

Shearlet Transform and Convolutional Neural Network for Histopathology Images in Breast Cancer Classification

Siti Shaliza Mohd Khairi^{a,b}, Mohd Aftar Abu Bakar^{b*}, Mohd Almie Alias^b, Sakhinah Abu Bakar^b, Nurwahyuna Rosli^c, Mohsen Farid^d

^aFaculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia; Department of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia; ^cDepartment of Pathology, Faculty of Medicine, Hospital Canselor Tuanku Muhriz, Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Bandar Tun Razak, Cheras, 56000 Kuala Lumpur, Malaysia; ^dDepartment of Computing and Mathematics, University of Derby, Kedleston Road, Derby, DE22 1GB, United Kingdom

Abstract Breast cancer stands out as one of the global health threats, as it may cause death if improperly treated. Thus, detecting the illness at the early stage through precise diagnosis is important to prevent progression of tumors with effective treatments through medical imaging. Traditionally, manual diagnostic processes rely on the input data representation and expert knowledge, which consume much time and are prone to human error due to heavy workloads and fatigue. Recently, deep learning has shown distinguishing results in medical imaging analysis for image classification and detection. Nevertheless, the increasing demand to enhance the performance of image classification is becoming more prominent. In this study, a hybrid method of deep learning is proposed by combining Shearlet transform and convolutional neural network (CNN) for breast cancer histopathology image classification. First, the histopathology images are decomposed using Shearlet transform for Shearlet coefficients. Then, the CNN approach is used to classify the images into benign and malignant with minimal pre-processing procedure. The ability of Shearlet transform to address singularities helps to increase the quality of images. The proposed hybrid model improves the performance of the original basic CNN model. Results from the experiment show that the proposed hybrid model achieves an accuracy of 75%, an F1-score of 85% for malignant tumor, and a misclassification rate of 0.25%. This result shows that the use of Shearlet transform as the first feature extraction layer in the CNN architecture provides better feature extraction, consequently leading to improved accuracy for image classification.

Keywords: Shearlet transform, image classification, convolutional neural network, breast cancer.

***For correspondence:**
aftar@ukm.edu.my

Received: 04 Sept. 2024

Accepted: 25 June 2025

©Copyright Mohd Khairi.
This article is distributed
under the terms of the
**Creative Commons
Attribution License**, which
permits unrestricted use
and redistribution provided
that the original author and
source are credited.

Introduction

Breast cancer is one of the most common illnesses highlighted by researchers in medical field as it contributes to highest death rates [1]. This cancer predominantly occurs in women, hits them after puberty where the incidence rates rising steadily as people grow older. It is anticipated that the number of deaths from breast cancer will grow continuously from 2020 to 2030 for East and South Asian countries [2]. Breast cancer is not an airborne or spreading illness. The affected breast cells experience unregulated growth and can develop into benign or malignant tumors. Malignant tumors (cancerous tumors) can metastasize or spread to other parts of the body, conferring worse prognosis for the patients [3]. Early diagnosis and accurate treatment may help in decelerating the growth process and prevent spreading of uncontrolled abnormal cells to other parts of the body.

In healthcare, the availability of digital medical images such as ultrasound imaging, magnetic resonance imaging (MRI) and biopsy images have made research related to medical analysis rise dramatically over the past decade. Some factors that contributed to this increase includes technology advancement, accessible computing power and expansion of demand from physicians and patients [4, 5]. Due to the swift advancement in medical technologies, approaches for breast cancer diagnosis widely shift to histopathology images (biopsy) analysis but, the clinical research still depends on the subjectivity of the pathologists [6, 7]. Research on histopathology images is considered as the current gold standard procedure in cancer diagnosis because the results from this image analysis are able to help pathologists in making decision on tumors classification [8]. However, few issues may occur during the diagnosis process such as diagnostic limitation, ability and experience of the pathologists, and the complexity of histopathology images. Therefore, introducing efficient methods of classification for histopathological images is important for consistent and accurate diagnosis.

The literature has demonstrated breast cancer image classification through traditional methodologies, employing established methods in machine learning. Most of the machine learning methods proposed by previous research achieved promising accuracy for classification of medical images [9–12]. Traditional methods rely on the manual generation of low-level features which required expertise for image analysis and posing challenges in automatically handling the classification tasks [13, 14]. At present, a multi-layer neural network, known as deep learning methods, is gaining attention among researchers because its ability to extract features automatically from images with minimal processing tasks and overcome the misclassification issues [15, 16].

With rapid development of deep learning approach, analyses using histopathology images has grown popular and able to assist pathologists in achieving better accuracy, reliable and reproducible diagnosis [6, 17]. In deep learning, the convolutional neural network (CNN) is a common technique among researchers due to its ability to learn features representation automatically and perform image classification [18, 19]. Various applications of CNN exist for histopathology images such as cancer diagnosis [20], image segmentation [21, 22], and illness classification [23]. Due to the advantages of CNN approach in avoiding hand-crafted features extraction with high resolution images, CNN showed more reliable results as compared to other traditional machine learning models. The classification process hinges on the effectiveness of the learned features that capture the fundamental characteristics of the images [24, 25]. While earlier mentioned CNN methods achieved successful classification results, they were limited to tuning different processes or strategies for better feature information to enhance the model's efficiency such as utilizing transfer learning, regularization methods, hyperparameter tuning and optimization algorithms.

The complexity of histopathology images and the dependence on image quality for classification performance contribute to the increased appeal of implementing CNN models with other algorithms. In recent study, Shearlet transform has been used for digital pathology image analysis due to its abilities to make the application of image processing more significant, such as ability to analyze images at different orientations in multiscale framework and produce sparse representation which overcomes the limitation in wavelet transform. Apart from that, Shearlet transform also manages to address singularities using effective and computational methods. Singularities happen when points, lines, or surfaces of image on certain areas have abrupt changes on the brightness or texture. Hence, addressing singularities on medical images can assists on accuracy improvement of classification model and provide more reliable results [26]. In 2016, a model was proposed with the Shearlet transform as feature extraction method on breast cancer images where Shearlet coefficients were used to construct the feature vectors before passed into traditional SVM classifier to classify the mammogram images as normal or abnormal and benign or malignant [27]. Their model shows that the usage of fast finite Shearlet transform leads to credible model performance.

Similarly, by taking the advantage of sparse representation on high dimensional data from Shearlet transform, Zhou *et al.* [28] compared the performance of SVM and AdaBoost classifiers using texture features descriptors based on Shearlet, curvelet, contourlet, wavelet and gray level co-occurrence matrix (GLCM) methods. Feature descriptors that emphasize the similarities and distinction of image texture are used as enhancement methods for image classification. Both classifiers showed good results on Shearlet representation with accuracy of 91% and 90% for SVM and AdaBoost models, respectively. Therefore, this study proposes a model that can extract the most features from histopathology images and accommodate the findings as an input to the CNN as the first features extraction layer, thus preserving the accuracy of classification model. Shearlet transform have advantages on ability to analyze complex images with multiscale and multidimensional structures such that the shear operators in Shearlet transform help to capture features at different scale and directional which eventually leads to sparse representation. The contributions of the study are summarized as follows:

- The proposed model uses Shearlet transform as the first features extraction layer for the CNN architecture. Thus, the performance of shearlet-CNN model does improve as compared to wavelet-CNN model.
- The proposed model applies data augmentation and hyperparameters tuning (e.g., Adam optimizer, learning rate) to reduce overfitting and improve model generalization.
- To the best of our knowledge, this is the first application of histopathology breast cancer image classification using Shearlet coefficients (Shearlet transform) and deep learning methods on small and imbalanced dataset.

Literature Review

Approach for Histopathology Image Analysis

In this section, the related works on histopathology images for breast cancer diagnosis were reviewed. We divided this section into two, such as Convolutional Neural Network (CNN) and Shearlet transform approach on histopathology images.

Breast Cancer Histopathology Images Using Convolutional Neural Network

As mentioned in the previous section, research on deep learning received some spotlight among the researchers especially in solving problems related to computer-vision and pattern recognition. Deep learning is a subset of machine learning methods approach involving multiple layers of architecture that connected to each other with learning representation [15]. CNN is among the popular methods in deep learning because the nature of the architecture that manage to extract features automatically with minimal prior manual image processing tasks, hence reducing the computation cost [29–31]. Pre-processing image tasks for input data preparation and enhancement include image resize, data augmentation, color conversion, normalization, and noise reduction. Classifying histopathology images can be challenging due to complex textures, inter-intraclass variability and geometric structures.

Numerous studies have applied CNN methods to improve image classification for breast cancer using histopathology images. For example, Yan *et al.* [7] proposed a combination method of CNN and recurrent neural network (RNN) to classify breast cancer using histopathology images. They used CNN to obtain higher level hierarchical representation of features and to preserve information, RNN helps to capture both short and long-term correlations between the patches. Sudharshan *et al.* [32] presented a multiple instance learning of CNN (MILCNN) framework to cater the labelling and classification problems of histopathology images. Their idea was to organize the patches into bags (patients) instead of labelling all the patches or instances.

Recently, researchers [30–32] have focused their studies on histopathology images for breast cancer classification using pre-trained CNN models such as VGG-16 and GoogleNet. However, none feature engineering method is used for better image representation before the classification process. Contrary, Barzekar and Yu [36] proposed a CNN architecture to classify the histopathology images into two classes using multiple CNNs known as Concatenation of multiple Networks (C-Net). The architecture comprised of Outer, Inner and Middle networks. Inspired by the VGG19 model, they used the Outer network for features extraction with total parameters of the architectures less than 30 million for 224×224 pixels of images.

On the other hand, Majumdar *et al.* [37] presented an ensemble method from three CNN models with Gamma function for image classification. The outputs from GoogleNet, VGG11 and MobileNetV3 are combined and transferred into rank-based ensemble method to classify the images into benign and malignant based on the confidence scores of each image. Their models showed significant performance as compared to other ensemble methods.

Even though various studies have employed convolutional neural network (CNN) for image classification, there remain some unresolved issues for model improvement. For example, Yan *et al.* [7] and Barzekar and Yu [36] used deep learning methods to classify histopathology images without focusing on any feature extraction method. Past studies have shown that extraction of multi-scale features from histopathology images helps to improve the image representation for further classification. While Sudharshan *et al.* [32] uses multiple instance learning to classify histopathology images, however the noise in images is not properly handled and filtered.

This study demonstrates a hybrid shearlet-CNN model for breast cancer histopathology image classification. In contrast with current methods, hybrid shearlet-CNN uses Shearlet transform as the first feature extraction layer in CNN architecture. Characteristics of Shearlet transform that able to capture

multi-scale features from complex images is beneficial for histopathology image classification. In addition, Shearlet transform also manage to eliminate irrelevant information while keeping the important features in images such as edges and curves due to its anisotropic properties.

Shearlet-Based for Histopathology Images

The quality of medical images is essential during information extraction for further analysis. There are several studies that aimed to enhance the quality of medical images for better results on the image analysis. High-quality images provide good visibility on the cancer cell structures which helps to extract important features so that issue on image misinterpretations can be avoided and reduce diagnostics error. In recent studies, Shearlet transform was the reliable method for multivariate data such as images. Histopathology images consist of small cells from the affected tissue where directional sensitivity is an important feature while extracting the information from the images [38]. However, not many studies applied Shearlet transform on histopathology image analysis.

In recent study, a model is proposed by combining the texture features with two other features (color channel histograms and morphological features) using multiple kernel learning (MKL) and classified them into different grading using SVM classifier [9]. Shearlet transform is used to extract texture features from prostate cancer images. Shearlet transform is believed to be better than other traditional approaches such as wavelet and Fourier transforms which are supported with experiment and finding in their previous study [39]. In image analysis, incorporating directional sensitivity methods provides advantages to the model in terms of texture analysis and quality of features extracted from the image. Shearlet transform is one of the well-suited methods for image processing. Also in the same year, Rezaeilouyeh *et al.* [40] introduced a framework for breast cancer detection based on magnitude and phase of Shearlet coefficients. A total of 58 histopathology images was extracted from the University of California, Santa Barbara Bio-Segmentation Benchmark dataset [41] and the classification process is done using CNN method.

On the other hand, Budak and Guzel [42] utilized the SVM classifier on BreakHis dataset with texture features from gray-level co-occurrence matrix (GLCM) of the Shearlet coefficients. They also combine the texture features with color histogram features for better results. For model comparison, models without any Shearlet transform application were compared with their proposed model and surprisingly, the metric of classification model for accuracy achieved promising values for 40×, 100×, 200× and 400× magnification levels.

As far as our knowledge, there are few researches used Shearlet transform and CNN on breast cancer histopathology images. Table 1 highlights the differences between our study and others. Studies by Alinsaif and Lang [47] and Rezaeilouyeh *et al.* [40] performed image decomposition using Shearlet transform such that magnitude and phase of Shearlet coefficients are extracted from histopathology images. Alinsaif and Lang [47] used support vector machine (SVM) to classify breast cancer images (BreakHis dataset) into benign and malignant tumors with large input size of 700×460 pixels. Meanwhile, Rezaeilouyeh *et al.* [40] applied University of California, Santa Barbara (UCSB) benchmark dataset with smaller sample size and lower input resolution for image classification using CNN method.

In comparison, this study proposed a hybrid shearlet-CNN by using magnitude of Shearlet coefficient as input for the model which can provide a less complex deep learning architecture. Input images are resized into 224×224 pixels to reduce the computational cost (time and memory usage) and ensure the proposed model compatible for other pretrained deep learning models. Unlike prior study, shearlet-CNN is applied to smaller sample size and imbalanced dataset for image classification.

Table 1. The comparison of Shearlet-based method for image classification

	Shearlet coefficients	Image	Input size	Model	BC dataset	Sample size
Alinsaif and Lang [47]	Magnitude and phase	RGB	700×460	SVM	BreakHis	2480 benign 5429 malignant
Rezaeilouyeh <i>et al.</i> [40]	Magnitude and phase	RGB	120×120	CNN	UCSB Benchmark dataset	32 benign 26 malignant
Proposed Shearlet-CNN	Magnitude	RGB	224×224	CNN	BreakHis	625 benign 1370 malignant

Shearlet Transform

In the transform methods family, curvelet transform was the pioneer approach which provide sparse approximation that display anisotropic features introduced in 2004 [43]. The representation of curvelet relies on rotation/orientation, scale, and location. The drawback from wavelet on directional sensitivity has been overcome in curvelet. However, the rotation used to control directional sensitivity led to another problem where the discrete domain could not be directly implemented from continuous setting because the rotation does not preserve the discrete lattice structures. Therefore, Shearlet transform was introduced in 2005 that apply shearing parameter instead of rotation for directional sensitivity [44].

The Shearlet multiscale framework incorporates directional sensitivity which allows the extraction of anisotropic features on images where wavelet theory serves as the foundation set of the Shearlet framework. In wavelet transform, the images are decomposed into high-frequency and low-frequency components based on various scales. One of the key factors of wavelet achievement in the past three decades is the competency of wavelet to approximate data sparsely and handle the singularities better than Fourier transform. However, the isotropic nature of wavelet make it perform successfully only in one-dimensional data because of limitation on capturing the geometric information of edges in an image. Most multivariate data (higher dimension data) such as images is controlled by anisotropic features where the data or signals are approximated with distributed singularities compared to pointwise singularities as in wavelet.

This study implemented Shearlet transform to decompose images into Shearlet coefficients and used them for feature extraction in the first layer of deep learning architecture, CNN. Figure 1 illustrates the singularities distribution along the curve based on isotropic and anisotropic representations. It shows that shearlets can capture the edges faster and more smoothly than wavelets due to the parabolic scaling properties, which means that they require fewer coefficients to represent the curves. Usually, histopathology images contain small contours of cancer cells, including edges and curves. This advantage justifies the choice of Shearlet transforms in our work for better feature extraction.

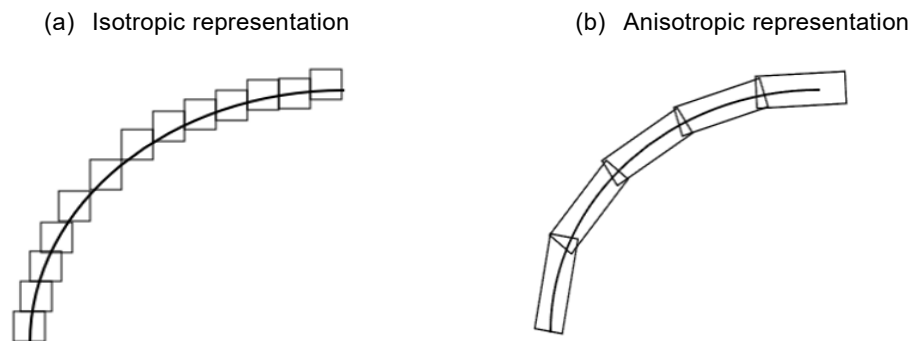


Figure 1. Approximation of curve using wavelet transform (isotropic) and Shearlet transform (anisotropic) [45]

Approximation Discrete Shearlet Transform

A discrete Shearlet transform is generated by sampling the continuous Shearlet transform using proper parameters during the process of transferring from continuous to discrete system. Easley *et al.* [46] had discussed on the details of discrete Shearlet transform. The general equation for discrete Shearlet transform with composite dilations in 2D is presented as follows,

$$SH_{\psi}f(j, k, m) = \langle f, \psi_{j,k,m} \rangle \quad \text{for } f \in L^2(\mathbb{R}^2) \quad (1)$$

where f is the function of image and $\psi_{j,k,m}$ is the Shearlet filter/generator. Let $a = 2^{-j}$, $s = -k$ with $j \in \mathbb{Z}$, $k \in \mathbb{R}$, $m \in \mathbb{Z}^2$. Therefore,

$$\begin{aligned} M_{a,s} &= M_{2^{-j}, -k}^{-1} = M_{2^j, k} \\ &= \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2^j & 0 \\ 0 & \sqrt{2^j} \end{pmatrix} \\ &= \begin{pmatrix} 2^j & k\sqrt{2^j} \\ 0 & \sqrt{2^j} \end{pmatrix} = B_0^k A_0^j \end{aligned} \quad (2)$$

where $A_0 = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{pmatrix}$ and $B_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, A_0 is the anisotropic dilation matrix and B_0 is shear matrix. Hence, the discrete shearlet coefficient is given by,

$$\psi_{j,k,m}(x) = |\det A_0|^{\frac{j}{2}} \psi(B_0^k A_0^j x - m) \quad (3)$$

This study is motivated by studies from Alinsaif and Lang [47]. They use four different descriptors to build the feature vector based on magnitude and relative phase of Shearlet coefficients. In our study, we compute the Shearlet coefficients using Shearlet transform on histopathology images and apply them as primary features for image classification. From Eq. (1), we can compute the Shearlet coefficients $f(j,k,m)$ by taking the inner product of input image and Shearlet function, $\langle f, \psi_{j,k,m} \rangle$. The detail process of Shearlet-CNN model is as tabulated in Algorithm 1 as follows,

Algorithm 1 Proposed Shearlet-CNN model

- 1 **Input:** Histopathology images from BreakHis dataset.
- 2 **Initialize:**
Divide the histopathology images into training dan testing dataset, randomly.
- 3 **First layer feature extraction:**
 - Generate the Shearlet filter for Shearlet system.
 - Calculate the Shearlet coefficients by using Eq. (3).
- 4 **Second layer feature extraction:**
 - Define the CNN architecture using convolution, max pooling, and fully connected layers.
 - Use Shearlet coefficients as input data for CNN model.
- 5 **Model training:**
 - Conduct model training using training dataset.
 - Save model.
- 6 **Model testing:**
 - Calculate the model performance using testing dataset.
 - Apply metric measures.
- 7 **Output:**
Results: Confusion metric, accuracy, precision, recall and F1-score.

Methodology

As described in previous section, the Shearlet transform is used as first feature extraction layer for deep learning method. Then, the CNN model is used to classify breast cancer from histopathology images into benign and malignant. The following sections provide an in-depth explanation of the proposed method based on the considered CNN model. The Jupyter Notebook 6.3.0 is applied for the experiments with Intel Core i7-8565U CPU processor.

Dataset

This study used histopathology images of breast cancer to evaluate the performance of the proposed framework. The dataset is known as Breast Cancer Histopathological dataset (BreakHis) that contains high resolution images of size 460×760 pixels. This public dataset is introduced by Spanhol *et al.* [48] with a total of 7909 images. The data is collected from P&D Laboratory – Pathological Anatomy and Cytopathology in Parana, Brazil where the main class categories of breast cancer tumors are benign and malignant. Furthermore, each histology slide is presented in several magnification levels (40×, 100×, 200× and 400×) which gives option to researchers either to use each magnification level separately or combine them all as one dataset for further image processing and analysis.

The distribution of benign and malignant images is divided into 2480 and 5429 images, respectively. Further images distribution based on magnification level can be obtained from Table 2. As can be seen, the division for benign and malignant images is imbalanced. Consequently, this can lead to bias image classification where the model will tend to classify the images into malignant. Therefore, data partition is applied such that 80:20 ratio for training and validation datasets, respectively.

Table 2. The distributions of images from BreakHis in different magnification levels

Magnification Level	Benign	Malignant
40×	625	1370
100×	644	1437
200×	623	1390
400×	588	1232
Total	2480	5429

Pre- Processing and Data Augmentation

For image analysis, pre-processing is important to make sure the images from dataset are consistent and normalized before further analysis for better result. Image normalization assured the data of pixel from images distributed similarly which helps to make a speed convergence during training process. Originally, the images from BreakHis dataset were in the size of 760×460×3 pixels, then resized into 224×224×3 pixels. The purpose of image resizing is to reduce the dimension calculation so that the training process will consume less time for computation. For preliminary analysis, this study only covered the 40× magnification level of BreakHis dataset. Next, we shuffled the dataset randomly before splitting them into two partitions, namely training and testing dataset. The percentage for both datasets is 80% and 20%, respectively. The training dataset should have a larger portion as compared to the testing dataset so that the training model has the ability to learn meaningful patterns from the data. Table 3 tabulated the distribution of training and testing dataset for this study.

Table 3. The distribution of histopathology images (magnification of 40×)

Class	Dataset Partition	Sample size
Benign/ Malignant	Training	1596
	Testing	399
	Total	1995

Since the medical dataset is small, data augmentation was applied to reduce the possibility of overfitting and build more accurate models. Generally, deep learning models need large datasets for model development, where data augmentation plays important role to overcome problem related to small datasets [13]. For this study, several data augmentation techniques that follow previous study in [49] have been applied during training and the parameter is illustrated in Table 4.

Table 4. Parameter for data augmentation techniques

Data augmentation	Values
Vertical flip	0.5
Horizontal flip	0.5
Rotation range	20.0
Shift range	0.2
Brightness and contrast	0.3

Architecture of CNN

Our main attributes consist of Shearlet coefficients in conjunction with RGB input. The BreakHis dataset is originally in the form of RGB image format which has advantage on color information that help the learning process upgrade the performance of deep learning models. Image conversion into grayscale is a straightforward technique, but nevertheless it could overlook some essential information in color representation. In medical imaging, significant clinical features are often necessary to accurately represent the image, making automatic feature learning crucial for medical image analysis. Consequently, we propose an automatic feature learning approach based on CNN. The proposed framework of the study is as presented in Figure 2.

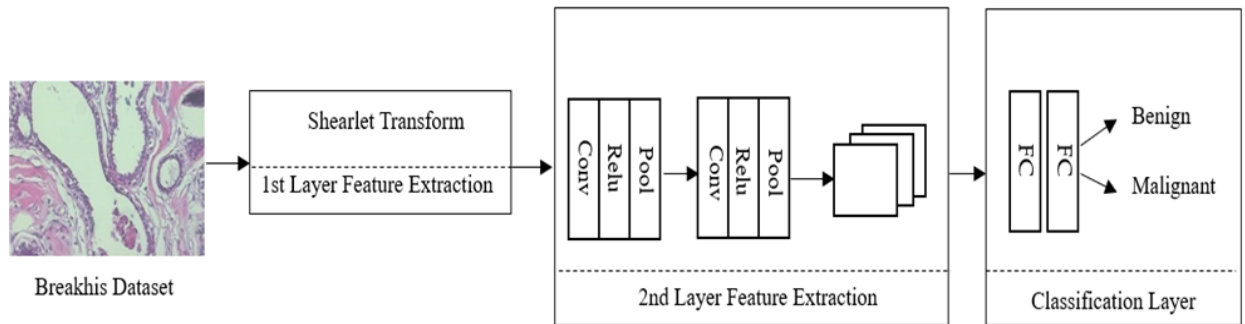


Figure 2. Proposed method for histopathology image classification (Framework)

Traditional machine learning techniques have limited capability when it comes to analyzing natural data in its raw form. This limitation arises because shallow network architecture requires precise techniques for feature extraction and representation which able to cater important characteristics of the input data while being constant to irrelevant features. On the contrary, deep learning (DL) methods employ multiple layers of processing to directly learn representations from raw data [50], capturing multiple levels of abstraction. In the case of images, lower abstraction layers might identify edges, while higher layers recognize objects.

CNN are forward-propagating networks that consist of convolutional and pooling layers consecutively, as well as fully connected layers which specifically designed for 2-D data includes images data [15]. The input data undergoes a series of convolutional and pooling layers. Here is a detailed description of each layer:

- **Convolutional Layer:** Acting as initial layer after the input layer, it performs a 2-D convolution operation on the input features. It employs 32 filters of size 3×3 that uses Gaussian distribution for weight initialization. The filters are initialized with a standard deviation of 0.0001 and a bias value of zero. The choice of these parameters relies on the aim of study such that smaller filter size can effectively capture and address the intricate features within the tissue regions [51]. The output of this layer is then passed through a rectified linear unit (ReLU) function, a non-linear activation function. The ReLU function enables the network to learn abstract representations efficiently with a small number of nodes. Without this non-linearity, the whole network is like a single-layer neural network.
- **Max-Pooling Layer:** This layer's purpose is to combine features based on character similarity and minimize the feature dimensionality. It operates on the output of the previous layer and performs max-pooling within a 2×2 region of the input feature map. By taking the maximum value within each local patch, this layer achieves invariance to shifts and distortions in the learned features.
- **Fully Connected Layer:** Also referred to as the inner product layer, it connects neurons from the previous layer to every neuron in the current layer. This layer resembles a multilayer perceptron network.
- **Dropout:** Dropout regularization is applied in this layer to randomly deactivate a portion of the neurons during the training phase. It overcomes overfitting problem and ensures that the learned results are not overly reliant on the specific structure of the network. In this study, a dropout threshold of 0.2 was chosen, indicating that 20% of the neurons are randomly dropped during training. The dropout choice is supported by Anisuzzaman *et al.* [17].
- **Classification Layer:** The terminal layer of the CNN architecture is a fully connected layer, two neurons for benign and malignant classes. This layer is activated by a softmax classifier, which assigns probabilities to each class based on the network's training labels. Ultimately, this layer provides the accuracy of the classification.

By employing these different layers and their specific functionalities, our CNN architecture aims to effectively learn and extract features from the input data, leading to accurate classification outcomes. The selection of parameters such as max-pooling layer, fully-connected layer and ReLU is due to the benefits stated in previous study by [16, 17, 54]. A simple CNN-based sequential model is employed for this study due to computing power limitation. The architecture is made up of three different layers with one fully connected layer. These layers work in sequence to learn the features of an input and classify the images as the output (benign and malignant). The details of the CNN-based model are shown in Table 5.

Originally, histopathology images served as the input image in input layer. But this study used the Shearlet coefficients as input to be fed into the CNN architecture. The input shape was $224 \times 224 \times 51$ where 51 is the total number of Shearlet coefficients for scales=2. A total of two convolutional layers are added, and the first convolutional layer has 32 filters of size 3×3 feature maps to the input layer. The second convolutional layer consists of 32 filters and ReLU activation function to transform the network from linear into a nonlinear structure. Then, a pooling layer is added in between the convolutional layer to minimize the number of parameters and size of network. A Max-pooling layer with size of 2×2 is chosen for the architecture. Lastly, a fully connected layer is added with a dropout rate of 0.2. This final layer used softmax function to classify the inputs into benign or malignant. The softmax function is commonly used for the probabilistic classification which ranges between 0 to 1. Initially, to save the computational time, the model is trained for 30 epochs, batch of size 32 and 0.0001 learning rate of Adam optimizer with total number of parameters 784 354.

Table 5. The architecture of simple CNN-based model

Layer	Filter	Kernel Size
Input		
Convolutional layer 1	32	3×3
ReLU		
Max_Pooling layer 1		2×2
Convolutional layer 2	32	3×3
ReLU		
Dropout		
Fully connected layer		

Shearlet Coefficients as the Image Representation

In the study, each image is represented using Shearlet coefficients as the input that will be fed into CNN during the classification task. For the Shearlet coefficients extraction, we utilized the *pyShearLab*, a Shearlet toolbox in Python by Stefan Looock [52]. In this preliminary study, we choose scales=2 for the Shearlet. The number of Shearlet filters is 17 for each channel of images. Since the study uses 3-channels images, the total Shearlet filters are 51. Eq. (4) below is used to calculate the total number of Shearlet filter,

$$n_{\text{Shearlet}} = 2 \sum_{j \in J} (|k_j| - 1) + 1 \quad (4)$$

where $|k_j| = 2 \times \lceil 2^{j/2} \rceil + 1$ and $J = \text{scales}$.

First, we need to construct the Shearlet system to compute the Shearlet coefficients of the images. As depicted in Eq. (4), the number of Shearlet is referring to the possible number of Shearlet filters of the corresponding scale, j . This Shearlet system creates a cone-adapted band-limited system for Shearlet. Then, after we have the Shearlet system, we can compute the Shearlet coefficients or Shearlet decomposition.

Performance Evaluation

For quantitative measurement, the elements from confusion matrix are used for classification. The matrix is made up of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). In BreakHis dataset, TP is the tumor images that are classified as benign correctly, while TN shows the tumor images which classified as malignant correctly. On the other hand, the FP measures the malignant images that are misclassified as benign, and FN represents the benign images that are mistakenly classified as malignant.

Next, the performance of classification model is evaluated using different performance measures known as accuracy, sensitivity, specificity, and F1-score. The details on these measures are as described below,

- Accuracy is the percentage of images that are classified correctly from the total number of images.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (5)$$

- Sensitivity and specificity measure the percentage of images that are classified correctly as single class, benign and malignant, respectively. Here, sensitivity is also known as recall where the percentage of benign (correctly positive class) is predicted out of all benign images.

$$Sensitivity = \frac{TP}{TP+FN} \quad (6)$$

$$Specificity = \frac{TN}{TN+FP} \quad (7)$$

- Precision is a metric that measures the percentage of correctly classified benign out of all predicted benign images.

$$Precision = \frac{TP}{TP+FP} \quad (8)$$

- F1-score evaluates the average of precision and recall (sensitivity) which the maximum value of F-score shows optimal balance from precision and recall.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

Results and Discussions

In this section, the performance of proposed framework is evaluated by taking into consideration the performance metrics of the model based on the confusion matrix concept. The model is trained using a publicly available dataset (BreakHis), with 80% training and 20% testing datasets for the data partition. In this study, we used a twofold cross-validation method. For model comparison, we also run the same CNN model with the original image (without transformation). This study purposely used basic CNN method with four layers depth of architecture in order to highlight the contribution of Shearlet transform as the first features extraction layer. In our previous work [49], we manage to show that Shearlet transform plays an important role in preserving the quality of histopathology images for further classification task.

Table 6 and Table 7 visualize the confusion matrices obtained from the experiment of using breast cancer histopathology images of the original CNN model and proposed model. From these tables, we can clearly see that the use of Shearlet coefficients as inputs to the CNN model for image classification provides better results where the proposed model is able to distinguish the images into benign or malignant.

Table 6. The confusion matrix of CNN model

		Predicted	
		Benign	Malignant
Actual	Benign	3	144
	Malignant	0	351

Table 7. The confusion matrix of the proposed model, shearlet-CNN

		Predicted	
		Benign	Malignant
Actual	Benign	26	121
	Malignant	5	346

Table 8 reports on the classification results for both models in terms of sensitivity, specificity, average accuracy, and F1-score values over 4-fold cross-validation. For the measurement of sensitivity and specificity values, the sensitivity should be higher as compared to specificity because the number of malignant images is higher than benign images. Based on the results in Table 8, the sensitivity for proposed model is higher than the specificity as expected due to imbalanced data distribution.

Furthermore, we can conclude that the average accuracy of the proposed model is better by including the Shearlet coefficients in the framework.

However, since the dataset is imbalanced, precision and recall are the more suitable metrics for the discussion because these metrics are able to distinguish between types of errors, for example, false positive and false negative. Therefore, one way to check the performance of imbalanced data is by looking at the F1-score where the calculation is made of precision and recall. The total F1-score for the proposed model is 0.60 which shows good precision (0.79) and recall (0.58) as compared to the original CNN model.

Table 8. The comparison of performance metrics for breast cancer classification

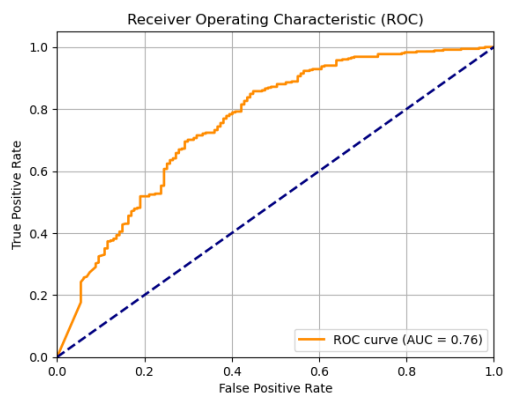
Model	Accuracy	Sensitivity	Specificity	F1-Score
CNN	0.71	1.00	0.02	0.43
Shearlet-CNN	0.75	0.99	0.18	0.60

We also compared the Shearlet-based feature extraction model with wavelet-CNN using the same parameters (e.g. batch size, learning rate, Adam optimizer) as in shearlet-CNN for better comparison results as reported in Table 9 and Figure 3. Our wavelet-CNN used Symlet wavelet. In general, Shearlet transform is an improvement of wavelet transform, therefore the shearlet-CNN model is expected to be better than wavelet-CNN model. Table 9 shows that shearlet-CNN has higher accuracy of 75% with better precision and recall values as compared to wavelet-CNN. On the other hand, receiver operating characteristic (ROC) curve illustrates the performance of binary classification based on true positives and false positives. The ROC curve is also computed as a measure of effectiveness. In summary, we concluded that the proposed model is able to perform successfully as compared to the other.

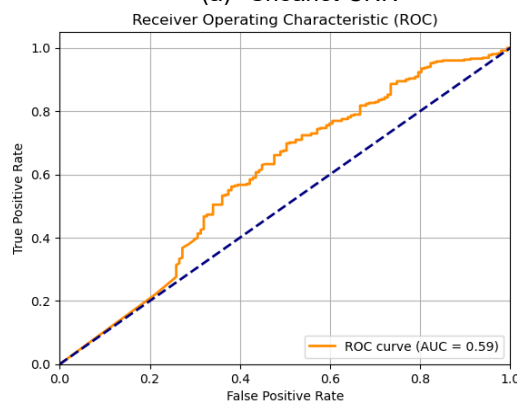
Table 9. Comparison of performance metrics using different first feature extraction layers

Model	Wavelet-CNN	Shearlet-CNN
Accuracy	0.70	0.75
Precision	0.51	0.79
Recall	0.50	0.58
F1-score	0.42	0.60

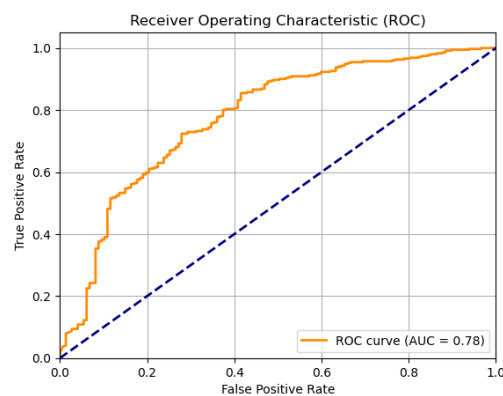
The area under curve (AUC) of ROC illustrates different progress for different models. Based on the rule of thumb, a higher AUC indicates an improved tumor classification for breast cancer. The classification performance of the shearlet-CNN model was assessed using the ROC curve. As shown in Figure 3, the shearlet-CNN model achieved an AUC of 0.76, reflecting its strong ability to discriminate between benign and malignant histopathology images. A higher AUC value indicates a more robust performance in distinguishing the two classes across various decision thresholds. In comparison, the wavelet-CNN model yielded a lower AUC of 0.59, demonstrating that the Shearlet-CNN offers improved classification effectiveness. In addition, Table 10 shows the performance metrics based on tumors category (benign and malignant) for the proposed model. For imbalanced classification problem, the aimed is to increase the recall value without losing the precision value and vice versa for an optimal evaluated model [53].



(a) Shearlet-CNN



(b) avelet-CNN



(c) CNN

Figure 3. Receiver operating characteristics (ROC) curve for different models**Table 10.** The performance metrics for shearlet-CNN model by category of tumors

Class	Precision	Recall	F1-Score
Benign	0.84	0.18	0.29
Malignant	0.74	0.99	0.85

In this study, recall is more crucial than precision because if malignant images are falsely predicted as benign indicate it is not good and risky for the patients. As can be seen in Table 10, the precision and recall for malignant images are both high due to the majority class of malignant in the dataset. The F1-

scores are reasonably good for benign class and malignant class, 0.29 and 0.85 respectively. In order to test on the generalizability and robustness of the shearlet-CNN model, another dataset is applied to the model. Table 11 presents the performance of the shearlet-CNN model with another histopathology image dataset [55]. Same metric measures are used to evaluate the performance of the model. The shearlet-CNN model achieved promising outcomes where the accuracy is 0.79, sensitivity 0.68, specificity 0.83 and F1-score 0.74. The model also demonstrated low misclassification rates during malignant tumor classification which is proven based on recall (0.75) and precision (0.74).

Table 11. Performance of shearlet-CNN using different dataset [55]

Model	Shearlet-CNN
Accuracy	0.79
Precision	0.74
Recall	0.75
F1-score	0.74

This shows that the model with Shearlet coefficients in the CNN architecture is better in predicting benign and malignant for imbalanced dataset. In contrast with Rezaeilouyeh *et al.* [40], our proposed model shows competitive performance in spite of the fact that having unbalanced category of tumors (benign and malignant) and small dataset.

Conclusion

Breast cancer is part of the major concerns illnesses that affecting women, globally. Advancement of technology helps medical experts to study the existence of breast cancer using various approaches. One of them is breast histopathology image examination where the sample tissue collected from patients were processed before being visualized under the microscope for image analysis and diagnosis. In this study, we aimed to enhance the classification performance by improving the image representation of the subject. This study proposes Shearlet coefficients as the inputs to be fed into convolutional neural network (CNN) based on breast cancer histopathology images. Instead of using histopathology images of size $224 \times 224 \times 3$, we transform the image into Shearlet coefficients with size $224 \times 224 \times 51$ at scales=2. The experiment shows that the proposed approach achieved promising classification results in terms of accuracy and F1-score as compared to the original CNN method.

In the future, additional experiments on hyperparameters tuning will be conducted to optimize the performance of the proposed model with an imbalanced dataset. This includes adjusting parameters such as the kernel size, number of epochs, batch size, learning rate and dropout probability. Additionally, the proposed model should be extended to other categories of histopathology images related to human body tissues, such as dermatopathology and neuropathology images which focus on skin tissue and the nervous system, respectively.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgement

S. S. M. Khairi thanks Ministry of Higher Education of Malaysian and Universiti Teknologi MARA for the scholarship of Skim Latihan Akademik Bumiputera. Also sincere thanks to Faculty of Computer and Mathematical Sciences, UiTM Shah Alam and Faculty of Science and Technology, UKM Bangi. This work was supported by funding from Universiti Kebangsaan Malaysia (TAP-K017073).

References

- [1] Wilkinson, L., & Gathani, T. (2022). Understanding breast cancer as a global health concern. *British Journal of Radiology*, 95(1130), 20211033. <https://doi.org/10.1259/bjr.20211033>
- [2] Mubarik, S., *et al.* (2022). Breast cancer mortality trends and predictions to 2030 and its attributable risk factors in east and south asian countries. *Frontier in Nutrition*, 9(2022). <https://doi.org/10.3389/fnut.2022.847920>

- [3] Ibrahim, S., Nazir, S., & Velastin, S. A. (2021). Feature selection using correlation analysis and principal component analysis for accurate breast cancer diagnosis. *Journal of Imaging*, 7(11), 225. <https://doi.org/10.3390/jimaging7110225>
- [4] Smith-Bindman, R., Miglioretti, D. L., & Larson, E. B. (2008). Rising use of diagnostic medical imaging in a large integrated health system. *Health Affairs (Millwood)*, 27(6), 1491–1502. <https://doi.org/10.1377/hlthaff.27.6.1491>
- [5] Vuong, T. T. L., *et al.* (2022). Multi-scale binary pattern encoding network for cancer classification in pathology images. *IEEE Journal of Biomedical and Health Informatics*, 26(3), 1152–1163. <https://doi.org/10.1109/JBHI.2021.3099817>
- [6] Gao, Z., *et al.* (2022). A convolutional neural network and graph convolutional network based framework for classification of breast histopathological images. *IEEE Journal of Biomedical and Health Informatics*, 26(7), 3163–3173. <https://doi.org/10.1109/JBHI.2022.3153671>
- [7] 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>
- [8] Xu, Y., *et al.* (2017). Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics*, 18(1), 281–297. <https://doi.org/10.1186/s12859-017-1685-x>
- [9] Rezaeilouyeh, H., & Mahoor, M. H. (2016). Automatic gleason grading of prostate cancer using shearlet transform and multiple kernel learning. *Journal of Imaging*, 2(3), 25. <https://doi.org/10.3390/jimaging2030025>
- [10] Kavitha, M., Lavanya, G., Janani, J., & Balaji, J. (2018). Enhanced SVM classifier for breast cancer diagnosis. *International Journal of Engineering Technologies and Management Research*, 5(3), 67–74. <https://doi.org/10.29121/ijetmr.v5.i3.2018.178>
- [11] Asri, H., *et al.* (2016). Using machine learning algorithms for breast cancer risk prediction and diagnosis. *Procedia Computer Science*, 83, 1064–1069. <https://doi.org/10.1016/j.procs.2016.04.224>
- [12] Al-Salihi, N. K., & Ibrikci, T. (2017). Classifying breast cancer by using decision tree algorithms. In *Proceedings of the 6th International Conference on Software and Computer Applications* (pp. 144–148). ACM Press.
- [13] Al-antari, M. A., *et al.* (2018). A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. *International Journal of Medical Informatics*, 117, 44–54. <https://doi.org/10.1016/j.ijmedinf.2018.06.003>
- [14] Md Idris, N., *et al.* (2020). Feature selection and risk prediction for patients with coronary artery disease using data mining. *Medical & Biological Engineering & Computing*, 58, 3123–3140. <https://doi.org/10.1007/s11517-020-02268-9>
- [15] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
- [16] Sharma, A. K., *et al.* (2023). Brain tumor classification using the modified ResNet50 model based on transfer learning. *Biomedical Signal Processing and Control*, 86, 105299–105312. <https://doi.org/10.1016/j.bspc.2023.105299>
- [17] Anisuzzaman, D. M., *et al.* (2021). A deep learning study on osteosarcoma detection from histological images. *Biomedical Signal Processing and Control*, 69, 102931–102939. <https://doi.org/10.1016/j.bspc.2021.102931>
- [18] AlZoubi, A., *et al.* (2024). Classification of breast lesions in ultrasound images using deep convolutional neural networks: Transfer learning versus automatic architecture design. *Medical & Biological Engineering & Computing*, 62(1), 135–149. <https://doi.org/10.1007/s11517-023-02922-y>
- [19] Khan, S. U., *et al.* (2019). A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recognition Letters*, 125, 1–6. <https://doi.org/10.1016/j.patrec.2019.03.022>
- [20] Alkhathlan, L., & Saudagar, A. K. J. (2022). Predicting and classifying breast cancer using machine learning. *Journal of Computational Biology*, 29(6), 497–514. <https://doi.org/10.1089/cmb.2021.0236>
- [21] Peng, B., *et al.* (2018). Fully convolutional neural networks for tissue histopathology image classification and segmentation. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (pp. 1403–1407). IEEE.
- [22] Priego-Torres, B. M., *et al.* (2020). Automatic segmentation of whole-slide H&E stained breast histopathology images using a deep convolutional neural network architecture. *Expert Systems with Applications*, 151, 113387–113400. <https://doi.org/10.1016/j.eswa.2020.113387>
- [23] Murtaza, G., *et al.* (2020). Ensembled deep convolution neural network-based breast cancer classification with misclassification reduction algorithms. *Multimedia Tools and Applications*, 79(25), 18447–18479. <https://doi.org/10.1007/s11042-020-08692-1>
- [24] Al-antari, M. A., Han, S. M., & Kim, T. S. (2020). Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms. *Computer Methods and Programs in Biomedicine*, 196, 105584–105598. <https://doi.org/10.1016/j.cmpb.2020.105584>
- [25] Arafa, D. A., *et al.* (2024). A deep learning framework for early diagnosis of Alzheimer's disease on MRI images. *Multimedia Tools and Applications*, 83, 3767–3799. <https://doi.org/10.1007/s11042-023-15738-7>
- [26] Khoramshahi, E., *et al.* (2020). An image-based real-time georeferencing scheme for a UAV based on a new angular parametrization. *Remote Sensing*, 12(19), 1–27. <https://doi.org/10.3390/rs12193185>
- [27] Gedik, N. (2016). A new feature extraction method based on multi-resolution representations of mammograms. *Applied Soft Computing*, 44, 128–133. <https://doi.org/10.1016/j.asoc.2016.04.004>
- [28] Zhou, S., *et al.* (2013). Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image. *Biomedical Signal Processing and Control*, 8(6), 688–696. <https://doi.org/10.1016/j.bspc.2013.06.011>
- [29] Talo, M., *et al.* (2019). Convolutional neural networks for multi-class brain disease detection using MRI images. *Computerized Medical Imaging and Graphics*, 78, 101673–101684. <https://doi.org/10.1016/j.compmedimag.2019.101673>

- [30] Hashemzahi, R., *et al.* (2020). Detection of brain tumors from MRI images base on deep learning using hybrid model CNN and NADE. *Biocybernetics and Biomedical Engineering*, 40(3), 1225–1232. <https://doi.org/10.1016/j.bbe.2020.06.001>
- [31] Farhan, A. M. Q., & Yang, S. (2023). Automatic lung disease classification from the chest X-ray images using hybrid deep learning algorithm. *Multimedia Tools and Applications*, 82(25), 38561–38587. <https://doi.org/10.1007/s11042-023-15047-z>
- [32] Sudharshan, P. J., *et al.* (2019). Multiple instance learning for histopathological breast cancer image classification. *Expert Systems with Applications*, 117, 103–111. <https://doi.org/10.1016/j.eswa.2018.09.049>
- [33] Kumar, A., *et al.* (2020). Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer. *Information Sciences*, 508, 405–421. <https://doi.org/10.1016/j.ins.2019.08.072>
- [34] Hameed, Z., *et al.* (2020). Breast cancer histopathology image classification using an ensemble of deep learning models. *Sensors*, 20, 4373–4389. <https://doi.org/10.3390/s20164373>
- [35] Lin, C. J., & Jeng, S. Y. (2020). Optimization of deep learning network parameters using uniform experimental design for breast cancer histopathological image classification. *Diagnostics*, 10, 662–672. <https://doi.org/10.3390/diagnostics10090662>
- [36] Barzekar, H., & Yu, Z. (2022). C-Net: A reliable convolutional neural network for biomedical image classification. *Expert Systems with Applications*, 187, 1–9. <https://doi.org/10.1016/j.eswa.2021.116003>
- [37] Majumdar, S., Pramanik, P., & Sarkar, R. (2023). Gamma function based ensemble of CNN models for breast cancer detection in histopathology images. *Expert Systems with Applications*, 213, 1–14. <https://doi.org/10.1016/j.eswa.2022.119022>
- [38] Chakraborty, J., *et al.* (2015). Detection of the nipple in mammograms with Gabor filters and the Radon transform. *Biomedical Signal Processing and Control*, 15, 80–89. <https://doi.org/10.1016/j.bspc.2014.09.001>
- [39] Liu, Y., *et al.* (2020). Breast tumors recognition based on edge feature extraction using support vector machine. *Biomedical Signal Processing and Control*, 58, 101825–101832. <https://doi.org/10.1016/j.bspc.2019.101825>
- [40] Rezaeilouyeh, H., Mollahosseini, A., & Mahoor, M. H. (2016). Microscopic medical image classification framework via deep learning and shearlet transform. *Journal of Medical Imaging*, 3(4), 1–23. <https://doi.org/10.1117/1.jmi.3.4.044501>
- [41] Drelie, Gelasca, E., *et al.* (2008). Evaluation and benchmark for biological image segmentation. In *2008 15th IEEE International Conference on Image Processing* (pp. 1816–1819). IEEE.
- [42] Budak, Ü., & Güzel, A. B. (2020). Automatic grading system for diagnosis of breast cancer exploiting co-occurrence shearlet transform and histogram features. *IRBM*, 41(2), 106–114. <https://doi.org/10.1016/j.irbm.2020.02.001>
- [43] Candès, E. J., & Donoho, D. L. (2004). New tight frames of curvelets and optimal representations of objects with piecewise C2 singularities. *Communications on Pure and Applied Mathematics*, 57(2), 219–266. <https://doi.org/10.1002/cpa.10116>
- [44] Guo, K., Kutyniok, G., & Labate, D. (2005). Sparse multidimensional representations using anisotropic dilation and shear operators. In *International Conference on the Interaction between Wavelets and Splines* (pp. 189–201).
- [45] Grohs, P., *et al.* (2014). Parabolic molecules: Curvelets, shearlets, and beyond. In G. E. Fasshauer & L. L. Schumaker (Eds.), *Springer Proceedings in Mathematics and Statistics* (pp. 141–172). Springer International Publishing.
- [46] Easley, G., Labate, D., & Lim, W. Q. (2008). Sparse directional image representations using the discrete shearlet transform. *Applied and Computational Harmonic Analysis*, 25(1), 25–46. <https://doi.org/10.1016/j.acha.2007.09.003>
- [47] Alinsaif, S., & Lang, J. (2019). Shearlet-based techniques for histological image classification. In *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 1424–1431). IEEE.
- [48] Spanhol, F. A., *et al.* (2016). A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, 63(7), 1455–1462. <https://doi.org/10.1109/TBME.2015.2496264>
- [49] Khairi, S. S. M., *et al.* (2023). Comparative analysis of image denoising techniques for histopathology images. In *The 7th International Conference on Quantitative Sciences and its Applications (ICOQSIA2022)* (pp. 040007-1–040007-6). AIP Publishing.
- [50] Sajid, U., *et al.* (2023). Breast cancer classification using deep learned features boosted with handcrafted features. *Biomedical Signal Processing and Control*, 86(C), 105353–105364. <https://doi.org/10.1016/j.bspc.2023.105353>
- [51] Sharma, H., *et al.* (2017). Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. *Computerized Medical Imaging and Graphics*, 61, 2–13. <https://doi.org/10.1016/j.compmedimag.2017.06.001>
- [52] Looock, S. (2011). *pyShearLab - A Python 2D Shearlet Toolbox*. <https://github.com/stefanloock/pyshearlab>
- [53] Hoens, T. R., & Chawla, N. V. (2013). Imbalanced datasets: From sampling to classifiers. In *Imbalanced Learning* (pp. 43–59). John Wiley & Sons, Inc.
- [54] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* (pp. 1097–1105).
- [55] Janowczyk, A., & Madabhushi, A. (2016). Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 7(1), 29. <https://doi.org/10.4103/2153-3539.186902>