

Breast cancer classification with histopathological image based on machine learning

Jia Rong Leow, Wee How Khoh, Ying Han Pang, Hui Yen Yap

Faculty of Information Science and Technology, Multimedia University, Malacca, Malaysia

Article Info

Article history:

Received Nov 15, 2022

Revised Mar 29, 2023

Accepted Apr 7, 2023

Keywords:

Breast cancer classification
Convolution neural network
Image processing
Machine learning
Transfer learning

ABSTRACT

Breast cancer represents one of the most common reasons for death in the worldwide. It has a substantially higher death rate than other types of cancer. Early detection can enhance the chances of receiving proper treatment and survival. In order to address this problem, this work has provided a convolutional neural network (CNN) deep learning (DL) based model on the classification that may be used to differentiate breast cancer histopathology images as benign or malignant. Besides that, five different types of pre-trained CNN architectures have been used to investigate the performance of the model to solve this problem which are the residual neural network-50 (ResNet-50), visual geometry group-19 (VGG-19), Inception-V3, and AlexNet while the ResNet-50 is also functions as a feature extractor to retrieve information from images and passed them to machine learning algorithms, in this case, a random forest (RF) and k-nearest neighbors (KNN) are employed for classification. In this paper, experiments are done using the BreakHis public dataset. As a result, the ResNet-50 network has the highest test accuracy of 97% to classify breast cancer images.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Wee How Khoh

Faculty of Information Science and Technology, Multimedia University

St. Ayer Keroh Lama, Bukit Beruang, 75450, Melaka, Malaysia

Email: whkhoh@mmu.edu.my

1. INTRODUCTION

Breast cancer has a high death rate [1]. However, nowadays the chances of cure are excellent with the ongoing advancement of advanced treatment levels and the upgrading of equipment. Current breast-conserving surgery has a therapeutic result comparable to major resection and modified radical surgery, which not only protects the structural integrity of women's breasts but also reduces their physical and psychological trauma. [2]. Early identification of breast cancer has a greater survival percentage than medium and late stages. Due to the fact that the cancer cells in the early stages of breast cancer have not disseminated, they are amenable to treatment with local surgery, radiation, chemotherapy, hormone therapy, and other comprehensive treatments with a very high cure rate.

Current medical technology and pharmaceuticals for treating breast cancer have advanced significantly in comparison to the past, as it has the rate of early identification of breast cancer. As a result, early diagnosis of breast cancer may boost the success rate of therapy and assist to decrease the death rate of breast cancer patients [3]. This work proposes the use of pre-trained convolutional neural network (CNN) algorithms for distinguishing breast cancer histological images of patients. BreakHis is the dataset that employed in this study. This paper also demonstrates how the CNN performance could be mixed with machine learning methods for classification. Finally, a state-of-art result of the models to categorize breast cancer as benign or malignant, and compared the performances of all of the pre-trained models that are employed.

2. RELATED WORK

Reza and Ma [4] employed CNN to classify imbalanced data from histopathological photos of malignant and non-cancerous tissues. The dataset is heavily imbalanced since the malignant class image is about 3 times as much as the benign class. The oversampling and under-sampling methods were used to address this imbalance problem. Under sampling resamples the data by eliminating majority classes until the minority class reaches a predetermined proportion of the majority class. However, it may result in the loss of important information. Utilizing oversampling to replicate samples from the minority class. The results demonstrated that oversampling provided the best recognition rate of 86.13% and the oversampling method beat the under-sampling strategy in all circumstances.

Hirra *et al.* [5], employed a deep belief network (DBN) to create a patch-based deep learning system for recognizing breast cancer. DBN is constructed by stacking the restricted Boltzmann machine (RBM) layers and initializing the weights in the network with a greedy learning method. Using the two hidden layers, histopathological image characteristics were retrieved. The first RBM was constructed by joining the input layer and the first hidden layer, whereas the second RBM was built by merging the first and second hidden layers. The PA-DBN-BC model produced an accuracy of 86%.

Zou *et al.* [6] presented a breast cancer histopathology image classification model based on GoogLeNet's Inception V1 CNN model. The global average pooling (GAP) technique was used to integrate images of any size into CNN, while the spatial pyramid pooling (SPP) approach was employed to integrate those images into the network and to a certain number of nodes by showing the feature map in the final convolution layer. Overall, the findings indicated that GAP outperformed SPP because it consisted of fewer parameters, trains quicker, and could reduce overfitting.

Spanhol *et al.* [7] improved the pre-trained networks of various CNN types by enriching the dataset using transfer learning by splitting it into numerous patches. It discovered the optimal results of 85% of accuracy. This strategy successfully handled the problems of computation time and a small training set. In the instance of residual neural network (ResNet), this method yielded a maximum accuracy of 85%.

Alzubaidi *et al.* [8] designed a network with two branches to create a deep network with adequate feature representation for feature extraction and the capacity to accomplish the job with high performance. It was created by expanding the network's breadth and depth without considerably increasing its computational cost. In terms of accuracy, it outperformed earlier approaches and was proven to be more resistant to change in cell volume, color, and structure, which achieved 83.6% for patch-wise classification accuracy and 91.3% for image-wise classification accuracy.

Wang *et al.* [9] proposed a prototype transfer generative adversarial network (PTGAN) to bridge the domain disparity at the style and feature levels. According to the experimental outcomes, the PTGAN achieved around 90% accuracy in categorizing benign and malignant tissues. Xu *et al.* [10] proposed a stacked sparse autoencoder architecture for detecting breast cancer nuclei. The authors compared the stacked sparse autoencoder to other encoders in terms of encoding high-level information from image pixels. Stacked sparse autoencoder with softmax classifier was also tested. In a cohort of 500 histopathology pictures and around 3,500 humanely created distinct nuclei, SSAE had an F-measure of 84.49% and an average area under precision-recall curve (AveP) of 78.83%.

Ahmad *et al.* [11] adapted an existing CNN architecture AlexNet created to classify breast cancer. The training algorithms were designed to extract patches obtained randomly and used a sliding window strategy to allow them to handle high-resolution pictures without modifying the CNN design. The trials demonstrated that CNN outperformed other models on the same dataset, and they also combined CNNs using fusion rules, resulting in a modest increase in recognition rates. Besides that, Parvin *et al.* [12] classified breast cancer histopathology pictures using different types of CNN architectures such as LeNet-5, AlexNet, VGG-16, ResNet-50, and Inception-v1. The Inception-v1 network scored the highest test accuracy.

Nguyen *et al.* [13] are scaling the original images in order to construct a CNN model and categorize breast cancer classes. A building model achieves an accuracy of 73.68%, which is satisfactory in relation to the image's unique characteristics. On the other hand, deep learning can extract features from a dataset without the need to construct feature extractors.

Besides that, Anupama *et al.* [14] provided several CNN classification designs, including AlexNet, Inception-Net, and ResNet. Due to the limitations of standard CNN, capsule networks that capture spatial and orientational information were selected as the architecture. This research reveals that histology image preprocessing increases the capsule network model's precision. It is evident from this study that the performance of traditional designs may be enhanced by data preprocessing and parameter adjustment.

Moreover, Ahmed *et al.* [15] suggested using a single CNN model to concurrently accept input from the same image at four different magnification levels and a CNN model with multiple inputs to classify breast histopathology images as benign or malignant. The model presented here has fewer parameters than earlier

complex models. It allows the model to be trained quicker than other models. In addition, real-time picture enhancement minimizes storage complexity.

Xu and Dong [16], have conduct transfer learning research in breast cancer image categorization. Most feature-representation-transfer algorithms for inductive transfer learning strive to eliminate domain differences. It also provided a method to extract characteristics from breast cancer pictures utilizing multiple pre-trained neural networks, GoogLeNet, visual geometry group neural network (VGGNet) and residual neural network (ResNet). The classifier was then trained using feature data. GoogLeNet, VGGNet, ResNet, and proposed network obtained 93.5%, 94.15%, 94.35%, and 97.525% classification accuracy. DCNN was used for feature extraction, while is method was used for feature selection. The algorithm's reduced calculation cost seems promising. In this survey, we observed fine-tuned pre-trained network achieves greater performance than full-trained network. Because it is challenging to get substantial, well-annotated data for breast cancer classification network training.

Naderan and Zaychenko [17], proposed a convolutional autoencoders were developed to minimize model complexity, parameter number, and overfitting. This work trained autoencoder models using multiple architectures and hyperparameters. The model could rebuild the original picture with 84.72% accuracy with 37% noise of input images. Convolutional autoencoder contains fewer parameters than DenseNet, preventing overfitting. The autoencoder is unsupervised and does not require labelled data. In the experiment, input photos included 37% noise. The model could recreate pictures from input data.

After that, Ukwuoma *et al.* [18] provided an attention mechanism based deep-learning model. The suggested attention method uses pre-trained models as inputs and a multilinear perceptron and softmax as outputs. The proposed model combined DenseNet-201 and VGG-16 for feature extractors into the attention mechanism, yielding more accuracy than pre-trained and ensemble models. The accuracy of classification increased by 1% to 6%.

3. METHOD

3.1. Convolutional neural network models

CNN's [19] are deep learning algorithms in image recognition and processing. A CNN is similar to a multilayer perceptron in that it is based on the same technology as a multilayer perceptron but is optimized for minimal processing needs. A CNN consists of the input layer, output layer, and a hidden layer comprising several convolutional, pooling, fully connected, and normalizing layers. Figure 1 illustrates a sample of conventional CNN architecture. A brief explanation of each layer is included in the following sections.

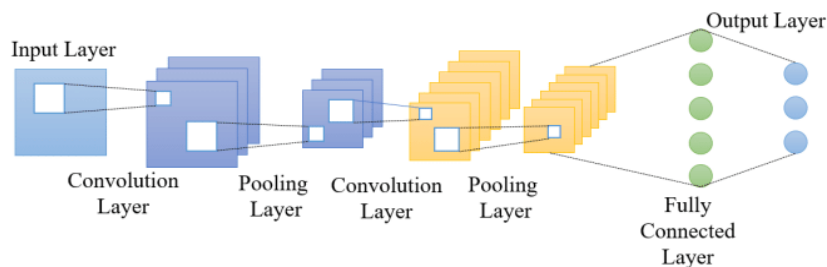


Figure 1. Sample of CNN architecture [20]

3.1.1. Convolutional layer

CNN's essential component is the convolutional layer. It bears the lion's share of the network's computational cost. Besides that, this layer performs a dot product on two matrices, one of which includes the kernel's set of learnable parameters and the other of which contains the receptive field's restricted area. Kernels are smaller than pictures yet contain a higher amount of information. So, if the image is under three red green blue (RGB) channels, the kernel height and width are minimized in terms of spatial dimensions, but the depth includes all three channels.

3.1.2. Pooling layer

A pooling layer plays an important role in CNNs in reducing the complexity of the feature maps produced by the convolutional layer. This is typically achieved through down sampling process which involves summarizing the local features of the filter's output. The pooling process operates on a small window of the input data, selecting a pre-defined value (commonly average or maximum value) in the window to create a simplified

and smaller output feature map. Figure 2 depicts examples of the two most common pooling used in CNN, maximum pooling (left), and average pooling (right). The maximum pooling selects the maximum value in the window (2×2 pool size), emphasizing the most salient feature in the windows while the average pooling computes the mean value in each window to reduce the impact of small variations in the input data. Both possess the effect of reducing the resolution of the output feature map by a factor of the pooling window size.

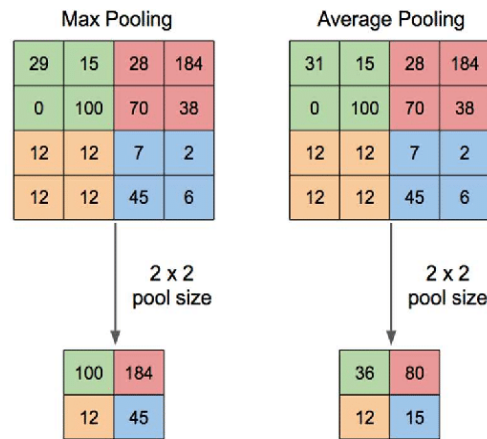


Figure 2. Max and average pooling [21]

3.1.3. Fully-connected layer (FC layer)

As with feedforward neural networks, neurons in the fully-connected layers are entirely dependent on all activations in the previous layer. FC layers are always the last to be implemented in a network. This layer usually receives input from the pooling layer that has been flattened after the image's features have been extracted. The output is then sent to the layer responsible for picture classification.

3.2. Proposed models

The pre-trained CNN models utilized for the BreakHis dataset in this study include the ResNet-50, VGG-19, AlexNet, and Inception-v3. Besides, the ResNet-50 is also used as a feature extraction to obtain the feature from the BreakHis images, then the extracted features are fed into the (RF) random forest and k -nearest neighbors (k -NN) for classification. The architectures of the adopted pre-trained CNN models are described in the subsequence sections.

3.2.1. ResNet-50

The ResNet-50 has 50 layers. The input components of the ResNet network are all made up of a huge convolution kernel with a stride of 2 and a maximum pooling with a size of 3×3 and a stride of 2. In this phase, the input image must be changed to $224 \times 224 \times 3$ dimensions. It contains five convolutional layers for producing feature maps, a rectified linear unit (ReLU) layer for dimensionality reduction, a FC layer, and a softmax activation function for categorizing the images of the breast cancer dataset into two categories (benign or malignant). The first convolution layer provides low-level properties like color, edges, and gradient operation while the deeper convolution layer delivers high-level characteristics of the input data. When the input is received, it processes by the first convolutional layer, which contains 7×7 with 64 filters and has a stride value of 2. The input size is then reduced to 112×112 and the data is processed through the following convolutional layers: 1×1 with 64 filters, 3×3 with 64 filters, and 1×1 with 256 filter stride values of 3. The image is then scaled down to 56×56 pixels and sent through the next convolutional layer, which is 1×1 with 128 filters, 3×3 with 128 filters, and 1×1 with 512 filters with a stride value of 4. The dimension is then reduced to 28×28 and the dimension is transmitted through the next convolutional layer, which is 1×1 with 256 filters, 3×3 with 256 filters, and 1×1 with 1,024 filters with a stride value of 6. The image is then reduced again to a size of 14×14 and transmitted through the next convolutional layer, which is 1×1 with 512 filters, 3×3 with 512 filters, and then 1×1 with 2,048 filters with a stride value of 3.

In the implementation phase, the final layer's block is made trainable alone, and the learning rate is set as 0.00001 to compile the frozen and unfrozen top blocks. The epochs are repeated 15 times, and the data batch size is set as 16. The ResNet-50 architecture is shown in Figure 3 and the configuration of the model is tabled in Table 1.

```

Model: "sequential_1"
-----
Layer (type)                Output Shape                Param #
-----
resnet50 (Functional)       (None, 7, 7, 2048)         23587712
-----
flatten (Flatten)           (None, 100352)             0
-----
dropout (Dropout)           (None, 100352)             0
-----
dense_4 (Dense)              (None, 64)                  6422592
-----
dropout_1 (Dropout)         (None, 64)                  0
-----
dense_5 (Dense)              (None, 2)                   130
-----
Total params: 30,010,434
Trainable params: 29,957,314
Non-trainable params: 53,120
-----
None

```

Figure 3. ResNet-50 architecture

Table 1. Configuration of ResNet-50

ConvNet configuration of Resnet-50
50 layers
Input (224×224×3 RGB image)
Conv1
7×7, 64, stride 2
3×3 max pool, stride 2
Conv2
1×1, 64
3×3, 64×3
1×1, 256
Conv3
1×1, 128
3×3, 128×4
1×1, 512
Conv4
1×1, 256
3×3, 256×6
1×1, 1024
Conv5
1×1, 512
3×3, 512×3
1×1, 2048
Flatten (100352)
FC (64)
soft-max

3.2.2. ResNet-50 with machine learning classifier

The ResNet-50 is combined with another machine learning model to create a hybrid model comprising two blocks each shown in Figure 4. Within the first block, the weight of the pre-trained ResNet-50 which was trained by ImageNet is adopted to retrieve the properties from the input images. Then, the deep features extracted are passed to the second block. The second block uses a machine learning algorithm, including the random forest and k -NN as classifiers to classify deep features. It is utilized in the second segment to categorize the deep feature maps of the CNNs.

3.2.3. VGG-19

The size of the image that the VGG-19 receives as input is 224×224×3. The filter size for the convolutional layer is 3×3, followed by the stride and padding of the convolution layers and five max-pooling layers. The max-pooling was performed using a 2×2 window, and the stride is set at 2. For the entirely interconnected layers, the first layer included 256 neurons, followed by a layer containing 256 neurons. Categorization is achieved using a softmax output layer.

During the implementation phase, a learning rate of 0.001 is used. Additionally, it refines the model such that each layer is trainable. The epoch size of the experiment has been set to 20, and the call-backs function has been used to control the learning rate and early stopping. Figure 5 depicts the VGG-19 design, while Table 2 displays the VGG-19 model's configuration.

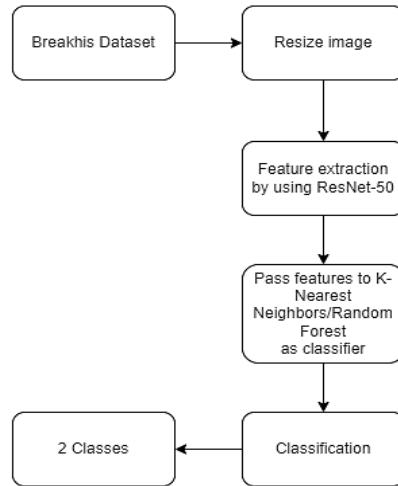


Figure 4. ResNet-50 with machine learning architecture

```

Model: "model"
-----
Layer (type)                Output Shape                Param #
-----
input_1 (InputLayer)        [(None, 224, 224, 3)]      0
-----
block1_conv1 (Conv2D)        (None, 224, 224, 64)       1792
block1_conv2 (Conv2D)        (None, 224, 224, 64)       36928
block1_pool (MaxPooling2D)   (None, 112, 112, 64)       0
block2_conv1 (Conv2D)        (None, 112, 112, 128)      73856
block2_conv2 (Conv2D)        (None, 112, 112, 128)      147584
block2_pool (MaxPooling2D)   (None, 56, 56, 128)        0
block3_conv1 (Conv2D)        (None, 56, 56, 256)        295168
block3_conv2 (Conv2D)        (None, 56, 56, 256)        590880
block3_conv3 (Conv2D)        (None, 56, 56, 256)        590880
block3_conv4 (Conv2D)        (None, 56, 56, 256)        590880
block3_pool (MaxPooling2D)   (None, 28, 28, 256)        0
block4_conv1 (Conv2D)        (None, 28, 28, 512)        1180160
block4_conv2 (Conv2D)        (None, 28, 28, 512)        2359808
block4_conv3 (Conv2D)        (None, 28, 28, 512)        2359808
block4_conv4 (Conv2D)        (None, 28, 28, 512)        2359808
block4_pool (MaxPooling2D)   (None, 14, 14, 512)        0
block5_conv1 (Conv2D)        (None, 14, 14, 512)        2359808
block5_conv2 (Conv2D)        (None, 14, 14, 512)        2359808
block5_conv3 (Conv2D)        (None, 14, 14, 512)        2359808
block5_conv4 (Conv2D)        (None, 14, 14, 512)        2359808
block5_pool (MaxPooling2D)   (None, 7, 7, 512)          0
global_max_pooling2d (Global (None, 512)          0
dense (Dense)                (None, 256)                131328
dropout (Dropout)            (None, 256)                0
dense_1 (Dense)              (None, 2)                  514
-----
Total params: 20,156,226
Trainable params: 131,842
Non-trainable params: 20,024,384
  
```

Figure 5. VGG-19 architecture

3.2.4. Inception V3

The required input image size for the Inception-V3 network is 224×224×3 resolution. This CNN architecture consists of 42-layer convolutional layers. It creates a basic yet robust deeper network, allowing us to reduce computational expenses to a minimum. This network design is summarized in Figure 6.

The factorization of the $n \times n$ convolutional is the inception module. By merging the outputs of all the layers, the output vector is formed. The output of the inception layer is subsequently passed on to the next layer. The inception layer aids the preceding layers in identifying the proper filter size.

The stochastic gradient descent (SGD) is utilized as the optimizer throughout implementation, and all the layers are frozen in order to train only the extra layers. It only fine-tunes the model by retraining a few of the end layers of the Inception-V3 model. The learning rate for the SGD is set to 0.001, decay to 0.00001, and the Nesterov to true. The experiment’s epoch is 20 times, and the batch size is set to 20.

Table 2. Configuration of VGG-19

ConvNet configuration of VGG-19
19 weight layers
Input (224×224×3 RGB image)
Conv 3×3(64)
Conv 3×3(64)
Maxpool
Conv 3×3(128)
Conv 3×3(128)
Maxpool
Conv 3×3(256)
Conv 3×3(256)
Conv 3×3(256)
Conv 3×3(256)
Maxpool
Conv 3×3(512)
Conv 3×3(512)
Conv 3×3(512)
Conv 3×3(512)
Maxpool
Conv 3×3(512)
Conv 3×3(512)
Conv 3×3(512)
Conv 3×3(512)
Maxpool
FC-256
FC-256
soft-max

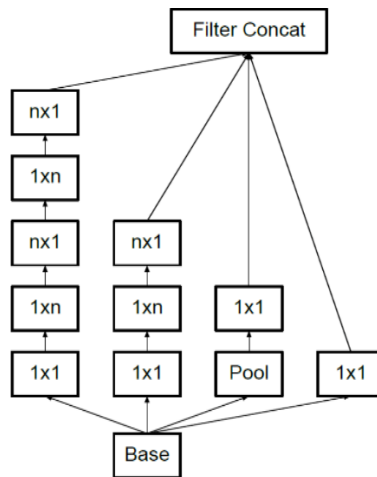


Figure 6. Concept of inception-V3 (5×inception) architecture [22]

3.2.5. AlexNet

The AlexNet is composed of eight convolutional layers. This model is provided with images with dimensions of 224×224×3. The first layer of AlexNet is a convolutional layer which consists of 96 filters with 11×11×3-pixel widths and strides of 4 pixels followed by a pooling layer and a batch normalization layer. Then a ReLU activation layer is used in the convolutional layer. The output feature map now is 27×27×96. Convolutional layers make up the second layer, as do the first. In this layer, the convolutional contains 256 filters with 11×11 filter sizes and 1-pixel strides. The third layer is also a convolutional layer, with 384 filters that are 3×3 pixels in size and the output feature map is 6×6×384. The fourth layer is a convolutional layer as well, with 384 filters that are 3×3 filters in size and the output feature map is 4×4×384. Similarly, the fifth convolutional layer consists of 256 filters that are 3×3 pixels in size. A layer of overlapping pooling and local

response normalization follows, and the output features map is $1 \times 1 \times 256$. In entirely connected layers, the sixth and seventh levels have 4,096 neurons with dropout function and ReLU activation. The classifications are then performed on an outer layer using a softmax activation function. After that, the top two complicated layers are selected to train, while the first eight layers are frozen, and the remaining layers are unfreezing. SGD is employed as the optimizer, using a learning rate of 0.001, a momentum of 0.9, and a decay of 0.05. Figure 7 shows the summary architecture of the AlexNet pre-trained model.

Layer (type)	Output Shape	Param #
conv2d_5 (Conv2D)	(None, 54, 54, 96)	34944
max_pooling2d_3 (MaxPooling2D)	(None, 27, 27, 96)	0
batch_normalization_8 (Batch Normalization)	(None, 27, 27, 96)	384
conv2d_6 (Conv2D)	(None, 17, 17, 256)	2973952
max_pooling2d_4 (MaxPooling2D)	(None, 8, 8, 256)	0
batch_normalization_9 (Batch Normalization)	(None, 8, 8, 256)	1024
conv2d_7 (Conv2D)	(None, 6, 6, 384)	885120
batch_normalization_10 (Batch Normalization)	(None, 6, 6, 384)	1536
conv2d_8 (Conv2D)	(None, 4, 4, 384)	1327488
batch_normalization_11 (Batch Normalization)	(None, 4, 4, 384)	1536
conv2d_9 (Conv2D)	(None, 2, 2, 256)	884992
max_pooling2d_5 (MaxPooling2D)	(None, 1, 1, 256)	0
batch_normalization_12 (Batch Normalization)	(None, 1, 1, 256)	1024
flatten_1 (Flatten)	(None, 256)	0
dense_4 (Dense)	(None, 4096)	1052672
dropout_3 (Dropout)	(None, 4096)	0
batch_normalization_13 (Batch Normalization)	(None, 4096)	16384
dense_5 (Dense)	(None, 4096)	16781312
dropout_4 (Dropout)	(None, 4096)	0
batch_normalization_14 (Batch Normalization)	(None, 4096)	16384
dense_6 (Dense)	(None, 1000)	4097000
dropout_5 (Dropout)	(None, 1000)	0
batch_normalization_15 (Batch Normalization)	(None, 1000)	4000
dense_7 (Dense)	(None, 2)	2002
Total params: 28,081,754		
Trainable params: 4,101,002		
Non-trainable params: 23,980,752		

Figure 7. AlexNet architecture

4. EXPERIMENTAL SETTINGS

The dataset named as BreakHis [23] is used in this study. This dataset consists of a collection of 9,109 samples of breast tumor tissue acquired from 82 cases at a range of magnification which 40, 100, 200, and 400. The database presently has 5,429 malignant as shown in Figure 8 and 2,480 benign as shown in Figure 9. The images from the dataset each have the characteristics of 700×460 pixels, three-channel RGB, and eight-bit depth for each channel. Details of the samples in the dataset are listed in Table 3.

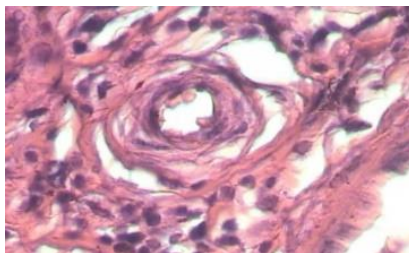


Figure 8. Malignant tumors slide

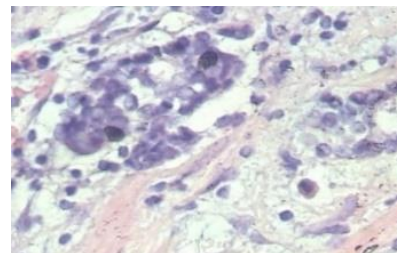


Figure 9. Benign tumors slide

Table 3. Details of BreakHis

Magnification	Benign	Malignant	Total
40×	625	1,370	1,995
100×	644	1,437	2,081
200×	623	1,390	2,013
400×	588	1,232	1,820
Total	2480	5,429	7,909

The BreakHis dataset has been constructed as shown in Table 3 for the training, validation, and testing phase. First, each image from the benign and malignant image folder is loaded and resized from 700×460 pixels to 224×224 pixels so that it is compatible with the input size of the pre-trained CNN models that are feeding in the next phase. Figures 10 and 11 illustrated the resized images in RGB format. Both the benign and malignant data are separately loaded and assigned for train and test variables. After that, the benign train and test variable are labelled as zero class while the malignant train test is labelled as one class. Then, the test data for benign and malignant are combined and mixed into one testing dataset which is used in the testing phase. Following the labelling and merging of the data, the train test split function is used to divide the data into 5,931 for training, 4,745 for validating, and 3,164 for testing the models, as shown in Table 4. The splitting ratio is set as 75:60:40 for training, validation and testing, respectively. After completing the dataset construction, begin training and testing the suggested model.

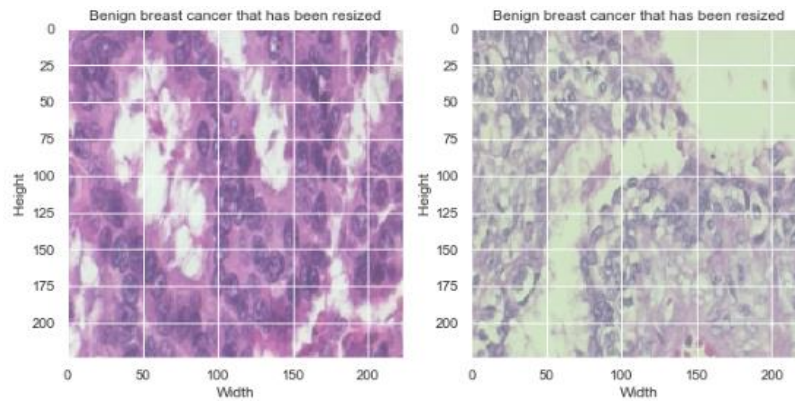
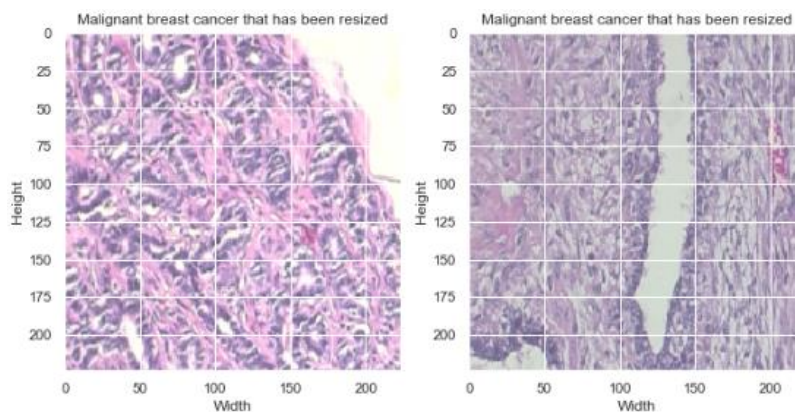
Figure 10. Resized benign image ($224 \times 224 \times 3$)Figure 11. Resized malignant image ($224 \times 224 \times 3$)

Table 4. Data splitting details

Class	Number
Training	5931
Validating	4745
Testing	3164

5. RESULT AND DISCUSSION

In this section, the experimental results and discussion in utilizing transfer learning are presented. Transfer learning is a powerful technique that enables the leverage of knowledge (weights, biases and other learned features) that was previously trained on CNN models to solve the new domain problems. To conduct the experiments, the Dell G15 Special Edition 5521 laptop equipped with the NVIDIA GeForce RTX 3050-Ti Laptop GPU were used. For implementation, the Python programming language version 3.9.12 and TensorFlow-GPU version 2.5.0 were used which allow us to take advantage of the powerful computational capabilities of the GPU to accelerate the experiments. A series of experiments and evaluation of the results based on various metrics, including accuracy, precision, recall, and F-feature are reported in this section.

Figure 12 shows the accuracy of all the pre-trained models for binary classification of breast cancer from the BreakHis database. Figure 12 demonstrates that the Resnet-50 network scored the highest test accuracy with 97.60%, followed by the VGG-19 models with an accuracy of 95.4%. The accuracy rate of the lightweight AlexNet model was the lowest of the pre-trained models, at 81.42%. This demonstrates that increasing the number of convolutional layers in a model may dramatically enhance its classification performance.

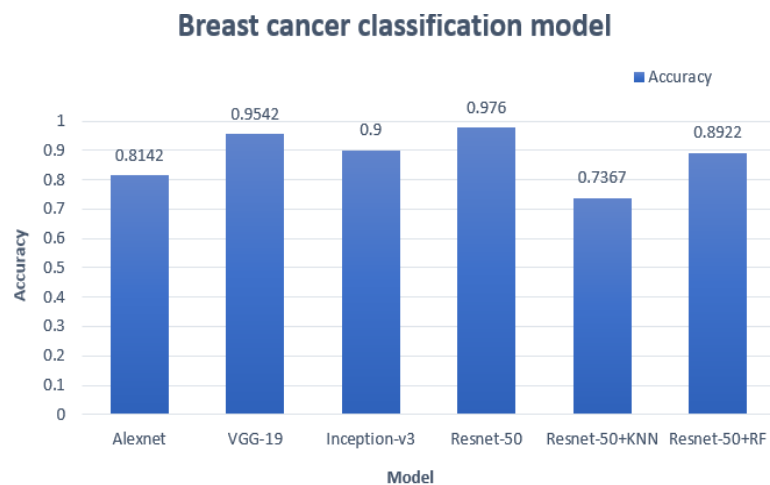


Figure 12. Accuracy performance of each pre-trained model

Intelligent systems often require a substantial quantity of training data to achieve human-level diagnostic abilities. For some illnesses, the quantity of data available for model training may be insufficient, leading to limited generalization of frequently used deep learning models. CNNs are an innovative technique for learning picture representation, with convolutional filters learning to recognize image information via backpropagation. Frequently, these attributes are given to a classifier, such as a softmax layer [24].

However, in this research, instead of using traditional backpropagation, a softmax layer is trained for classification. Deep forest, a decision trees that outperforms deep neural networks (DNN) in a broad variety of applications. Therefore, ResNet50 was used to retrieve information from breast cancer images before categorizing them using deep forest. Deep forest has been proved to be particularly efficient in circumstances when only small-scale training data is available, hence this technique was chosen. Furthermore, the deep forest network chooses its own complexity, it tackles the dataset imbalance problem that we observed in this circumstance. Aside from that, the standard k-NN [25] is a straightforward alternative technique to small class prediction was also employed.

The accuracy based on Figure 12 has shown that the ResNet-50 with RF is better than the ResNet-50 with k-NN since ResNet-50 with RF is achieved at the 89.22% and ResNet-50 with k-NN is 73.67%. Figure 13 shows the actual and predicted results. Table 5 details the area under the curve (AUC), recall, accuracy, and F1-score obtained from all models. It could be seen that the ResNet-50 network achieved the highest performance among the AUC of 97%, precision of 97% and 98%, recall of 95% and 98%, and F1-score of 96% and 98% for benign and malignant, respectively.

Figure 14 shows the confusion matrices for the ResNet-50 model, which achieved the best accuracy performance among the adopted models. This confusion matrix was created utilizing a 75% training and 40% testing technique. The results show that the true positive rate is 68.14%, the true negative rate is 29.27%, the false positive rate is 1.55%, and the false negative rate is 1.04% during the testing stage.

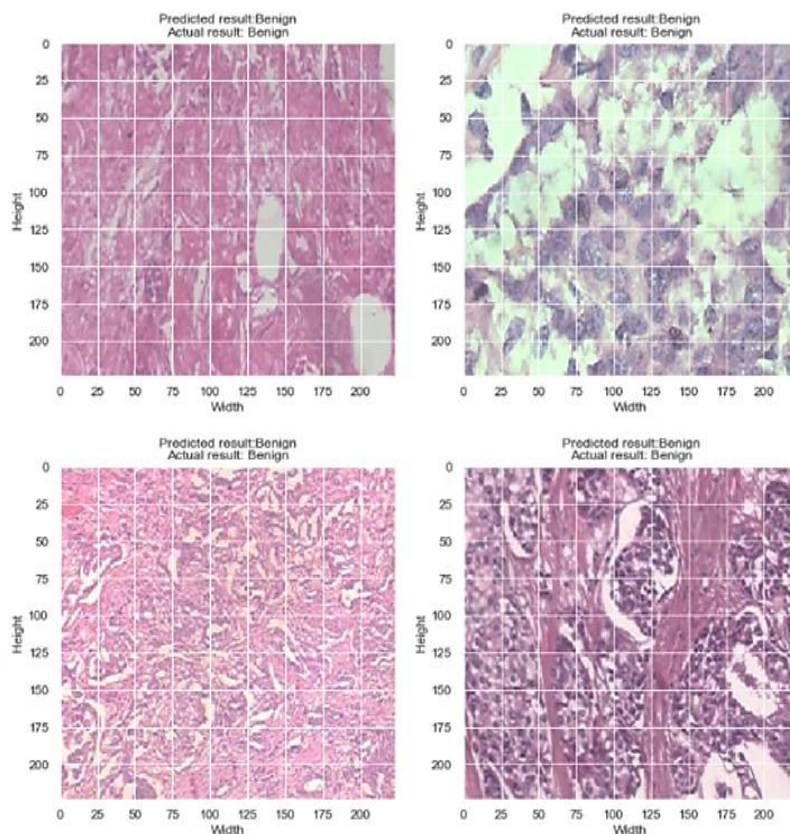


Figure 13. Actual and predicted results

Table 5. Performance of each pre-trained model

Model	Class	Precision	Recall	F-features	Accuracy
AlexNet	Benign	0.82	0.55	0.66	0.81
	Malignant	0.81	0.94	0.87	
VGG 19	Benign	0.82	0.88	0.85	0.90
	Malignant	0.94	0.91	0.93	
Inception V3	Benign	0.96	0.92	0.94	0.96
	Malignant	0.96	0.98	0.97	
ResNet-50	Benign	0.97	0.95	0.96	0.97
	Malignant	0.98	0.98	0.98	
ResNet50+ random forest	Benign	0.84	0.82	0.83	0.89
	Malignant	0.92	0.92	0.92	
ResNet50+KNN	Benign	0.62	0.46	0.53	0.74
	Malignant	0.77	0.87	0.82	

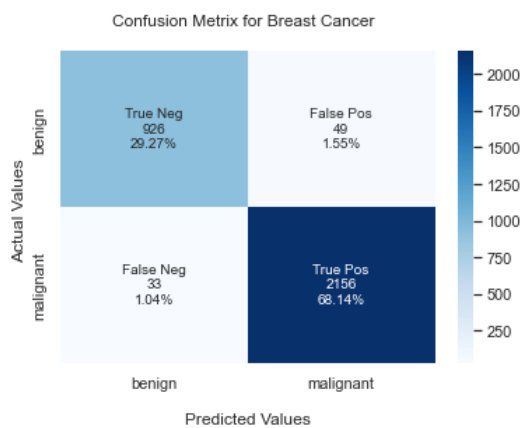


Figure 14. Confusion matrix of ResNet-50 model

6. CONCLUSION

Overall, this research has offered five distinct types of pre-trained CNN models on the classification that may be utilized to categorize benign and malignant breast cancer histopathology images. The models utilized in the study include ResNet-50, VGG-19, Inception-V3, and AlexNet, with the ResNet-50 also working as a feature extractor, extracting features from images and passing them to machine learning algorithms for classification, in this case, a RF and k -NN. The experiments of this study were based on a publicly available dataset named BreakHis.

In conclusion, the ResNet-50 has a higher accuracy for the classification of medical color images, but accurate cancer diagnosis is critical for protecting a person's life, and this research can aim to improve its accuracy as much as possible in the future. This study's proposed architecture might be used in real-world medical imaging in real time. More testing in real-world scenarios is required for clinical applications, and the system should be strengthened. Because of their speed, such devices may allow clinicians to consult with more patients. Furthermore, because the suggested approach is more accurate in cancer classification, the death rate from breast cancer might be greatly lowered.





REFERENCES

- [1] N. Azamjah, Y. Soltan-Zadeh, and F. Zayeri, "Global trend of breast cancer mortality rate: A 25-Year study," *Asian Pacific Journal of Cancer Prevention*, vol. 20, no. 7, pp. 2015–2020, Jul. 2019, doi: 10.31557/APJCP.2019.20.7.2015.
- [2] H. Qiu, W.-H. Xu, J. Kong, X.-J. Ding, and D.-F. Chen, "Effect of breast-conserving surgery and modified radical mastectomy on operation index, symptom checklist-90 score and prognosis in patients with early breast cancer," *Medicine*, vol. 99, no. 11, Mar. 2020, doi: 10.1097/MD.00000000000019279.
- [3] G. N. Sharma, R. Dave, J. Sanadya, P. Sharma, and K. K. Sharma, "Various types and management of breast cancer: an overview," *Journal of Advanced Pharmaceutical Technology and Research*, vol. 1, no. 2, pp. 109–126, 2010.
- [4] M. S. Reza and J. Ma, "Imbalanced histopathological breast cancer image classification with convolutional neural network," in *2018 14th IEEE International Conference on Signal Processing (ICSP)*, Aug. 2018, pp. 619–624, doi: 10.1109/ICSP.2018.8652304.
- [5] I. Hirra *et al.*, "Breast cancer classification from histopathological images using patch-based deep learning modeling," *IEEE Access*, vol. 9, pp. 24273–24287, 2021, doi: 10.1109/ACCESS.2021.3056516.
- [6] W. Zou, H. Lu, K. Yan, and M. Ye, "Breast cancer histopathological image classification using deep learning," in *2019 10th International Conference on Information Technology in Medicine and Education (ITME)*, Aug. 2019, pp. 53–57, doi: 10.1109/ITME.2019.00023.
- [7] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *2016 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2016, pp. 2560–2567, doi: 10.1109/IJCNN.2016.7727519.
- [8] L. Alzubaidi, R. I. Hasan, F. H. Awad, M. A. Fadhel, O. Alshamma, and J. Zhang, "Multi-class breast cancer classification by a novel two-branch deep convolutional neural network architecture," in *2019 12th International Conference on Developments in eSystems Engineering (DeSE)*, Oct. 2019, pp. 268–273, doi: 10.1109/DeSE.2019.00057.
- [9] D. Wang, Z. Chen, and H. Zhao, "Prototype transfer generative adversarial network for unsupervised breast cancer histology image classification," *Biomedical Signal Processing and Control*, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102713.
- [10] J. Xu *et al.*, "Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 119–130, Jan. 2016, doi: 10.1109/TMI.2015.2458702.
- [11] H. M. Ahmad, S. Ghuffar, and K. Khurshid, "Classification of breast cancer histology images using transfer learning," in *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, Jan. 2019, pp. 328–332, doi: 10.1109/IBCAST.2019.8667221.
- [12] F. Parvin, F. Parvin, M. A. M. Hasan, and M. Al Mehedi Hasan, "A comparative study of different types of convolutional neural networks for breast cancer histopathological image classification," in *2020 IEEE Region 10 Symposium (TENSYP)*, 2020, pp. 945–948, doi: 10.1109/TENSYP50017.2020.9230787.
- [13] P. T. Nguyen, T. T. Nguyen, N. C. Nguyen, and T. T. Le, "Multiclass breast cancer classification using convolutional neural network," in *2019 International Symposium on Electrical and Electronics Engineering (ISEE)*, Oct. 2019, pp. 130–134, doi: 10.1109/ISEE2.2019.8920916.
- [14] M. A. Anupama, V. Sowmya, and K. P. Soman, "Breast cancer classification using capsule network with preprocessed histology images," in *2019 International Conference on Communication and Signal Processing (ICCSP)*, Apr. 2019, pp. 0143–0147, doi: 10.1109/ICCSP.2019.8698043.
- [15] M. Ahmed and M. R. Islam, "A multiple-input based convolutional neural network in breast cancer classification from histopathological images," in *2021 24th International Conference on Computer and Information Technology (ICCIT)*, Dec. 2021, pp. 1–5, doi: 10.1109/ICCIT54785.2021.9689856.
- [16] J. Xu and X. Dong, "A survey of transfer learning in breast cancer image classification," in *2020 IEEE 3rd International Conference of Safe Production and Informatization (ICSPI)*, Nov. 2020, pp. 220–223, doi: 10.1109/ICSPI51290.2020.9332405.
- [17] M. Naderan and Y. Zaychenko, "Convolutional autoencoder application for breast cancer classification," in *2020 IEEE 2nd International Conference on System Analysis and Intelligent Computing (SAIC)*, Oct. 2020, pp. 1–4, doi: 10.1109/SAIC51296.2020.9239139.
- [18] C. C. Ukwuoma *et al.*, "Boosting breast cancer classification from microscopic images using attention mechanism," in *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, Mar. 2022, pp. 258–264, doi: 10.1109/DASA54658.2022.9765013.
- [19] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network-a deep learning approach," *Procedia Computer Science*, vol. 132, pp. 679–688, 2018, doi: 10.1016/j.procs.2018.05.069.
- [20] H. Gu, Y. Wang, S. Hong, and G. Gui, "Blind channel identification aided generalized automatic modulation recognition based on deep learning," *IEEE Access*, vol. 7, pp. 110722–110729, 2019, doi: 10.1109/ACCESS.2019.2934354.





- [21] M. Yani, B. Irawan, and C. Setiningsih, "Application of transfer learning using convolutional neural network method for early detection of terry's nail," *Journal of Physics: Conference Series*, vol. 1201, no. 1, May 2019, doi: 10.1088/1742-6596/1201/1/012052.
- [22] Jahandad, S. M. Sam, K. Kamardin, N. N. A. Sjarif, and N. Mohamed, "Offline signature verification using deep learning convolutional neural network (CNN) architectures GoogLeNet inception-v1 and inception-v3," *Procedia Computer Science*, vol. 161, pp. 475–483, 2019, doi: 10.1016/j.procs.2019.11.147.
- [23] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016, doi: 10.1109/TBME.2015.2496264.
- [24] S. Ray, "Disease classification within dermoscopic images using features extracted by ResNet50 and classification through deep forest," *Prepr. arXiv.1807.05711*, Jul. 2018.
- [25] J. Zhuang, J. Cai, R. Wang, J. Zhang, and W.-S. Zheng, "Deep kNN for medical image classification," in *MICCAI 2020: Medical Image Computing and Computer Assisted Intervention-MICCAI 2020*, 2020, pp. 127–136, doi: 10.1007/978-3-030-59710-8_13.

BIOGRAPHIES OF AUTHORS







Jia Rong Leow     has taken his major in Bachelor of Computer Science Artificial Intelligence at the Multimedia University of Malaysia began in 2020. He is presently studying at Multimedia University. He was do his research through the Centre for Ubiquitous Computing and Communication Research Centre (CUCC) at Multimedia University. He conducted research on categorization breast cancer using several types of convolutional neural networks, and he has also developed a hybrid model that combines convolutional neural networks with other machine learning algorithms. He can be contacted at email: leow1017@gmail.com.







Wee How Khoh     received his Master of Science in Information technology and Ph.D degree from Multimedia University. He is also working as a lecturer in the Faculty of Information Science and Technology (FIST), Multimedia University. He is currently the deputy chairperson of the Centre for Ubiquitous Computing and Communication Research Centre (CUCC) at the university. His research interests are machine learning, biometrics authentication and security, and signal and image processing. He can be contacted at email: whkhoh@mmu.edu.my.



Ying Han Pang     received her B.E. (Hons) degree in Electronic Engineering in year 2002, Master of Science degree in year 2005 and Ph.D. degree in year 2013 from Multimedia University. She is currently a lecturer at the Faculty of Information Science and Technology, Multimedia University. Her research interests include biometrics, human activity recognition, machine learning, inertial data processing, image processing and pattern recognition. She can be contacted at email: yhpang@mmu.edu.my.



Hui Yen Yap     is currently a lecturer in Faculty of Information Science and Technology (FIST), Multimedia University. She is a Ph.D. candidate in Computer Science at the Technical University of Malaysia (UTEM), Malacca. She received her Master's Degree in Computer Science in the year 2013. Her current research focuses on user recognition with brainwaves. She can be contacted at email: hyyap@mmu.edu.my.