, regenerate measures, and quantifiable. Various technological approaches and expertise are employed by diverse personnel in preparing pathological slices, encompassing tissue arrangement,

fixation, and staining. However, a notable challenge arises from the inherent variability in these processes, leading to differences in pathological slices among individuals. The images related to *BC* pathological are always categorized by slight interclass variance and more significant intra-class variance; these categories make it difficult to classify with the fine-grained model, and feature selection is a challenging issue in *BC* pathological image classification [4].

Automatic feature identification and selection are crucial factors in *BC* prediction. *DL* architectures have demonstrated proficiency in addressing these concerns, yielding remarkable results. Using extensive datasets and the availability of GPUs significantly enhances the performance of *DL* models in *BC* prediction. This is why *DL* architecture is most frequently used in pathological images. The *CV*-computer vision community is using *DL* techniques in biomedical image analysis. The different biomedical tasks, such as segmentation tissue, pathological image classification [5–8], and image enhancement, can be exploited with the *DL* architecture. On the other hand, *DL* architecture suffers with an over-fitting problem when a dataset is not large enough. This is the major problem with histopathological image diagnosis. If pathological images are set apart by reducing their dataset volumes, then we are limiting the performance of traditional *CNN* for the classification task. For over-fitting problems, transfer *TL*-learning plays a vital role in handling them. It also knows learning about learning, which transfers knowledge from one domain to another with a similar context. Some awareness-sharing approach is needed to transfer awareness from one domain to another. For this transfer of knowledge, *TL* plays an important role. For a tuned pre-trained network, it depends on the problem whether it requires re-trained whole arguments of the network or requires training only some layers of the network.

The AlexNet model, based on *CNN* architecture, has achieved a 6% improved accuracy compared to traditional models [5]. The problem related to this model is that it requires a large amount of data to train itself [6]. The various structure of traditional *CNN* has been proposed to optimize architecture [7–12]. Reducing the model number of arguments while developing more profound *CNN* architectures was the primary purpose of different structures. To address more complex issues related to trainable arguments. A new technique, ResNet [7], was proposed to handle the gradient vanishing issue. This approach generates a new residual connection to handle the vanishing problem between *CNN* layers. Other architectures are introduced for the embedded vision and mobile application, such as MobileNet [8]. The VGG-NET [9] proposed the concept of convolution layers' exploitation with the little field of receptive (3×3). The new approach has been shown in [10], where instead of using traditional *CNN* layers, inception modules are introduced to combine modules and 22 layers of deep network generate that will decrease the AlexNet's arguments 12 times less.

Tajbask et al. [13] are one of the early studies about the proposed malignant level-related fine-tuning networks. After that, the *BreakHis* dataset has been used by [14], which is the extension of the [13] methodology. The finding stated in [14] is that with the help of fine-tuning, the model with different levels of malignant performs much better than training the model with the initial stage. Another study stated that VGG-16 and AlexNet work better when first using some feature extraction and

performing classification on *BC* pathological images with extracted features [15]. In this *TL* approach, the concern related to the suitable resizing of the pathological images at a different level of malignant. The information loss can occur in the expected resizing of the images at a different level of malignant. The *WSI*-Whole Slide Images are introduced to resolve this issue. In *WSI*, the image is split into sub-images of non-intersecting. These sub-images are known as patches. Due to the patches model, the training size image increased. With the help of these patches, the *TL* model fine-tunes better. To handle the complexity of *WSI* images and reduce computing power, only different patches are provided to train the model. Various selection techniques are proposed to select different patches of pathological images [16].

This work proposes a novel *DL* architecture, *GLNET*, to predict multi-label *BC* categories. This model has the strength to incorporate detailed problem features (*BreakHis*, *Bach Challenge*). Our main contribution to this research has been summarized as follows:

- The proposed model addresses challenges associated with the multiclass problem, including classifying attributes more considerably across various object categories, challenging category vulnerability, and capturing features in a single step with features related to two-stage learning objectives.
- We avoid using the traditional way of extracting features from a slide; here, we have used two methods to extract dense features, both with a specific purpose. Global feature extractor $W_{GF(color, shape, tissue\ size)}$ captures the whole slide's global information. Our model can combine different information (cancer categories global features) by this. Local feature extractor $W_{LF(object\ of\ interest)}$ encapsulates the features (attributes) corresponding *WSI* to patch^{3×4} and cancer categories.
- We have proposed the novel dense feature extractor *CNNs* model with five convolution layers along two filter sizes (256, 128), four batch normalization layers, two GAP layers, and two dense layers with two filter sizes (256, 64) after flattening layer along dropout (0.2) layer for each global–local descriptor extracted. This model can fuse on *GL* features of the slides used for prediction.
- We employ an informed feature intention technique to learn the best weights for each cancer category. This technique will merge global–local feature $W_{GL(color, shape, interest\ region)}$ that helps fine-tune each former and proceeding layer in the proposed *CNN* model.
- This *GLNET* model can classify multi-label cancer categories for final class labels with reason stages. The model verifies its weight's steps belonging to *WSI* to patch^{3×4} categories. This is a notable contribution related to weights because different cancer categories do not have the same features.

2 Related work

Initial studies, which propose a variety of techniques to predict *BC* with high accuracy, have divided mainly into two categories: *ML* and *DL* models. Traditional *CNNs*, widely used in the *CV* community, have three modifications: use *CNNs* as

feature selectors, work as a base for *TL*, and use as a scratch to train models. After studying comparative analysis for different bio-medical applications, it has been approved that the performance of *TL* in many biomedical applications is much better than training models from the beginning [13]. Other studies also stated that *TL* works much better for automatic feature selection in complex problems via the *CNNs* model [14], and it has been observed that ImageNet is used in various histopathological cancer prediction applications with *TL* to pre-trained *CNNs* model. In [17], authors have proposed a *BC* prediction methodology with the help of a *TL* approach. In this approach, they have employed the pre-trained model on ImageNet to predict and classify the *BC* disease.

Recent studies stated that most of the work had used histopathological datasets as a sourcing dataset compared to ImageNet [14, 18–20]. Due to non-distinct tasks in *BC* prediction, *TL* works much better than other traditional approaches. Due to microscopic structures in pathological bio-images, histopathological datasets are considered non-distinct. In [21], authors have used the source *CRC-Colorectal Biopsy & Polypectomy* dataset to target another dataset made with the combination of three tissues (*breast*, *lung*, and *colon*). In this research, they have employed Siamese *NN* for *TL*. Too much computation power is needed to select the dataset source due to 5000 samples of the *CRC* dataset; on the other hand, this approach accurately predicts and classifies cancer patients after selecting the source dataset.

The *ML* approach has been used in previous studies to classify histopathological images. The *GLCM*-Gray Level Cooccurrence Matrix, *PWT*-Pyramid Structured Wavelet Transforms, *LBP*-Local Binary Pattern, and *TWT*-Tree Structure Wavelet Transform have worked as feature selectors and gathered various *ML* approaches for the identification of disease. In [22], authors have used texture descriptors to extract different features from pathological images and classify images into benign and malignant tumors. The *GLCM* feature selector has shown better results than another texture selector. To classify images, they used the *KNN*-K-Nearest Neighbor. Comparatively, the authors [23] have worked with different extractor techniques: *LEP*, *GLCM*, Oriented-FAST, *PWT*, and *TWT* used for feature selection with a variety of classification *ML* algorithms: *1-NN*, *QDA*-Quadratic Discriminant Analysis, *RF*-Random Forest, and *SVM*-Support Vector Machine to prediction benign and malignant tumors via histopathological images. The authors have achieved an accuracy of 80–85%. In [24], they have found similarities to features of the malignant tumor via an *ASVM*-adaptive vector machine and achieved notable accuracy. The same works have been proposed in other studies with various methodologies [25].

The problem stated in previous [25–27] studies with *ML* is the incorrect selection of features. Because the selected feature from the feature extractor works as an input to *ML* models, it is also highlighted in previous studies that most of the *ML*-base approaches fail to work correctly due to wrong feature selection. Therefore, it needs to use an approach that automatically gathers features and performs classification concurrently. The *DL* approach can handle *ML* problems efficiently. Many studies have used the *CNNs* architecture for feature selection, classification, and outperforming.

In [28], authors used the *CNNs* architecture model with fifty layers of neurons to extract features and predict cancer. After that, another ensemble-based technique

combines *CNN*-base mode with recurrent architecture to find the pixel relationships. This combined *CNN* model stack with *LSTM*-Long Short-Term Memory [26]. Due to complex cell structures in histopathological images, *DL* approaches have not worked well as models require more time to train themselves if the dataset is large, i.e., *CRC*, *BreakHis*. The solution to this problem is introducing different patch-based approaches, i.e., *MIP*-Maximum Intensity Projection with *CNN*. Extract discriminative patches from malignant images; these patches have been used to predict and reduce computation costs. In [3], the proposed model combines *CNN*-based architecture with *MIP* and then extracts information most relevant to discriminative patches. The *CNN*s-based architecture with multiple (convolutions and pooling layers) has been proposed [29] that is used to classify malignant tumors. Another *DL* technique, *SE-ResNet* [27], has been used for binary and multi-classification for the *BreakHis* dataset.

The accuracy of the *DL* model depends upon the amount of dataset that has been used during the training process. This is another drawback of the *DL* approach as compared to *ML* approaches. To handle this issue for the *DL* model, *TL* has been introduced. In [30], pre-trained *ResNet-50*, *GoogleNet*, and *AlexNet* models have been used with the *BreakHis* histopathological image dataset. The outcome of this paper stated that *ResNet-50* works much better in terms of accuracy [31].

In [31], authors have introduced similar work [30] with some modifications in *GoogleNet* (*InceptionV3*) architecture. This model's result has better accuracy than the shallow model. Early studies [14, 18] stated that most of the work had been proposed on binary classification. A gap still needs to be filled for histopathological multi-classification problems. The author [20] introduces the multiclass model. This model uses *TL* with a large dataset of histopathological images (900,000). This rebuilt dataset has assembled 22 classification tasks. The comparative analysis of this paper stated that the correct selection of features is more important than fine-tuning during comparative analysis between histopathological dataset and *ImageNet*; in [19], the author has introduced model's *inception* and *ResNet* to classify multiclass automatically. Authors [32] have suggested using a handcrafted feature in the classification [14] of malignant tumors with *DL* approaches.

The paper [33] presents *BG-3DM2F*, a novel *AD*-*Alzheimer's* disease diagnostic framework utilizing structural *MRI* scans. It outperforms existing methods with accuracies of 98.12%, 95.26%, and 96.97% for *AD* classification scenarios. Author [34] introduces *MLCA2F*, a *COVID-19* lesion segmentation *DL* model utilizing *MSCA-Net*-Multi-Scale Context-Attention Network with *MC2F*-Multi-Scale Contextual Feature Fusion and *MCAF*-Multi-Context Attentional Feature modules for enhanced lesion delineation and accuracy. Multiple *MSCA-Net* blocks are employed for intricate lesion structures.

Study [35] introduces the *IRRCNN*-*Inception* Recurrent Residual Convolutional Neural Network model for breast cancer classification, a fusion of *Inception*, *Residual*, and *Recurrent CNN*s. It outperforms existing sensitivity, *AUC*, *ROC* curve, and overall accuracy methods on two public datasets. Paper [36] provides an insightful overview of recent *CNN* advancements, covering eight critical architectural improvements. These encompass spatial exploitation, depth enhancement, multi-path structures, network breadth expansion, dimension manipulation, feature-map

optimization, channel boosting, and attention mechanisms. These advancements hold promise across domains like computer vision and healthcare.

The *DBGC*-Dimension-Based Generic Convolution Block [37] is a versatile unit compatible with various *CNN*-based network models. It offers dimension-wise kernel selection for flexible architectural design, with optimization options that balance *FLOPs*, speed, and accuracy.

Previous studies stated that adding different gated and attention procedures in early models [38] will work better when combined with *DL* approaches. An ensemble technique is also merged with *TL* to get better results [39], and Authors [40] have combined *CNN*-based architecture with *LSTM* models and outperformed them. *BreakHis*, *ICAR-2018*, *KimiaPath960*, *CRC*, *Bioimaging-2015*, *Lymphoma*, *BCDR*, *IDC*, *UCSB*, *BHI*, *Bach-CBCD*, and *CHD* datasets, which have been used in various histopathological applications to BC classification and prediction tasks. In this research, we have used the *BachCBC-Bach Challenge Breast Cancer* and *BreakHis* dataset. Because this is the most updated and widely used in pathological imaging applications, and these datasets are publicly available.

3 Methods and materials

3.1 Histopathological slides dataset

In this research, we have used two baseline histopathological datasets to validate our proposed model with notable results, namely *BreakHis* [14], and *ICAR2018-BachChallenge* [18]. The dataset *BreakHis* comprises a total of 7909 pathological slides, which include samples stained with *H&E*-Hematoxylin and Eosin. Additionally, the dataset *ICAR2018-BachChallenge* encompasses 400 whole-slide imaging (*WSI*) samples. In *BreakHis* 82 patients' slides were presented where 2480 images belong to benign tumors, and 5429 images are related to the malignant tumor. The different magnification levels have been used to capture various intensities of cancer in the dataset, e.g., 400×–200×–100×–40× as shown in Fig. 1. The *BreakHis* has been arranged into eight subcategories with two main categories. A malignant tumor is divided into Ductal Carcinoma (*M-DC*), Lobular

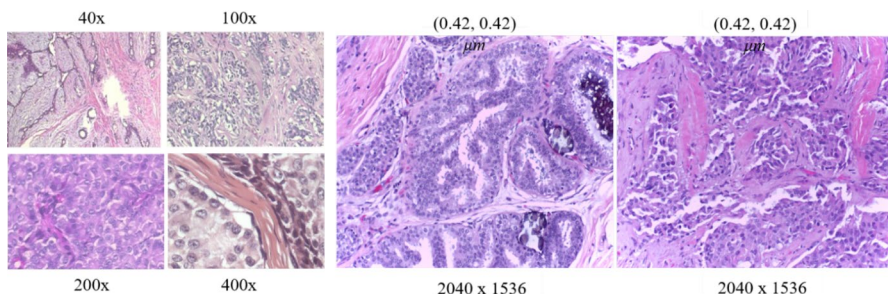
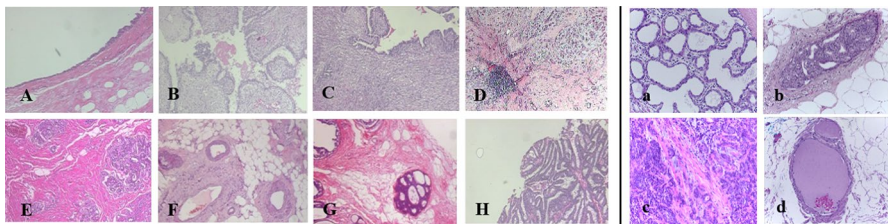


Fig. 1 *BreakHis* magnification levels and *ICAR2018-Bach Challenge* slides

Table 1 *BreakHis Pathological* slides updated 2022 VRI-databases (BreaKHis_v1.tar.gz)

Category	Subcategory	Pathology slide magnification levels				Total slides
		40×	100×	200×	400×	
Malignant	M-DC	864	903	896	788	3451
	M-LC	156	170	163	137	626
	M-MC	205	222	196	169	792
	M-PC	145	142	135	138	560
Benign	B-AA	114	113	111	106	444
	B-FA	253	260	264	237	1014
	B-PT	149	150	140	130	569
	B-TA	109	121	108	115	453
Total		1995	1995	2013	1820	7909

**Fig. 2** BreakHis subgroup of benign; **A** adenosis adenoma, **B** fibro adenoma, **C** phyllodes tumor, **D** tubular adenoma and malignant; **E** ductal carcinoma, **F** lobular carcinoma, **G** mucinous carcinoma, **H** papillary carcinoma and *ICIAR2018-Bach Challenge* slides sub-select; **a**: B, **b**: IS, **c**: IV, **d**: N)

Carcinoma (*M-LC*), Mucinous Carcinoma (*M-MC*), and Papillary Carcinoma (*M-PC*) subcategories, while benign tumor subcategories into Adenosis Adenoma (*B-AA*), Fibro Adenoma (*B-FA*), Phyllodes Tumor (*B-PT*), and Tubular Adenoma (*B-TA*). The distribution of samples to each subcategory is shown in Table 1. The sample of eight subcategories of the main category is shown in Fig. 2. The *ICIAR2018-BachChallenge* contains 400 *H&E*-stained histopathological slides which have been same accretion environment during digitized slides. The samples have the same accrete matrix (2040, 1536) dimension. While the size of the in-plane is (0.42, 0.42) μm as shown in Fig. 1. The pathologists have classified *ICIAR2018-BachChallenge* into two main branches with further four preceding branches. These branches: Benign (*B*), InSitu (*IS*), InVasive (*IV*), and Normal (*N*). Table 2 shows a divided sample of each category in distribution order.

In our research, we conducted validation using two fundamental histopathological datasets, namely *BreakHis* [14] and *ICIAR2018-BachChallenge* [18], to assess the efficacy of our proposed model. The *BreakHis* dataset comprises 7909 pathological slides stained with *H&E*. This dataset has 82 patient cases, resulting in 2480 images associated with benign tumors and 5429 images related to malignant tumors. The dataset encompasses various magnification levels (400×

Table 2 *ICIA2018-BachChallenge* pathological slides (ICIA2018_BACH_Challenge.zip)

Category	Subcategory	Pathology slide WSI Image size (2048 × 1536)
BC	B	100
	IS	100
	IV	100
	N	100
Total		400

200×, 100×, 40×) to capture different intensities of cancer, as illustrated in Fig. 1. *BreakHis* is further categorized into eight subcategories, with two primary categories: malignant tumors, which include *M-DC*-Ductal Carcinoma, *M-LC*-Lobular Carcinoma, *M-MC*-Mucinous Carcinoma, and *M-PC*-Papillary Carcinoma; and benign tumors, which include *B-AA*-Adenosis Adenoma, *B-FA*-Fibro Adenoma, *B-PT*-Phyllodes Tumor, and *B-TA*-Tubular Adenoma. The distribution of samples among these subcategories is detailed in Table 1, while Fig. 2 displays representative samples from the eight subcategories.

The *ICIA2018-BachChallenge* dataset comprises 400 H&E-stained histopathological slides, all digitized under uniform conditions, resulting in consistent image dimensions (2040, 1536). The in-plane size of each image is (0.42, 0.42) μm , as depicted in Fig. 1. Pathologists have classified the *ICIA2018-BachChallenge* dataset into two main branches: *B*-Benign and *IV*-Invasive, further branching into *IS*-InSitu and *N*-Normal (N). The distribution of samples across these categories is presented in Table 2.

3.2 Dataset preprocessing and augmentation

We employed two fundamental filters, Unsharp Masking and Conservative Smooth Masking, to enhance image edges and quality [41]. The application of the Unsharp filter proved particularly effective in rectifying blurred versions of pathological slides, yielding satisfactory outcomes during the restoration process. However, to mitigate the potential overshoot issue, we judiciously introduced a conservative smooth mask in the subsequent image processing steps.

Histopathology slides undertake a crucial digitization process involving the application of *H&E* staining to enhance image clarity and detail. This staining technique is predicated on the absorption of light by specific components within the images. Nonetheless, it is essential to note that variations exist in staining procedures, *WSI* techniques, and fixation processes across different biological laboratories. Consequently, *H&E* images can exhibit notable variations from one laboratory to another. As a result, in the preprocessing phase, we standardized the stain image by transforming histopathology slides into a consistent color space with a standardized variance range.

To address the inherent challenges associated with *H&E* staining variations, we adopted the *H&E* removal process proposed in [42]. This process involves a series of transformations, beginning with a logarithmic transformation to compute the optical density for each stained slide. Subsequently, a decomposition method called single was applied to the density-optical slide. This method facilitated the identification of degrees of freedom and higher variance by developing a 2-dimensional matrix projection. The comprehensive preprocessing steps are illustrated in Fig. 3. In order to tackle potential overfitting issues within our model, we incorporated data augmentation techniques into our datasets. Specifically, we implemented rotation by 20 degrees after each training cycle, shifted images by 0.5 units in width and height dimensions, and applied horizontal and vertical flips to diversify the dataset. These augmentation strategies collectively contributed to our model's robustness and generalization capability.

3.3 Proposed methodology

Breast cancer tissue identification constitutes a prominent domain of research that has garnered considerable attention in previous investigations. This research focus has emerged due to the intricate challenges posed by histopathology slides, the variability stemming from tissue fixation techniques, and issues related to the precision of tissue sectioning. Some attributes present in these slides necessitate comprehensive scene comprehension during model training, while others mandate a focus on densely concentrated optical features. In our study, we introduce a novel GLNET-informed CNN to predict distinct categories of benign and malignant lesions with high accuracy. To facilitate this, we utilizing two foundational histopathological datasets.

Figure 4 illustrates the comprehensive architecture of the GLNET model, including the validation process. Our model initiates by capturing global features, offering an understanding of the entire image structure, encompassing aspects such as color, shape, and tissue patterns. This is achieved by utilizing ResNet101 [43] with additional layers designed for global feature extraction. Following global feature acquisition, we once again employ ResNet101 to extract scene context through 3×4 patches encompassing the entirety of the slide. The proposed multi-label categorization model boasts two notable attributes: Firstly, it identifies the objects present within the slides using a global feature extractor, and secondly, it pinpoints the regions of interest associated with these objects. This innovative approach represents

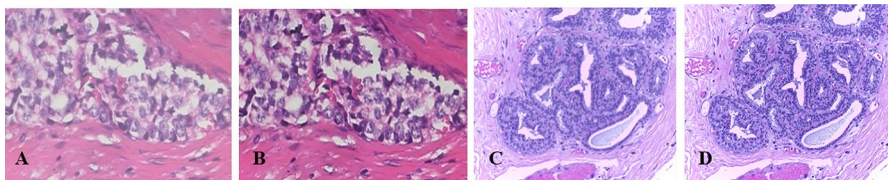


Fig. 3 *BreakHis* **A** unprocessed slide **B** process slide and *ICIAR2018-Bach Challenge* slides **C** unprocessed slide, **D** process slide

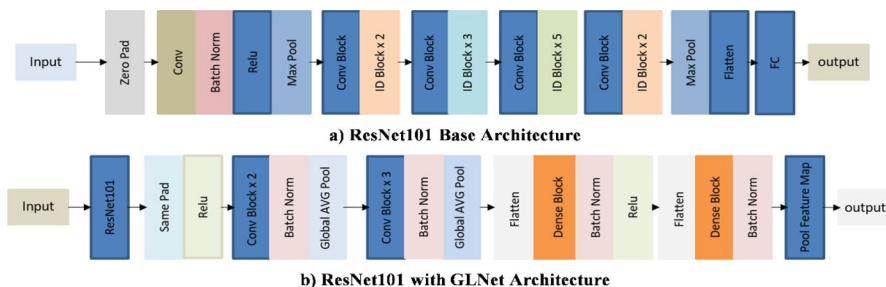


Fig. 4 Proposed GLNET ResNet101-based architecture

a substantial advancement in accurately classifying breast cancer tissue, particularly distinguishing between benign and malignant lesions.

- ResNet101 model

ResNet101, a pre-trained convolutional neural network (CNN) model, has undergone extensive training on a vast dataset comprising over one million images from the ImageNet database [14]. This rigorous training enabled the model to exhibit robust classification capabilities across various object categories, including but not limited to ducks, horses, computers, and keyboards. Leveraging its dense feature representation learned during training, ResNet101 demonstrates competence in handling a broad spectrum of object classes.

In this research, we harnessed the foundational architecture of ResNet101 as the cornerstone for our GLNET model, which features various configurations customized to our specific objectives. The augmentation involved incorporating five convolutional layers, each equipped with two distinct filter sizes (256 and 128), four *BN*-batch normalization layers, two *GAP*-global average pooling layers, and two densely connected layers employing filter sizes of 256 and 64. These additional layers were strategically introduced onto the ResNet101 model while keeping the base pre-trained weights of ResNet101 frozen. Consequently, the pre-trained weights of the ResNet101 model remained unaltered throughout, and the trainable parameter was explicitly set to false to ensure that only the newly introduced layers would undergo weight adjustments. This approach effectively preserved the integrity of the pre-trained ResNet101 model while enhancing its feature extraction capabilities. For a visual representation, refer to Fig. 4, which illustrates the base ResNet101 architecture (a) and its extension featuring the proposed additional layers (b).

Figure 4 illustrates the architectural framework of GLNET, a model designed to analyze histopathological slides. Initially, the model takes the original dimension histopathological slides as input and subjects them to a preprocessing stage. This preprocessing step is essential for capturing global features, denoted as (W_{FG}), retained for subsequent use. Additionally, the model generates patches of size 3×4 from the entire scene to facilitate the comprehension of *ROI*-regions of interest within the histopathological slide (W_{FL}). This task is achieved using

ResNet101 [43], a robust *CNN*. By acquiring global and local features, the model combines the entire scene with the regions corresponding to the areas of interest.

The resultant combined feature set is then presented as input to the model as a vector feature, which is further processed through NET Convolution during the training phase. This step aims to learn and refine the global and local features extracted from the preceding histopathological slides. It is important to note that our proposed GLNET model encompasses two distinct phases related to necrotized information. The first phase entails global–local feature extraction followed by their integration, while the second phase focuses on generating interpretable predictions. In the subsequent sections, we will discuss each component within these phases, providing a comprehensive understanding of the GLNET model's inner workings.

3.3.1 Feature

Generating valuable features holds significance within the classification framework for accurate multi-category prediction. The role of feature generation is multifaceted, ranging from its influence on the final classification model's performance to its ability to enhance the understanding of complex classification problems. Depending on the nature of the problem, feature generation may encompass a comprehensive grasp of the entire image or a focused understanding limited to specific regions of interest associated with an object. In multi-category classification, the generation of precise and relevant features is paramount, as these features contribute significantly to the model's learning process by establishing correlations with the specific problem at hand.

To address this critical aspect, we have devised a two-step feature extraction methodology consisting of global feature extraction (W_{FG}) and local feature extraction (W_{FL}). Each feature extraction step is to address distinct considerations related to feature generation. By integrating global and local features, our model gains a comprehensive view of the visual data while focusing on *ROI* associated with relevant objects. This approach to feature extraction ensures that the model possesses a well-rounded and detailed understanding of the input data, ultimately enhancing its ability to perform accurate multi-category predictions.

3.3.1.1 GFE-Global feature extractor In this study, we exploit the power of the *GFE* (W_{GF}^1 , W_{GF}^2 , and W_{GF}^3) to acquire a comprehensive set of histopathological image-related features from *WSI*, which we refer to as *WSI Features* (W_{FG}). These features encompass critical aspects such as color, shape, tissue composition, dimensions (height and width), and categorical attributes. To execute this process, we employed the ResNet101 architecture [43], which has five convolutional layers equipped with two distinct filter sizes (256 and 128), four *BN* layers, two *GAP* layers, and two dense layers featuring filter sizes of 256 and 64. Additionally, a drop-out layer with a rate of 0.2 was incorporated into the backbone model to enhance robustness. The architecture of GFE is visually represented in Fig. 5, clarifying its role in feature extraction from the *WSIs*.

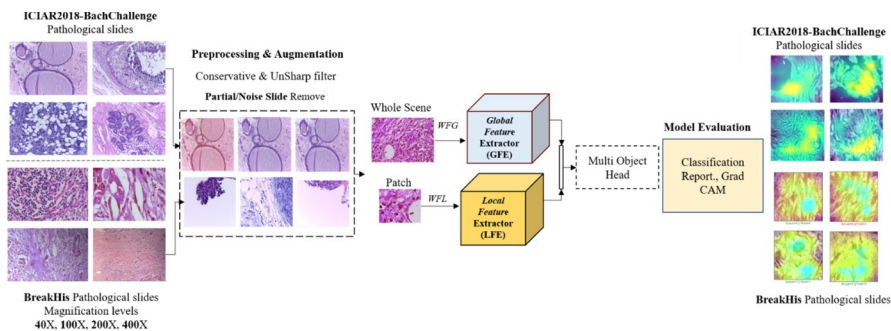


Fig. 5 Architecture diagram of GLNET

3.3.1.2 LFE-Local feature extractor We devised a comprehensive strategy for object identification within *WSI*. The next phase involved the generation of 3×4 patches, which considered not only the central *WSI* patch but also the surrounding edges and the vicinity near the object of interest. The feature extraction process was crucial, utilizing five convolutional layers with two different filter sizes (256 and 128) to capture varying details. Four batch normalization layers were employed to ensure efficient training, and the model incorporated two *GAP* layers. Following this, two dense layers (with 256 and 64 filters) were applied after flattening the feature maps. To prevent overfitting, we introduced a dropout layer with a rate of 0.2 on top of the backbone ResNet101 model.

A fundamental aspect of our methodology is the *LFE* module, which specializes in highlighting the object of interest area. By focusing on this critical region, *LFE* significantly enhances the model's ability to comprehend crucial information for subsequent analysis. The *LFE* architecture's importance in our approach to object identification in *WSI* images is depicted in Fig. 6, illustrating its pivotal role in our overall system.

3.3.2 Batch normalization and global average pooling

In the context of pathological slide analysis, it becomes evident that the distribution of feature weights varies nonlinearly throughout the training process. This phenomenon arises due to the inherent complexity and complexity of pathological slides. Consequently, research indicates that the model's learning capacity should encompass the ability to adapt and fine-tune feature weights continually as the input

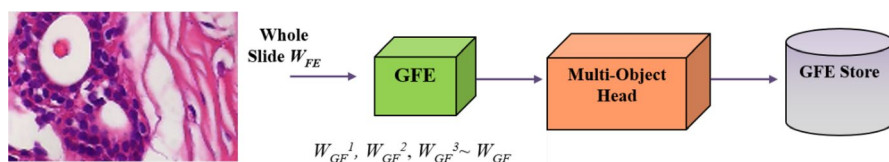


Fig. 6 Entire architecture of proposed GFE methodology

features of preceding layers evolve. This necessitates carefully selecting initial parameters and learning rates to facilitate stable model convergence.

Our proposed model enhances its robustness by incorporating two crucial components within the GLNET model architecture. Firstly, we introduce *BN* [44] layers, which play a central role in normalizing feature distributions during each training iteration. This normalization process contributes to stabilizing the training procedure and accelerating convergence. Secondly, we integrate two *GAP* layers [45], within the model. These layers help combine the information from each multi-label category into a single, concise feature vector. This combination facilitates the efficient representation of complex multi-label features and aids in the overall effectiveness of our model during training and inference.

3.3.3 Regularization-pooled feature map layer

CNN has significant attention in bio-imaging due to its exceptional capability for robust feature extraction. However, when applied to WSI, it has been observed that *CNNs* are valuable to overfitting issues that delay the generalization ability of their fully connected layers (*FCLs*). To mitigate these overfitting challenges, our proposed model incorporates a regularization layer of a unique combination of components. This includes two *BN* layers, two dense layers, one dropout layer, and two ReLU-Activation layers. Furthermore, the model optimizes feature dimensions by introducing a *GAP* layer. Our experimental results indicate the efficacy of GLNet model, demonstrating its superior performance compared to previous studies. The integration of these components addresses the overfitting problem and enhances the model's capacity to extract meaningful features from *WSIs*, ultimately contributing to more accurate and reliable bio-imaging analysis.

3.3.4 Details proposed model architecture

Figure 4 illustrates the comprehensive architecture of our proposed model for multi-label breast cancer diagnosis. The initial step involves preprocessing the slide images, which includes tasks such as removing H&E staining, applying Unsharp Mask to enhance lesion tissue edges, using Conservative Smooth Mask to smooth disordered tissues and reduce overfitting, and introducing augmentation techniques to enhance image quality and interpretation.

For feature extraction, we employ the ResNet101 [43] model, comprising five convolutional layers with two different filter sizes (256 and 128), four *BN* layers, two *GAP* layers, two dense layers with filter sizes of 256 and 64, followed by a flattening layer and a dropout layer (with a dropout rate of 0.2) onto the ResNet101 backbone model. Prior research primarily focused on unidirectional feature extraction from *CNN* models [12, 15, 31]. However, we observed that more than fine-tuning *CNNs* alone for feature extraction was needed to achieve high accuracy on both the *BreakHis* and *Bach Challenge* datasets.

To address this limitation, we introduce the GLNET model, which extracts features from the entire *WSI* through patches of size 3×4 . After obtaining the global–local features, these features are merged into a single feature vector.

Subsequently, we design a novel neural network that initializes with the *GL* features as its initial weights. These fused features traverse two *CNN* layers, incorporating *BN* [44] to handle non-linearity within the slides and applying regularization layers [45] to mitigate overfitting during the weight adjustment in the preceding layers. Finally, feature map layers summarize the trained model's predictions across the multi-label categories of breast cancer, including B-AA, B-FA, B-PT, B-TA, M-DC, M-LC, M-MC, M-PC, B, IS, IV, and N as per the *BreakHis* and *Bach Challenge* Datasets. Our proposed model demonstrates promising accuracy, achieving approximately 86% to 93% accuracy on both datasets, showcasing its effectiveness in multi-label breast cancer diagnosis.

4 Experiments results and discussion

4.1 Experimental setup details

It has been studied that the majority of the research [2, 4, 7, 8] have used only features extracted by *CNNs* in one way. We have changed this traditional method by extracting global–local features with the *CNNs* model. In global feature extraction, all the images have passed with the same image dimension for the *BreakHis* dataset ($400 \times 200 \times 100 \times 40 \times$), while for the *Bach Challenge* dataset, we have to downsize the slides into three dimensions ($1000, 800, 600$) ^{$W \times H$} . After that, we generated the *WSI* patches 3×4 for both datasets with the size $250\text{--}200^{H \times W}$ by using the SVDGEO Desic method. A connected flatten layer trained with activation function (*Relu*) and batch normalization layer [46], dropout rate (0.2) to handle the model overfitting problem. For the rate of learning, our model has proved acceptable evaluation metrics (Sensitivity, Specificity, F1-score, Precision, accuracy) by adjusting the rate of learning of the Adam optimizer (0.001). Net Convolution model, weights trained with *GL* feature extractor Pool Feature Map. Our novel GLNET model experiments have been run on the Co-lab environment, developing the model in Python language using Tensor Flow Packages (Framework of *DL* backend) with Keras. This model has run on NVIDIA GeForce MX350 Ti GPU with 6028 MB RAM. In contrast to traditional approaches employed by prior research [2, 4, 7, 8], which primarily utilized *CNNs* for singular feature extraction, we have introduced a novel methodology using *CNNs* to extract both *GL* features. In the global feature extraction phase, we standardized the image dimensions for the *BreakHis* dataset to $400 \times 200 \times 100 \times 40 \times$. In contrast, for the *Bach Challenge* dataset, due to variations in slide dimensions, we downsized them into three dimensions ($1000, 800, 600$) $W \times H$. Subsequently, we generated *WSI* patches in a 3×4 grid format for both datasets, resulting in patch sizes of 250×200 $H \times W$, accomplished through the application of the SVDGEO Desic method.

To enhance the model's robustness and mitigate overfitting, we introduced a connected flatten layer integrated with an activation function (*Relu*) and a *BN* layer [46]. Furthermore, a dropout rate of 0.2 was implemented. To optimize the learning process, we adjusted the learning rate of the Adam optimizer to 0.001, thereby achieving commendable evaluation metrics such as sensitivity, specificity, F1-score,

precision, and accuracy. The GLNET model architecture was based on *CNN*, with weights trained using the *GL* feature extractor Pool Feature Map.

Our proposed GLNET model experiments were conducted within the Google Colab environment, implemented in the Python programming language with the aid of TensorFlow packages, explicitly utilizing the Keras framework for deep learning. The model's computations were executed on an *NVIDIA GeForce MX350 Ti GPU* with 6028 MB of RAM, ensuring efficient and accelerated processing for our *GL* feature extraction methodology.

4.2 Implementation details of GLNET model

In this study, we assessed the efficacy of our proposed GLNET model using two well-established histopathological datasets, *BreakHis* and *ICIA2018-BachChallenge*. The *BreakHis* dataset comprises images at four distinct magnification levels (400×, 200×, 100×, and 40×). It encompasses eight breast cancer categories. The *ICIA2018-BachChallenge* dataset consists of *WSI* with dimensions (2040, 1536) and encompasses four cancer categories. Detailed sample distribution information for both datasets is presented in Tables 1 and 2.

All experiments were conducted within the *Python IDE* environment, specifically Google Colab. Parameters for the novel GLNET model were selected based on the observed results. As depicted in Fig. 4, our methodology commenced with image preprocessing techniques to enhance image quality. Subsequently, we applied the SVDGEO Desic technique to eliminate noise and partial slides during the global–local feature extraction phase. The dataset category distributions following these preprocessing steps for both datasets are summarized in Tables 3 and 4 (Fig. 7).@@@

Table 3 presents a distribution of histopathological samples at various magnification levels within the eight cancer categories contained in the *BreakHis* dataset. Notably, before analysis, a preprocessing step was implemented to eliminate images containing noise and partial features. Specifically, the SVDGEO Desic method was

Table 3 After remove NOISE & PARTIAL & STAIN color induction during slide creation *BreakHis*

Category	Subcategory	Pathology slide magnification				Total slides
		40×	100×	200×	400×	
Malignant	M-DC	240	260	290	250	1040
	M-LC	120	110	132	103	465
	M-MC	130	159	126	125	540
	M-PC	119	130	105	120	474
Benign	B-AA	99	100	99	92	390
	B-FA	105	140	156	148	549
	B-PT	100	110	120	110	440
	B-TA	96	105	94	101	396
Total		1009	1114	1122	1049	4294

Table 4 After remove NOISE & PARTIAL & STAIN color induction during slide creation
Bach Challenge

Category	Subcategory	Pathology slide WSI Image size (2048 × 1536)
BC	B	79
	IS	97
	IV	76
	N	98
Total		350

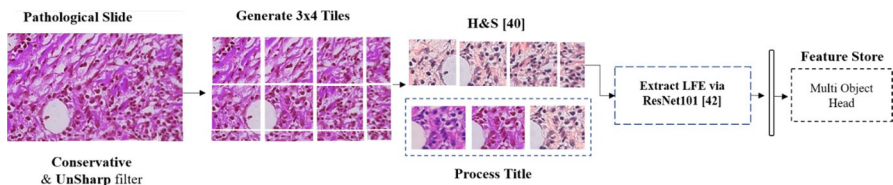


Fig. 7 Architecture of proposed LFE methodology

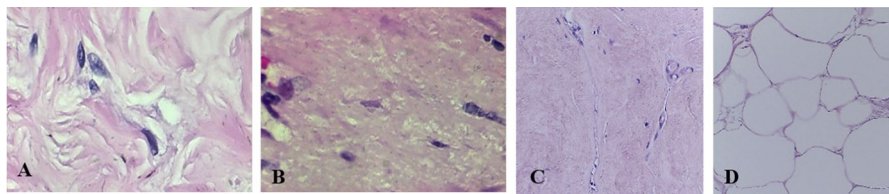


Fig. 8 NOISE & PARTIAL & STAIN color induction slides; **A** BFA-400×, **B** BPT-400×, **N**-1000×, **IS**-1000×

applied, removing 1415 images across different magnification levels. Furthermore, Table 4 provides a detailed distribution of slides within four distinct categories after removing undesirable slide number 50 from the *BachChallenge* dataset. This data refinement process ensures the dataset's quality and relevance for subsequent analysis. For visual clarity, Figs. 8 and 9 offer illustrative examples of images identified as noisy or containing partial information and consequently excluded from the dataset. These steps of preprocessing and data curation are crucial to enhancing the dataset's integrity and suitability for accurate cancer classification tasks, as they help eliminate potential noise and bias sources.

Following the preprocessing of WSI within both datasets, our subsequent step involves extracting global features encompassing color, shape, and tissue size from these processed WSIs. We employed ResNet101 [43] as the foundational architecture to identify potential features of each category in both datasets. Subsequently, we generated a Pool Features Map Vector ($W_{GF}^1, W_{GF}^2, W_{GF}^3 \dots$) to encapsulate these features (1). Here, n represents the total number of samples, and W signifies

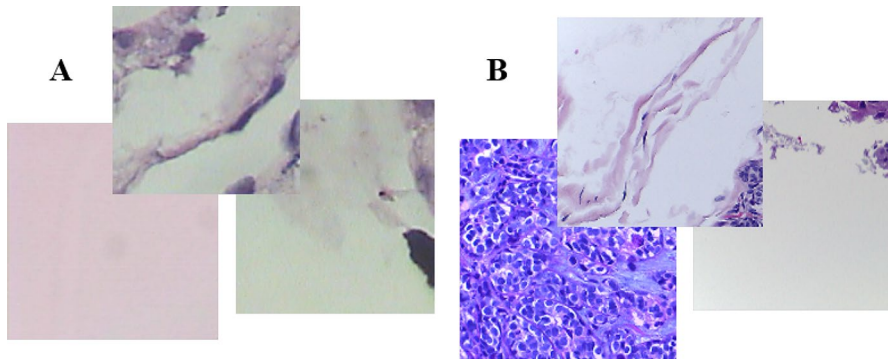


Fig. 9 NOISE & PARTIAL & STAIN color induction slides; **A** *BreakHis*, **B** *Bach Challenge*

the feature weights corresponding to the respective slide category. This approach enables us to capture and store essential global features of complex histopathological slides.

$$\sum_1^n W_{GF} = W_{GF}^1, W_{GF}^2, \dots, \dots, W_{GF}^n \quad (1)$$

LFE has generated the local features (object of interest) toward the interest area. For this purpose, we have generated the patches^{3×4} with patch dimensions 250–200^{W×H} for each category process *WSI* slide, and then, for local feature extraction, ResNet101 [43] has used to extract a region of interest to get a better understanding toward categories learning. Tables 5 and 6 are presenting the total number of patches after removing noise, and partial and unwanted patches from both datasets, while Table 7 is presenting the dataset distribution (*WSI*, *Patch*) in terms of the train (10), validate (10), and test (10) for both datasets. Noise, Stain Color Induction, and Partial patch samples are shown in Fig. 8 from both datasets.

LFE has been employed to extract relevant features corresponding to the region of interest within the images. To achieve this, we generated patches with dimensions of 250 pixels in width (*W*) and 200 pixels in height (*H*) in a 3×4 grid for each category within the whole slide image (*WSI*) dataset. Subsequently, ResNet101 [43]

Table 5 After remove NOISE & PARTIAL & STAIN color induction (*WSI*, *Patches*) *BreakHis*

Category	Subcategory	1000×		800×		600×	
		WSI	Patches	WSI	Patches	WSI	Patches
Malignant and benign	B	135	1882	153	1182	434	1524
	IS	158	1930	186	1488	505	1754
	IV	139	1982	146	1252	426	1635
	N	155	1960	166	1329	468	1623
Total		587	7754	651	5251	1833	6536

Table 6 After remove NOISE & PARTIAL & STAIN color induction patches (WSI, Patches) *Bach Challenge*

Category	Subcategory	Pathology slide magnification levels							
		40×		100×		200×		400×	
		WSI	Patches	WSI	Patches	WSI	Patches	WSI	Patches
Malignant	M-DC	240	2880	260	3120	290	3480	250	3000
	M-LC	120	1440	110	1320	132	1584	103	1236
	M-MC	130	1560	159	1908	126	1512	125	1500
	M-PC	119	1428	130	1560	105	1260	120	1440
Benign	B-AA	99	1188	100	1200	99	1188	92	1104
	B-FA	105	1260	140	1680	156	1872	148	1776
	B-PT	100	1200	110	1320	120	1440	110	1320
	B-TA	96	1152	105	1260	94	1128	101	1212
Total		1009	12,108	1114	13,368	1122	13,464	1049	12,588

Table 7 Datasets distribution (WSI, Patches)

Dataset	Main category	Subcategory	Dataset distribution The ratio of train: validate: test (80:10:10)					
			WSI			Patch 3×4		
			Train	Validate	Test	Train	Validate	Test
<i>BreakHis</i>								
40×	Benign/Malignant	Eight	807	101	101	9686	1211	1211
100×			892	111	111	10,694	1337	1337
200×			898	112	112	10,772	1346	1346
400×			839	105	105	10,070	1258	1258
<i>Bach Challenge</i>								
600×	Benign/Malignant	Four	1467	183	183	5228	654	654
800×			521	65	65	4201	525	525
1000×			469	59	59	6202	776	776

was utilized for local feature extraction, enabling us to gain a deeper insight into the learning of various categories.

Tables 5 and 6 provide an overview of the total number of patches obtained after filtering out both dataset noise, partial, and unwanted patches. Additionally, Table 7 offers insights into the dataset distribution concerning WSI and patches, categorizing them into training (10), validation (10), and testing (10) subsets for both datasets. Figure 8 presents samples of noisy, stain-colored, and partial patches extracted from both datasets for a more visual representation. This comprehensive approach to local feature extraction and dataset curation ensures the quality and relevance of the data used in our study, facilitating effective category learning and analysis.

Table 8 GLNET informed proposed model accuracy (*BreakHis*)

Accuracy	Pathology slide magnification levels			
	40×	100×	200×	400×
<i>BreakHis slides</i>				
Train accuracy	88.13	87.47	86.81	88.76
Test accuracy	90.01	92.32	91.98	91.51

Table 9 GLNET informed proposed model accuracy (*Bach Challenge*)

Category	Pathology slide WSI		
	1000×	800×	600×
<i>Bach Challenge</i>			
Train accuracy	90.21	88.47	84.78
Test accuracy	90.33	87.67	85.44

Following the acquisition of the global–local pool feature map, the obtained *GL* features are merge to encompass the extracted features across various multi-label histopathological categories within the network. Notably, during the feature extraction process, regions of the object that are considered weak or non-essential are reduced. Specifically, one of the composite *GL* features comprises the attention feature derived from ResNet101 [43]. Subsequently, these informed features are fed into the novel *CNNs* model, facilitating training with optimized weights to different categories of breast cancer within both datasets. The GLNET model experiences substantial enhancement by incorporating combined features, encompassing whole-slide and object-of-interest data. The resulting improvements in training and testing outcomes are presented in Tables 8 and 9 for both benchmark histopathological datasets, indicating the efficacy of the GLNET model in the context of breast cancer classification.

4.3 Evaluation metrics details

4.3.1 Analysis of GLNET model performance

4.3.1.1 Classification report The effectiveness of our proposed GLNET model is validated through the results presented in Tables 10 and 11. In each column of the model evaluation matrix, the highlighted value represents the highest achieved result for the respective metric. These results illustrate the model's capability in classifying BC categories at different magnification levels, demonstrating notable sensitivity and specificity.

Specifically, Table 10 highlights that the proposed GLNET model attains a sensitivity and recall of 0.96 for both Papillary Carcinoma and Adenosis Adenoma categories at a 40× magnification level. However, at a 200× magnification level, sensitivity and recall were reduced to 0.85 for both categories. Notably, the GLNET model exhibits its highest recall and sensitivity values at the 400× magnification

Table 10 GLNET (Global, Local Features Informed Network) model evaluation on *BreakHis* histopathological slides

Metrics	40×								200×																											
	MDC				MLC				MMC				MPC				BAA				BFA				BPT				BTA							
	MDC	MLC	MMC	MPC	BAA	BFA	BPT	BTA	MDC	MLC	MMC	MPC	BAA	BFA	BPT	BTA	MDC	MLC	MMC	MPC	BAA	BFA	BPT	BTA	MDC	MLC	MMC	MPC	BAA	BFA	BPT	BTA				
Sensitivity		0.83	0.87	0.94	0.96	0.70	0.87	0.96		0.96	0.84	0.90	0.85	0.85	0.83	0.84		0.96	0.84	0.90	0.85	0.85	0.83	0.83	0.77	0.84		0.96	0.80	0.81	0.83	0.83	0.84	0.82	0.75	0.83
Specificity		0.95	0.80	0.90	0.97	0.88	0.90	0.93	0.85	0.85	0.83	0.85	0.85	0.85	0.83	0.84		0.94	0.85	0.83	0.85	0.97	0.84	0.84	0.79	0.95		0.96	0.84	0.90	0.85	0.85	0.83	0.77	0.84	
Precision		0.95	0.90	0.80	0.93	0.88	0.90	0.83	0.87	0.96	0.70	0.87	0.96	0.85	0.83	0.84		0.96	0.84	0.90	0.85	0.85	0.83	0.83	0.77	0.84		0.96	0.84	0.90	0.85	0.88	0.82	0.76	0.89	
Recall		0.83	0.87	0.94	0.96	0.96	0.70	0.87	0.96		0.96	0.84	0.90	0.85	0.85	0.83		0.96	0.84	0.90	0.85	0.85	0.83	0.83	0.77	0.84		0.96	0.84	0.90	0.85	0.88	0.82	0.76	0.89	
F-score		0.85	0.88	0.86	0.94	0.90	0.77	0.85	0.91		0.94	0.86	0.90	0.85	0.85	0.83		0.95	0.82	0.80	0.85	0.88	0.82	0.82	0.76	0.89		0.92	0.82	0.85	0.88	0.82	0.76	0.89		
Overall test accuracy 0.90																																				
Overall test accuracy 0.92																																				

Bold values represent the highest result achieved by our proposed model in different dataset magnification level

Table 11 GLNET (Global, Local Features Informed Network) Model Evaluation on *BreakHis* histopathological slides

Metrics	100×					400×										
	MDC	MLC	MMC	MPC	BAA	BFA	BPT	BTA	MDC	MLC	MMC	MPC	BAA	BPT	BFA	BTA
Sensitivity		0.89	0.83	0.85	0.95-	0.92	0.79	0.83	0.98	0.90	0.90	0.87	0.88	0.94	0.87	0.91
Specificity		0.88	0.83	0.85	0.92	0.93	0.79	0.82	0.98	0.86	0.88	0.90	0.85	0.96	0.86	0.92
Precision		0.98	0.83	0.86	0.98	0.94	0.85	0.78	0.86	0.91	0.93	0.84	0.90	0.98	0.88	0.89
Recall		0.89	0.83	0.85	0.94	0.92	0.79	0.83	0.98	0.90	0.90	0.87	0.88	0.94	0.87	0.91
F-score		0.92	0.86	0.85	0.95	0.93	0.80	0.81	0.92	0.88	0.86	0.90	0.85	0.96	0.85	0.92
Overall test accuracy 0.92																
Overall test accuracy 0.91																

Bold values represent the highest result achieved by our proposed model in different dataset magnification level

level compared to other magnification levels. At a 100× magnification level, the model achieves its highest recall and sensitivity of 0.98 for Tubular Adenoma classification.

These results underscore the model's robust performance across different magnification levels, highlighting its ability to effectively identify and classify various BC categories with significant sensitivity and specificity, thus demonstrating its potential as a valuable tool in medical image analysis.

Table 12 presents the outcomes of the proposed model when applied to the down-sized Bach Challenge dataset, which is available in three distinct dimensions (1000, 800, 600) in terms of width (W) and height (H). Notably, the results demonstrate that the performance of the GLNET model is somewhat compromised. This can be attributed to inherent variations stemming from the diverse biological laboratories involved, discrepancies in staining tissue procedures, discrepancies in WSI processing, and differences in fixation processes. These disparities make noticeable distinctions in the $H&E$ -stained images across different laboratories.

However, it is worth highlighting that our proposed model consistently outperforms previous studies [2, 4, 7]. Specifically, our model exhibits superior classification capabilities, achieving the highest sensitivity and recall of 0.91 when identifying instances categorized as invasive within the histopathological slide images downsized to (800) $W \times H$. Furthermore, the GLNET model demonstrates comparable sensitivity and recall of 0.90, particularly in image sizes (1000, 600) $W \times H$, effectively identifying non-cancerous tumor tissues. These findings highlight the effectiveness of our proposed model in addressing the inherent challenges posed by variations in laboratory procedures and processes when analyzing histopathological images, showcasing its potential for robust and accurate classification.

4.3.1.2 Analysis with Grad-CAM Our study employs our novel GLNET model to elucidate the rationale behind classification outcomes. To further substantiate the explanatory power of GLNET, we employ the benchmark analysis tool *GRAD-CAM*, a visual technique. Compared to prior research [2, 4, 6], which needed more clarity in explaining model classification decisions, our GLNET model offers

Table 12 GLNET (Global, Local Features Informed Network) model evaluation on *ICIA2018-Bach-Challenge* histopathological slides

Metrics	1000×				800×				600×					
	(B)	(IV)	(IS)	(N)	(B)	(IV)	(IS)	(N)	(B)	(IV)	(IS)	(N)		
Sensitivity	0.88	0.91	0.81	0.90	0.83	0.92	0.81	0.88	0.80	0.90	0.85	0.90		
Specificity	0.88	0.90	0.82	0.90	0.83	0.92	0.81	0.88	0.79	0.89	0.85	0.90		
Precision	0.85	0.82	0.93	0.89	0.90	0.82	0.91	0.87	0.88	0.81	0.82	0.82		
Recall	0.88	0.91	0.81	0.90	0.83	0.92	0.81	0.88	0.80	0.90	0.85	0.90		
F-score	0.88	0.86	0.87	0.89	0.86	0.85	0.86	0.87	0.80	0.86	0.84	0.84		
Overall test accuracy				0.90	Overall test accuracy				0.88	Overall test accuracy				0.85

Bold values represent the highest result achieved by our proposed model in different dataset magnification level

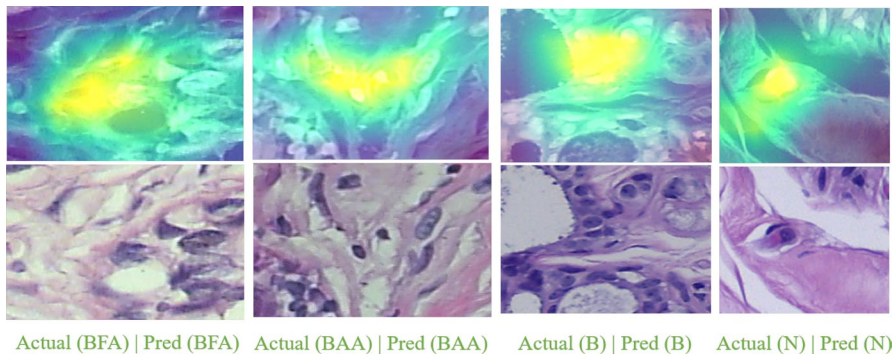


Fig. 10 Grad-CAM visual explanations about NETGL decision; *BreakHis*; **A** 400× B-FA (EV) **B** 200× B-AA (VE). *Bach Challenge*; **C** 1000^{w×h} B (VE), **D** (1000)^{w×h} N

Table 13 GLNET model comparison details with pre-trained base model (*BreakHis*)

Models	Pathology slide magnification levels			
	40× (%)	100× (%)	200× (%)	400× (%)
DenseNet201	54.00	70.00	65.00	70.00
ResNet101	80.00	83.00	75.00	74.00
VGG19	73.00	72.00	60.00	59.00
VGG16	74.00	75.00	57.00	58.00
GLNET informed model	90.01	92.32	91.98	91.51

enhanced transparency in this regard. To demonstrate the efficacy of GLNET, we leverage the *GRAD-CAM* technique to generate visual explanations for the classification of multi-label cancer categories, facilitating a more transparent validation of the model's decisions. Figure 10 illustrates how *GRAD-CAM* effectively highlights specific regions of interest within histopathological slides, corresponding to categories such as *B*, *N*, *B-FA*, and *B-AA*. These conspicuous regions, generated by *GRAD-CAM*, provide a coarse map localization that accentuates the reasoning behind our model's predictions for each cancer category, offering valuable *VE*-visual explanations.

4.3.1.3 GLNET model comparison details with pre-trained base model The superior performance of the GLNET model becomes evident upon closer examination of Tables 13 and 14, which present its consistently higher accuracy scores across all magnification levels and downsized images compared to the various pre-trained models listed. It is important to note that each pre-trained model possesses its own set of strengths and weaknesses in accuracy. For instance, VGG16 achieves the highest accuracy score of 75% at the 100× magnification level, while ResNet101 demonstrates the highest accuracy of 76.86% at a resolution of 1000×1000 pixels. However, it is noteworthy that these accuracy values, while commendable, fall

Table 14 GLNET model comparison details with pre-trained base model (*Bach Challenge*)

Models	Pathology slide WSI image sizes		
	1000× (%)	800× (%)	600× (%)
DenseNet201	45.61	42.42	45.65
ResNet101	76.86	63.63	76.08
VGG19	72.41	48.48	65.21
VGG16	65.51	72.72	72.82
GLNET	90.27	88.07	85.11

short of the performance achieved by the proposed GLNET model. The GLNET model consistently attains accuracy rates ranging from 90.01% to approximately 92.32% for the *BreakHis* dataset and 85.11% to around 90.02% for the *Bach Challenge* dataset. This remarkable accuracy underscores the GLNET model's efficacy in effectively distinguishing between the eight categories of breast cancer in the *BreakHis* dataset and the four categories of breast cancer in the *Bach Challenge* dataset, especially when equipped with global–local features.

4.3.1.4 GLNET model comparison details with existing model Table 15 provides a comprehensive overview of our proposed model's exceptional performance in *BC* diagnosis, surpassing previous studies in the field. Our model's experimental outcomes reaffirm its reliability in distinguishing between multi-label categories using a fusion of *GL* features. Our model outperforms existing models, yielding accuracy rates of 90.01%, 92.32%, 91.98%, and 91.51% for pathological images at magnification levels of 40×, 100×, 200×, and 400×, respectively. This substantial improvement

Table 15 GLNET model comparison details with existing model (*BreakHis*)

Reference	Model	Pathology slide magnification levels			
		40× (%)	100× (%)	200× (%)	400× (%)
Baumara et al. [47]	VGG16 with SVM	88.95	83.71	80.64	76.77
Bardou et al. [48]	CNN	86.34	84.00	79.93	79.74
	CNN With Augmentation	83.79	84.48	80.83	81.03
	SVM	82.89	80.94	79.44	77.94
	Ensemble CNN Model	88.23	84.64	83.31	83.98
	Bo W/D SFIT	66.72	69.06	62.42	52.75
	Bo With SURF	41.80	38.56	49.75	38.67
	LLC/DSIFT	60.58	57.44	70.00	46.96
	LLC/SIFT	80.37	63.84	74.54	54.7
Spanhol et al. [49]	CNN with Fusion Rules	89.00	85.00	84.00	80.00
Jain et al. [50]	Residual Learning + CNN	87.40	87.26	91.15	86.27
Kumar et al. [51]	Inception + Residual	80.8	82.76	86.55	85.80
Proposed model	GLNET Informed Model	90.01	92.32	91.98	91.51

Table 16 GLNET model comparison details with existing model (*Bach Challenge*)

Reference	Model	Accuracy (%)
Laxmisagar et al. [52]	MobileNet2.10ex	88.92
Aresta et al. [53]	CNN designs	87.00
	Resnet-101; Densenet-161	87.00
	ResNet, Inception-V3, Random Forests	79.00
	Custom CNN (Capsule Network)	72.00
	ResNet50, SVM	71.00
Golatkar et al. [54]	InceptionV3	79.00
Yan et al. [55]	hybrid convolutional and recurrent deep neural network	82.10
Proposed model	GLNET informed model	90.27

in accuracy can be attributed to the integration of *GL* features into our *CNN*-based GLNET model, setting it apart from *SOA*-state-of-the-art convolutional classifiers as represented by [4, 5, 24], and [3]. Table 16 further demonstrates the superior performance of our proposed model compared to *SOA* classifiers, maintaining an impressive accuracy rate of 90.27%. Importantly, our model excels in classifying four distinct categories of breast cancer in *Bach Challenge* histopathological slides. These findings underscore the effectiveness of our model in advancing *BC* diagnosis, signifying its potential as a valuable tool in the medical domain.

In Table 17, our proposed model demonstrates superior performance in the context of model complexity analysis compared to other pre-trained models. Specifically, our suggested network surpasses standard pre-trained models, underlining the efficacy of our approach. *CNN*-based algorithms generally outperform conventional techniques due to their inherent ability to learn and capture intricate features autonomously. In our case, employing *TL* with ResNet101, coupled with the proposed Convolution block, which exhibits a slightly larger scale than pre-trained models, yields the most favorable outcomes. The distinctiveness of the features extracted by this specified architecture is important in enhancing accuracy and balance between a model's expressive capacity and the risk of overfitting is crucial. A complex network often needs to be more comprehensive in grasping the nuances within the data, resulting in suboptimal accuracy. Conversely, an excessively complex network tends to be challenging to train and is prone to rapid overfitting, leading to diminished

Table 17 GLNET model complexity comparison with pre-trained model

Model	Total parameter (M)	Model size (MB)
DenseNet201	20.2	80
ResNet101	44.7	171
VGG19	143.7	549
VGG16	138.4	528
GLNET informed model	58.7	231

accuracy. Therefore, an optimal model size, in combination with other effective overfitting mitigation techniques like appropriate dropout rates and adept data augmentation, is essential to yield the best results.

Our proposed ResNet101 model incorporates five convolution layers, two *GAP* layers, four batch normalization layers, and two dense layers with varying filter sizes (256 and 64). Augmented by a dropout layer with a rate of 0.2 and followed by flattening, our model demonstrates the capacity to glean valuable insights from the data effectively. Combined with two objective learnings, this comprehensive architecture is highly accurate for our proposed ResNet101 [43] model.

4.3.1.5 Ablation studies Through our comprehensive ablation studies, we have substantiated the robustness and effectiveness of our novel GLNET model. These studies involved a detailed analysis of the parameters, focusing on the impact of both global and local features, aimed at establish the reliability and significance of these parameters.

For our in-depth investigation, we utilized the complete *BreakHis* dataset, comprising four magnification levels of slides and downsized images, in combination with our Global-Net and Local-Net architectures. The results, as presented in Tables 17, 18, 19, and 20, indicate the constant performance of both our architectural approaches, surpassing the *SOA* models [2, 4]. Table 17 highlights the exceptional achievement of the GLNET model, with the highest test accuracy recorded at

Table 18 NETG Global Feature Extractor (*GFE*)—accuracy *BreakHis*

Category	Pathology slide magnification levels			
	40×	100×	200×	400×
Total WSI	1009	1114	1122	1049
Train accuracy	88.43	86.87	89.10	89.01
Test accuracy	87.88	88.10	89.88	90.01

Table 19 NETL Local Feature Extractor (*LFE*)—accuracy *BreakHis*

Details	Pathology slide magnification levels			
	40×	100×	200×	400×
Total <i>Patch</i>	12,108	13,368	13,464	12,588
Train accuracy	87.13	88.97	90.01	90.56
Test accuracy	86.88	90.12	88.88	91.11

Table 20 NETG Global Feature Extractor (*GFE*)—accuracy *BachChallenge*

Category	Pathology slide WSI image sizes		
	1000×	800×	600×
<i>Malignant and Benign</i>			
Total WSI	587	651	1833
Train accuracy	86.78	84.33	85.10
Test accuracy	88.91	85.17	83.80

90.01% in the 400× magnification level, outperforming other levels, and an accuracy of 88.91% in (1000) $W \times H$ images when utilizing the global feature, as indicated in Table 19. It is important to note that in the NETG model, we exclusively employed feature extraction from the entire slide, encompassing attributes such as color, tissue size, and shape. Subsequently, we integrated these pool feature vectors into our proposed CNN model. The outcomes clarify in Tables 17 and 19 indicating the model's capacity to maintain remarkable accuracy across various magnification levels and downsized images.

Similar to the findings presented in Tables 18 and 20, the results emphasize the significance of specific parameters in optimizing model accuracy. Specifically, the highest accuracy, reaching 86.13 (1000) $W \times H$ and 91.11 (400× Magnification Level), is achieved by local features within the proposed CNNs model. Remarkably, this parameter modification does not compromise the model's efficacy. This observation is reinforced through ablation studies, indicating that both global and local features play important roles in enhancing the performance of the proposed model when compared to *SOA-CNN* models [2, 4, 6, 33]. These results highlight the critical influence of feature engineering and selection on model effectiveness and underline the value of these features in achieving superior performance (Table 21).

5 Strength and future work

In this study, the proposed model's strengths lie in its adept handling of multiclass challenges, innovative feature extraction techniques, sophisticated dense feature extractor model, informed feature integration, and ability to perform multi-label cancer category classification effectively. These attributes set it apart from existing methods and contribute to its performance in image classification tasks. The following section provides a comprehensive overview of the advantages offered by our proposed model compared to existing research in this field.

- Multiclass problem handling** The model effectively tackles the complexities of multiclass problems, offering enhanced classification of attributes across diverse object categories. This is particularly significant in scenarios where category vulnerabilities challenge accurate classification. By addressing these challenges, the proposed model contributes to improved classification performance.

Table 21 NETG Local Feature Extractor (LFE)—accuracy *BachChallenge*

Category	Pathology slide WSI image sizes		
	1000×	800×	600×
<i>Malignant and Benign</i>			
Total patch 250×	7754	5251	6536
Train accuracy	87.36	83.11	83.78
Test accuracy	86.13	85.81	84.19

- **Innovative feature extraction** Unlike traditional feature extraction methods, the proposed model adopts a unique approach by employing two distinct feature extraction techniques with specific purposes. The global feature extractor (W_{GF}) captures holistic, global information from whole slide images, encompassing attributes like color, shape, and tissue size. Simultaneously, the local feature extractor (W_{LF}) focuses on extracting features related to the object of interest within the image. This dual-feature extraction strategy allows for a more comprehensive understanding of the image content.
- **Dense feature extractor model** The model introduces a novel dense feature extractor CNNs model featuring five convolutional layers with two filter sizes: batch normalization layers, GAP layers, dense layers, and dropout layers. This sophisticated architecture enhances feature extraction and representation, improving the model's accuracy and robustness.
- **Informed feature integration** The proposed model incorporates an informed feature intention technique to determine optimal weights for each cancer category. By merging global–local feature W_{GL} (color, shape, interest region), the model fine-tunes its layers to align with the characteristics of different cancer categories. This adaptive approach ensures the model can effectively differentiate between diverse cancer features, enhancing classification accuracy.
- **Multi-label classification** The GLNET model excels in multi-label cancer category classification, providing granular insights into the attributes of each category. This is a significant contribution, as it recognizes that different cancer categories exhibit distinct features, allowing for more accurate and nuanced classification.

For future work, integrating temporal data or dynamic features into the classification process can offer a more comprehensive insight into histopathological cellular behavior. This augmentation holds the promise of significantly enhancing the diagnostic capabilities of the model.

6 Conclusion

The GLNET model, our proposed global and local features extractor, exhibits superior performance across two benchmark datasets such as *BreakHis* and the *ICAR2018-BachChallenge*, showcasing its efficacy in handling multi-label category taxonomies and attributes, surpassing previous studies [2, 4, 6, 8]. Through comprehensive model evaluation, it consistently outperforms *SOA* models due to integrating *GL* feature extraction methodologies. Both feature extractors are customized to address specific goals, ultimately enhancing accuracy. The model effectively fulfills its two-stage learning objectives.

Furthermore, our proposed model effectively tackles challenges associated with multiclass problems, including classifying attributes spanning multiple object categories, addressing category vulnerabilities, and capturing features in a single step. It accomplishes this by employing features aligned with the two-stage learning objectives, encompassing *WSI*-whole slide features and *ROI*-region of interest areas.

A noteworthy innovation is the dense fine-tuning process, where combined features from *WGL* fuse *WSI* to *patch 3×4 slides'* categories, along with shape and Pool Feature Map integration. This augmentation significantly boosts the model's predictive capabilities for cancer categories. The *GL-NET* module, rooted in informed *CNNs*, facilitates a more profound understanding of slide features within *WSI* to *patch 3×4*.

The GLNET model exhibits remarkable proficiency in multi-label cancer category classification, achieving impressive accuracy rates of 90.01% at 40× magnification, 92.32% at 100×, 91.98% at 200×, and 91.51% at 400×. Moreover, the model offers detailed insights into the distinctive attributes of each cancer category, acknowledging inherent variations and enabling a more precise and nuanced classification process. This achievement represents a significant leap forward, offering enhanced capabilities for classifying diverse cancer categories.

Acknowledgements This work was supported by the Natural Science Foundation of Hunan Province, China (Grant No. 2020JJ4757). This work is supported by the Intelligent annotation and fine-grained recognition of large-scale multimodal medical behavior belong to 2030 Innovation Megaprojects (to be fully launched by 2020)-New Generation Artificial Intelligence (Project No. 2020AAA0109600). This work is funded by the National Key R&D Program of China Under Grant 2021ZD0140301, the National Natural Science Foundation of China under Project No. 61902433, and the High-Performance Computing Center of Central South University.

Author contributions Saif Ur Rehman Khan was involved in Methodology, Formal analysis, Validation, and Writing—original draft. Ming Zhao contributed to Conceptualization, Formal analysis, Supervision, and Writing—review & editing. Sohaib Asif was involved in Conceptualization, Methodology, Formal analysis, and Writing—review & editing. Xuehan Chen contributed to Conceptualization, Formal analysis, Supervision, and Writing—review & editing. Yusen Zhu was involved in Formal analysis and Validation. All authors reviewed the manuscript.

Funding No funding.

Data availability Dataset is publicly available.

Declarations

Conflict of interest There is no conflict of interest between authors.

Ethics approval Approved.

References

1. American cancer society.about breast cancer.org—1.800.227.2345. <https://www.cancer.org/content/dam/CRC/PDF/Public/8577.00.pdf>
2. Deniz ES, Engür A, Kadiroglu Z, Guo Y, Bajaj V, Budak Ü (2018) Transfer learning based histopathologic image classification for breast cancer detection. *Health Inf Sci Syst* 6(1):18
3. Das K, Conjeti S, Roy AG, Chatterjee J, Sheet D (2018) Multiple instance learning of deep convolutional neural networks for breast histopathology whole slide classification. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE, pp 578–581
4. Li L et al (2020) Multi-task deep learning for fine-grained classification and grading in breast cancer histopathological images. *Multimed. Tools Appl.* 79(21):14509–14528

5. Ahmad N, Asghar S, Gillani SA (2022) Transfer learning-assisted multi-resolution breast cancer histopathological images classification. *Vis. Comput.* 38(8):2751–2770
6. Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mouggiakakou S (2016) Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Trans. Med. Imaging* 35(5):1207–1216
7. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 770–778
8. Howard A, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H (2017) Mobilenets: efficient convolutional neural networks for mobile vision applications. *CoRR*, [arXiv:1704.04861](https://arxiv.org/abs/1704.04861)
9. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In: *International Conference on Learning Representations*
10. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 1–9
11. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 2818–2826
12. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen L-C (2018) Mobilenetv2: inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 4510–4520
13. Tajbakhsh N, Shin JY, Gurudu SR, Todd Hurst R, Kendall CB, Gotway MB, Liang J (2016) Convolutional neural networks for medical image analysis: full training or fine tuning? *IEEE Trans Med Imaging* 35(5):1299–1312
14. Mehra R et al (2018) Breast cancer histology images classification: training from scratch or transfer learning? *ICT Exp* 4(4):247–254
15. Kablan EB, Dogan H, Ercin ME, Ersoz S, Ekinici M (2020) An ensemble of fine-tuned fully convolutional neural networks for pleural effusion cell nuclei segmentation. *Comput Electr Eng* 81:106533
16. Akhtar Z, Foresti GL (2016) Face spoof attack recognition using discriminative image patches. *J Electr Comput Eng* 2016:66
17. Khan S, Islam N, Jan Z, Din IU, Rodrigues JJC (2019) A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recogn Lett* 125:1–6
18. Chan A, Tuszynski JA (2016) Automatic prediction of tumour malignancy in breast cancer with fractal dimension. *R Soc Open Sci* 3(12):160558
19. Nawaz MA, Sewissy AA, Soliman THA (2018) Automated classification of breast cancer histology images using deep learning based convolutional neural networks. *Int J Comput Sci Netw Secur* 4:152–160
20. Mormont R, Geurts P, Maree R (2020) Multi-task pre-training of deep neural networks for digital pathology. *IEEE J Biomed Health Inform* 6:66
21. Medela A, Picon A, Saratxaga CL, Belar O, Cabezon V, Cicchi R, Bilbao R, Glover B (2019) Few shot learning in histopathological images: reducing the need of labeled data on biological datasets. In: *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, pp 1860–1864
22. Samah AA, Fauzi MFA, Mansor S (2017) Classification of benign and malignant tumors in histopathology images. In: *2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*. IEEE, pp 102–106
23. Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2015) A dataset for breast cancer histopathological image classification. *IEEE Trans Biomed Eng* 63(7):1455–1462
24. Kahya MA, Al-Hayani W, Algamal ZY (2017) Classification of breast cancer histopathology images based on adaptive sparse support vector machine. *J Appl Math Bioinform* 7(1):49
25. Sanchez-Morillo D, González J, García-Rojo M, Ortega J (2018) Classification of breast cancer histopathological images using kaze features. In: *International Conference on Bioinformatics and Biomedical Engineering*. Springer, Berlin, pp 276–286
26. Nahid A-A, Mehrabi MA, Kong Y (20185) Histopathological breast cancer image classification by deep neural network techniques guided by local clustering. *BioMed Res Int* 6, 66
27. Jiang Y, Chen L, Zhang H, Xiao X (2019) Breast cancer histopathological image classification using convolutional neural networks with small se-resnet module. *PLoS ONE* 14(3):e0214587

28. Nejad EM, Affendey LS, Latip RB, Ishak IB (2017) Classification of histopathology images of breast into benign and malignant using a single-layer convolutional neural network. In: Proceedings of the International Conference on Imaging, Signal Processing and Communication, pp 50–53
29. Kumar K, Rao ACS (2018) Breast cancer classification of image using convolutional neural network. In: 2018 4th International Conference on Recent Advances in Information Technology (RAIT). IEEE, pp 1–6
30. Sun J, Binder A (2017) Comparison of deep learning architectures for H&E histopathology images. In: 2017 IEEE Conference on Big Data and Analytics (ICBDA). IEEE, pp 43–48
31. Benhammou Y, Tabik S, Achchab B, Herrera F (2018) A first study exploring the performance of the state-of-the-art cnn model in the problem of breast cancer. In: Proceedings of the International Conference on Learning and Optimization Algorithms: Theory and Applications, pp 1–6
32. Sharma S, Mehra R (2020) Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images-a comparative insight. *J Digit Imaging* 33(3):632–654
33. Bakkouri I et al (2022) BG-3DM2F: Bidirectional gated 3D multi-scale feature fusion for Alzheimer's disease diagnosis. *Multimed Tools Appl* 81(8):10743–10776
34. Bakkouri I, Afdel K (2023) MLCA2F: multi-level context attentional feature fusion for COVID-19 lesion segmentation from CT scans. *Signal Image Video Process* 17(4):1181–1188
35. Alom MZ et al (2019) Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. *J Digit Imaging* 32:605–617
36. Bhatt D et al (2021) CNN variants for computer vision: history, architecture, application, challenges and future scope. *Electronics* 10(20):2470
37. Patel C et al (2022) DBGCC: dimension-based generic convolution block for object recognition. *Sensors* 22(5):1780
38. Guo Y et al (2020) DeepANF: a deep attentive neural framework with distributed representation for chromatin accessibility prediction. *Neurocomputing* 379:305–318
39. Singh J et al (2019) RNA secondary structure prediction using an ensemble of two-dimensional deep neural networks and transfer learning. *Nat Commun* 10(1):1–13
40. Guo Y et al (2019) DeepACLSTM: deep asymmetric convolutional long short-term memory neural models for protein secondary structure prediction. *BMC Bioinform* 20(1):1–12
41. Gao Q, Lim S, Jia X (2019) Spectral–spatial hyperspectral image classification using a multiscale conservative smoothing scheme and adaptive sparse representation. *IEEE Trans Geosci Remote Sens* 57(10):7718–7730
42. Macenko M et al (2009) A method for normalizing histology slides for quantitative analysis. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE
43. Metwaly K et al (2022) Glidenet: global, local and intrinsic based dense embedding network for multi-category attributes prediction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition
44. Sergey I, Christian S (202s1) Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv 2015. arXiv preprint [arXiv:1502.03167](https://arxiv.org/abs/1502.03167)
45. Kassani SH et al (2019) Breast cancer diagnosis with transfer learning and global pooling. In: 2019 International Conference on Information and Communication Technology Convergence (ICTC). IEEE
46. Al-Ameen Z, Muttar A, Al-Badrani G (2019) Improving the sharpness of digital image using an amended unsharp mask filter. *Int J Image Graph Signal Process* 11(3):66
47. Boumaraf S et al (2021) Conventional machine learning versus deep learning for magnification dependent histopathological breast cancer image classification: a comparative study with visual explanation. *Diagnostics* 11(3):528
48. Bardou D, Zhang K, Ahmad SM (2018) Classification of breast cancer based on histology images using convolutional neural networks. *IEEE Access* 6:24680–24693
49. Spanhol FA et al (2016) Breast cancer histopathological image classification using convolutional neural networks. In: 2016 International Joint Conference on Neural Networks (IJCNN). IEEE
50. Gour M, Jain S, Kumar TS (2020) Residual learning-based CNN for breast cancer histopathological image classification. *Int J Imaging Syst Technol* 30(3):621–635
51. Singh S, Kumar R (2022) Breast cancer detection from histopathology images with deep inception and residual blocks. *Multimed Tools Appl* 81(4):5849–5865

52. Laxmisagar HS., Hanumantharaju MC (2021) Design of an efficient deep neural network for multi-level classification of breast cancer histology images. In: *Intelligent Computing and Applications*. Springer, Singapore, pp 447–459
53. Aresta G et al (2019) Bach: grand challenge on breast cancer histology images. *Med Image Anal* 56:122–139
54. Golatkar A, Anand D, Sethi A (2018) Classification of breast cancer histology using deep learning. In: *International Conference Image Analysis and Recognition*, pp 837–844. https://doi.org/10.1007/978-3-319-93000-8_95
55. Yan R et al (2020) Breast cancer histopathological image classification using a hybrid deep neural network. *Methods* 173:52–60. <https://doi.org/10.1016/j.ymeth.2019.06.014>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Saif Ur Rehman Khan¹ · Ming Zhao¹ · Sohaib Asif¹ · Xuehan Chen¹ · Yusen Zhu²

✉ Ming Zhao
meanzhao@csu.edu.cn

Saif Ur Rehman Khan
saifurrehman.khan@csu.edu.cn

Sohaib Asif
punjabians1592@gmail.com

Xuehan Chen
xuehan2013@csu.edu.cn

Yusen Zhu
zhu_yusen@163.com

¹ School of Computer Science and Engineering, Central South University, Changsha, China

² School of Mathematics, Hunan University, Changsha, China