Transactions on Materials, Biotechnology and Life Sciences
ISSN: 2960-1932, eISSN: 2960-2319 | Volume 7
MISBP 2024

Breast Cancer Image Classification Based on Attention Mechanisms

Yuyuan Guo[#, 1, *], Sinuo Duan[#, 1], Fangqi Wang[#, 2]

[1]School of Medicine, Shihezi University, Shihezi, China, 832000

[2]School of Life Sciences, Shanghai University, Shanghai, China, 200444

* Corresponding Author Email: gyy030215@163.com

#These authors contributed equally.

Abstract. Convolutional neural networks have been frequently utilized in computer-aided diagnosis (CAD). Breast cancer image classification is one of the vital applications of CAD. The purpose of this study was to explore the role of attention mechanism in breast cancer image classification. Specifically, this study introduces attention mechanism into the classical image classification deep learning network and constructs a new breast cancer image classification model. Test results indicate that convolutional neural networks perform better in classification when an attention mechanism is added, and they also perform better in terms of training loss and accuracy.

Keywords: Deep learning; Image Classification; Attention Mechanisms.

1. Introduction

Data on the worldwide cancer burden, published in 2022 by the World Health Organization's International Agency for Research on Cancer, indicates that the number of instances of breast cancer has surpassed 2.3 million, making it the most common cancer among women and endangering the lives of countless more. At present, breast cancer is mainly treated by surgical removal of the breast. The premise of surgical treatment and good prognosis is timely detection and early treatment of breast cancer. Consequently, early identification of breast cancer is critical and imperative, raising the bar for breast cancer diagnosis as well. Currently, breast ultrasonography, magnetic resonance imaging, and mammography are typically utilized in the clinical diagnosis of breast cancer. These techniques provide powerful tools for diagnosing breast cancer. For instance, pathological examination of breast biopsies is the gold standard for diagnosis [1] and gives medical professionals reliable information for making diagnoses.

Breast cancer has a variety of pathological forms and types, and breast tumors can be divided into benign tumors and malignant tumors [2]. Benign tumors can be classified into four subtypes: tubular adenoma (TA), fibroadenoma (F), adenosis (A), and phyllodes tumor (PT). There are four subtypes of malignant tumors: papillary carcinoma (PC), ductal carcinoma (DC), lobular carcinoma (LC), and mucinous carcinoma (MC).

With the growth of computer technology and the establishment and popularization of large-scale medical image data sets, computer-aided diagnosis (CAD) technology based on artificial intelligence has developed rapidly. Artificial neural networks and machine learning algorithms were the foundation of early artificial intelligence CAD technologies [3]. Using the computational analysis and learning capabilities of computer systems, the application of machine learning in CAD provides pathologists with a reliable reference for early diagnosis, further reducing the pressure on pathologists in reading diagnosis.

Deep learning has experienced rapid progress in recent years, and the use of convolutional neural networks (CNNs), a typical deep learning technique, in the field of computer-aided design (CAD) is crucial. Owing to CNN's superior extraction of features capabilities, the integration of CAD technology with CNN enhances the precision of pathological diagnosis of image and significantly lowers the rate of missing and incorrect diagnoses. There are some examples where excellent CNN

methods easily outperform human experts in providing diagnostic results. As a result, deep learning applications in pathological tissue pictures are crucial for early breast cancer prevention, diagnosis, therapy, and prognosis.

Therefore, based on the strong feature extraction ability and learning analysis ability of convolutional neural network in deep learning, this paper deeply studies the classification method of breast cancer pathological images in order to obtain more accurate classification results, which is of great significance to further relieve the doctor's reading pressure and reduce the misdiagnosis rate, and has a good application prospect.

2. Related word

2.1. Breast cancer classification

For cancer auto-aided diagnosis applications, the main objective of image classification systems is frequently to classify histopathological pictures into distinct histopathological patterns that correlate to the non-cancerous or malignant condition of the tissue being investigated. [4] A method for diagnosing breast cancer using fine needle biopsy cytological analysis of images was presented by Filipczuk et al. [5]. This has the ability to identify benign from malignant photos. Using 25-dimensional vectors of features, they trained four distinct classifiers and reported 98 percent accuracy on 737 photos.

2.2. Convolutional neural networks

CNN are deep learning algorithms inspired by the visual cortex of animal brains and designed to mimic animal visual machines [6]. CNN receives the image directly, processes it through fully connected, pooling, and convolutional layers, and then outputs class-based probability estimates of the image [7].

In medical image analysis, CNN is a popular deep learning architecture that is used for image classification [8], segmentation [9], detection [10], and other tasks. Starting with AlexNet[11], various end-to-end models have been developed with deeper and deeper networks and greater representation compactness for image classification, such as VGG[12], ResNet[13], and DenseNet[14]. Deep learning is being utilized extensively in medical image processing because to the exceptional outcomes these models have delivered.

2.3. The attention mechanism

Neural networks use the attention mechanism as a resource allocation technique to direct computing power toward regions that are deemed more critical.

According to input picture attributes, the attention mechanism in computer vision can be thought of as a dynamic weight adjustment process [15].

A low-cost, low-complexity channel attention mechanism module called SENet was proposed by Hu et al. [16]. It has the ability to adaptively calibrate channel level characteristic response in order to increase accuracy. In order to accomplish spatial alignment feature mapping, Roy et al [17]. developed a spatial attention mechanism for spatial feature compression and excitation. In general, a lot of visual tasks heavily rely on attention.

3. The basic fundamental of the method

3.1. The structure of ResNet

Deep residual network was proposed by He Kaiming and other four scholars from Microsoft Research Institute in 2016. The network solves the gradient disappearance problem that occurs when the neural network deepens. ResNet introduces residual structure to realize identity mapping, so that the network

can always learn the characteristics of shallow network, which ensures that the network performance will not degrade with the increase of the number of layers. A common ResNet model structure is shown in Figure 1. For low-level networks, convolution blocks are stacked with two layers, and for deeper networks, three-layer bottleneck convolution structures are used. The residual structure blocks of the core in each residual network are shown in Figure 2 and Figure 3. The network fits a residual mapping F(x), where the expected mapping H(x)=F(x)+x, rather than the expected mapping H(x) of the data, as a result of the author's short-cut connection in the residual block. In order to prevent a decline in network performance, the stacking layer is equivalent to performing identity mapping when the residual is 0. But since the residual error won't be zero, the network will be able to learn new features from the input features, which translates to improved performance.

layer_name	ouput_size	18-layer	34-layer	50-layer	101-layer	152-layer
Conv1_x	112×112	7×7, 64, stride2				
Conv2_x	56×56	3×3 max_pool, stride2				
		$\begin{bmatrix}3\times3, & 64\\ 3\times3, & 64\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, & 64\\ 3\times3, & 64\end{bmatrix}\times2$	$\begin{bmatrix}1\times1, & 64\\ 3\times3, & 64\\ 1\times1, & 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 64\\ 3\times3, & 64\\ 1\times1, & 256\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 64\\ 3\times3, & 64\\ 1\times1, & 256\end{bmatrix}\times3$
Conv3_x	28×28	$\begin{bmatrix}3\times3, & 128\\ 3\times3, & 128\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, & 128\\ 3\times3, & 128\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, & 128\\ 3\times3, & 128\\ 1\times1, & 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, & 128\\ 3\times3, & 128\\ 1\times1, & 512\end{bmatrix}\times4$	$\begin{bmatrix}1\times1, & 128\\ 3\times3, & 128\\ 1\times1, & 512\end{bmatrix}\times8$
Conv4_x	14×14	$\begin{bmatrix}3\times3, & 256\\ 3\times3, & 256\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, & 256\\ 3\times3, & 256\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, & 256\\ 3\times3, & 256\\ 1\times1, & 1024\end{bmatrix}\times6$	$\begin{bmatrix}1\times1, & 256\\ 3\times3, & 256\\ 1\times1, & 1024\end{bmatrix}\times23$	$\begin{bmatrix}1\times1, & 256\\ 3\times3, & 256\\ 1\times1, & 1024\end{bmatrix}\times36$
Conv5_x	7×7	$\begin{bmatrix}3\times3, & 512\\ 3\times3, & 512\end{bmatrix}\times2$	$\begin{bmatrix}3\times3, & 512\\ 3\times3, & 512\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 512\\ 3\times3, & 512\\ 1\times1, & 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 512\\ 3\times3, & 512\\ 1\times1, & 2048\end{bmatrix}\times3$	$\begin{bmatrix}1\times1, & 512\\ 3\times3, & 512\\ 1\times1, & 2048\end{bmatrix}\times3$
	1×1	average pool, 1000dTc, softmax				
FLOPs		1.8×10e9	3.6×10e9	3.8×10e9	7.6×10e9	11.3×10e9

Figure 1. Residual Network model structure

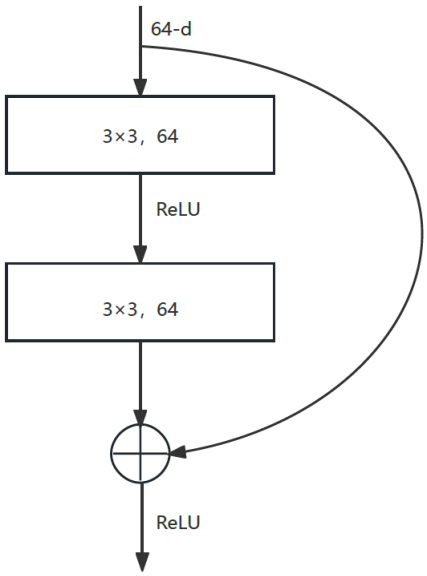

Figure 2. Shallow network residuals block

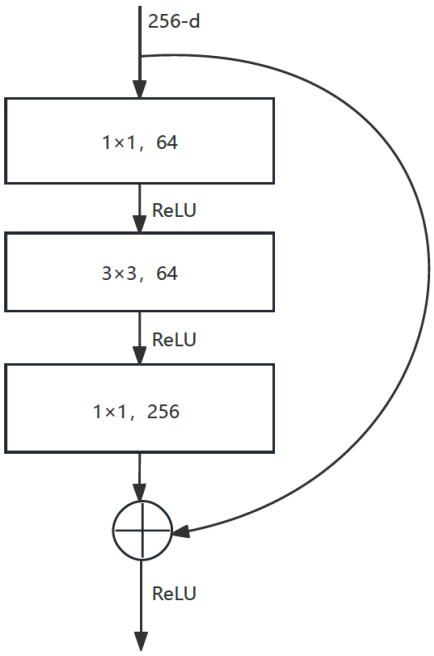

Figure 3. Deep network residuals block

Considering the powerful capability of resnet, we use resnet as a baseline network in this paper to carry out research on breast cancer image classification.

3.2. Attention mechanisms method

The attention mechanism in deep neural networks mimics the human visual attention mechanism by allocating resources to focus on critical information, utilizing strategies such as self-attention, spatial attention, and channel attention in computer vision tasks.

For this reason, Hu et al. [16] originally introduced the idea of channel attention and suggested SENet (shown in Figure 4). The SE block, which is utilized to better convey information and record channel

relationships, is the central component of SENet. Squeeze and Excitation modules make up the SE block. First, global average pooling is used in the Squeeze module to compress global information. Secondly, the Excitation module uses fully connected layers and activation functions to capture channel relationships and output attention vectors. Lastly, the matching element in the attention vector is multiplied by each channel of the input feature to scale it. SE block captures channel information while suppressing noise, emphasizes key information contained in channel, and helps to control global information by network model. However, SE blocks are not perfect. First, the Squeeze module directly compresses global information by employing the global average pool, which may result in the loss of some complicated global information. Second, the fully linked layer employed in the Excitation module has a large amount of parameter information, which raises the computational complexity of the model.

Woo et al. [18] introduced CBAM (as shown in Figure 5), a method integrating channel attention and spatial attention, where the channel attention module employs maximum and average pooling to generate weight vectors for each channel, mapping them through a multilayer perceptron network and multiplying the resultant weights with the input feature map to enhance information channels and important regions. The formula of the channel attention module acting on the input characteristic graph X is:

$$XC = X\ Sigmoid(FC(Max\_pool(X)) + FC(Avg\_pool(X))) \tag{1}$$

where XC is the output characteristic graph of the input characteristic graph X after passing through the channel attention module. For the spatial attention module, the feature graph XC is maximally pooled and averagely pooled by channel to form two [1,H,W] weight vectors. The dimension of the feature graph is changed from [C,H,W] to [1,H,W], i.e. pooling all channels of the same feature point. Stacking the obtained two feature maps to form a feature map spatial weight of [2,H,W], and then changing the feature map dimension from [2,H,W] to [1,H,W] through a convolution layer. Then multiply the spatial weights [1,H,W] obtained by the feature map XC. The formula that the spatial attention module acts on the feature map XC is:

$$Y = XC\ Sigmoid(F3 \times 3([Max\_pool(XC); Avg\_pool(XC)]) \tag{2}$$

where the $F_{3\times3}$ convolution kernel size is a 3×3 two-dimensional convolution and Y is the final output feature map. CBAM uses channel attention and spatial attention in sequence to find more effective features, extract spatial features and channel features from input images. Compared with single attention, CBAM further narrows the attention range and clarifies the location of key feature areas.

In this study, we consider adding SE module and CBAM module to resnet to enhance the performance of breast cancer classification.

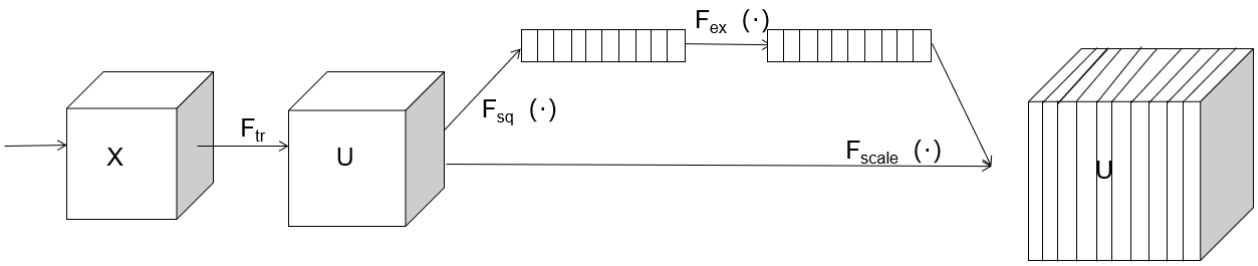

Figure 4. SE module

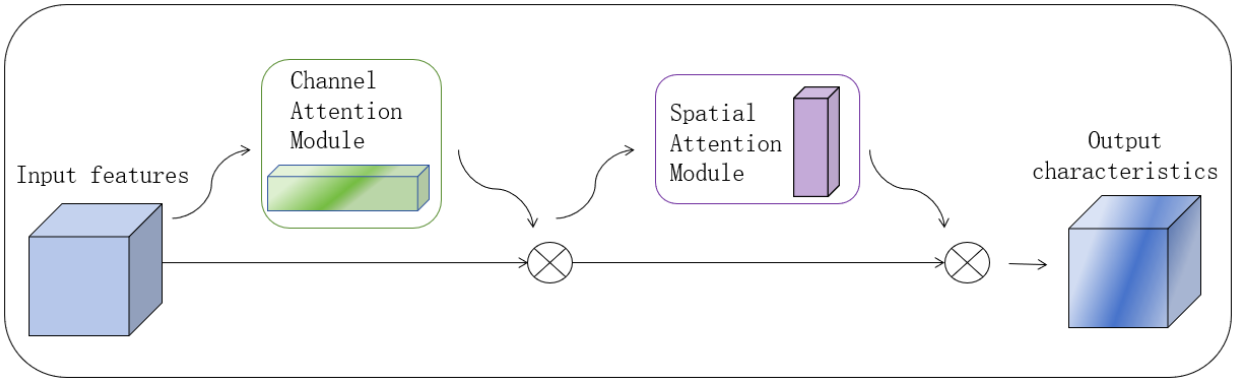

Figure 5. Channel-Spatial Attention Module

4. Results

4.1. The dataset

The Breakhis Dataset [19], which comes from the online image collection for breast pathology, is used in this work. This dataset, which includes digital pictures of histological sections, has been extensively utilized in machine learning-based breast cancer classification research in the past. Images from a clinical research were gathered for the Breakhis dataset between January and December of 2014. Together with the P &D Pathological Anatomy and Cytopathology Laboratory in Parana, Brazil, this dataset was produced [20]. This dataset includes 7909 microscopic images of breast tumor tissue obtained at various magnifications (40x, 100x, 200x, and 400x) from 82 patients. There are currently 5,429 malignant samples and 2480 benign samples in the PNG format (700*460 pixels, 3 channels RGB, 8 bits depth per channel).

The dataset currently contains four histologically distinct types of benign breast tumors: adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenoma (TA); and four malignant tumors (breast cancer): carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC), and squamous carcinoma (PC). [29] In the course of the experiment, we used all 7909 images from the training set and split them into test and training sets using the validation_split=0.1 ratio, which is roughly 9:1.

4.2. Analysis of experimental results

We used ResNet and attention based senet and CBAM neural networks to train and test the dataset. At the same time, we tested the three types of neural networks in 12 different scenarios based on the model size (20, 32, 56, 110). The results of the experiment are shown in Table 1.

Table 1. The results of the experiment

	loss	**accuracy**		**loss**	**accuracy**		**loss**	**accuracy**
resnet20	0.484	0.819	senet20	0.381	0.861	cbam20	0.414	0.861
resnet32	0.432	0.833	senet32	0.329	0.896	cbam32	0.414	0.861
resnet56	0.566	0.877	senet56	0.304	0.895	cbam56	0.458	0.851
resnet110	0.444	0.876	senet110	0.323	0.859	cbam110	**0.258**	**0.910**

Based on the experimental results, we can conclude that the cbam110 model has the lowest val_loss, approximately 0.2583, indicating the lowest cross entropy loss and the highest val-accuracy, approximately 0.9101, indicating the highest accuracy. Overall, senet and cbam with added attention mechanisms have smaller val_loss and higher val-accuracy compared to resnet without attention mechanisms, resulting in smaller cross entropy loss, higher accuracy, and better recognition performance.

5. Conclusions

In this paper, we focus on breast cancer image classification methods in integrated convolutional neural network, and use Breakhis dataset to study the task of breast cancer pathological image classification. In order to improve the accuracy of the study, we divide the three types of neural networks into 12 cases according to the size of the model (20,32,56,110). The results show that the size of the model has a positive correlation with the accuracy and cross-entropy loss of the neural network. At the same time, by introducing attention mechanism, the classification effect and classification performance of convolutional neural network can be further improved, so as to diagnose the patient's disease in time, so as to take treatment.

References

[1] Zaalouk AM,Ebrahim GA,Mohamed HK, et al. A Deep Learning Computer-Aided Diagnosis Approach for Breast Cancer. Bioengineering (Basel). 2022;9 (8):. doi:10.3390/bioengineering9080391

[2] Li J, Chen Z, Su K, Zeng J. Clinicopathological classification and traditional prognostic indicators of breast cancer. Int J Clin Exp Pathol. 2015 Jul 1;8(7):8500-5. PMID: 26339424; PMCID: PMC4555752.

[3] Francesco Tacchino,Chiara Macchiavello,Dario Gerace, et al. An artificial neuron implemented on an actual quantum processor. npj Quantum Information. 2019;5 (1):0-0. doi:10.1038/s41534-019-0140-4

[4] Jiang, Yun; Chen, Li; Zhang, Hai; Xiao, Xiao; Hu, Jianjun (2019). Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. PLOS ONE, 14(3), e0214587–. doi:10.1371/journal.pone.0214587 no. 10, pp. 1563–1572, 2013.

[5] P. Filipczuk, T. Fevens, A. Krzy˙zak, and R. Monczak, “Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies,” IEEE Transactions on Medical Imaging, vol. 32,no. 12, pp. 2169–2178, 2013.

[6] Sarvamangala DR, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intell. 2022;15(1):1-22. doi: 10.1007/s12065-020-00540-3. Epub 2021 Jan 3. PMID: 33425040; PMCID: PMC7778711.

[7] Chen X, Wang X, Zhang K, Fung KM, Thai TC, Moore K, Mannel RS, Liu H, Zheng B, Qiu Y. Recent advances and clinical applications of deep learning in medical image analysis. Med Image Anal. 2022 Jul;79:102444. doi: 10.1016/j.media.2022.102444. Epub 2022 Apr 4. PMID: 35472844; PMCID: PMC9156578.

[8] Zuo T, Zheng Y, He L, Chen T, Zheng B, Zheng S, You J, Li X, Liu R, Bai J, Si S, Wang Y, Zhang S, Wang L, Chen J. Automated Classification of Papillary Renal Cell Carcinoma and Chromophobe Renal Cell Carcinoma Based on a Small Computed Tomography Imaging Dataset Using Deep Learning. Front Oncol. 2021 Nov 18;11:746750. doi: 10.3389/fonc.2021.746750. PMID: 34868946; PMCID: PMC8637858.

[9] Zopes J, Platscher M, Paganucci S, Federau C. Multi-Modal Segmentation of 3D Brain Scans Using Neural Networks. Front Neurol. 2021 Jul 14;12:653375. doi: 10.3389/fneur.2021.653375. PMID: 34335436; PMCID: PMC8318570.

[10] Zhou X, Ma L, Mubarak HK, Little JV, Chen AY, Myers LL, Sumer BD, Fei B. Automatic detection of head and neck squamous cell carcinoma on pathologic slides using polarized hyperspectral imaging and deep learning. Proc SPIE Int Soc Opt Eng. 2022 Feb-Mar;12039:120390G. doi: 10.1117/12.2614624. Epub 2022 Apr 4. PMID: 36798940; PMCID: PMC9930132.

[11] Krizhevsky A , Sutskever I , Hinton G . ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in neural information processing systems, 2012, 25(2).

[12] Simonyan K , Zisserman A . Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Computer Science, 2014.

[13] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[14] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243.

[15] Meng-Hao Guo1, Tian-Xing Xu1, Jiang-Jiang Liu2.Attention mechanisms in computer vision: A survey..Computational Visual Media.2022;https://doi.org/10.1007/s12065-020-00540-3.

[16] HU J, SHEN L, SUN G． Squeeze－and－excitation networks[C]/ /Proceedings of the IEEE conference on computer vision and pattern recognition． 2018: 7132－7141.

[17] ROY A G, NAVAB N, WACHINGER C. Concurrent spatial and channel‘squeeze & excitation’in fully convolutional networks［C］/ /International conference on medical image computing and computer－assisted intervention. Springer, Cham, 2018: 421－429.

[18] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 3-19.

[19] Spanhol F A, Oliveira L S, Petitjean C, et al. A dataset for breast cancer histopathological image classification[J]. Ieee transactions on biomedical engineering, IEEE, 2015, 63(7): 1455-1462.

[20] Abdallah N, Marion JM, Tauber C, Carlier T, Hatt M, Chauvet P. Enhancing histopathological image classification of invasive ductal carcinoma using hybrid harmonization techniques. Sci Rep. 2023 Nov 16;13(1):20014