



ARTICLE

ResNet50-Based Effective Model for Breast Cancer Classification Using Histopathology Images

Nishant Behar* and Manish Shrivastava

Department of Computer Science & Engineering, GGV, Bilaspur, 495009, India

*Corresponding Author: Nishant Behar. Email: nishant.itggvonline@gmail.com

Received: 26 April 2021 Accepted: 11 August 2021

ABSTRACT

Breast cancer is considered an immense threat and one of the leading causes of mortality in females. It is curable only when detected at an early stage. A standard cancer diagnosis approach involves detection of cancer-related anomalies in tumour histopathology images. Detection depends on the accurate identification of the landmarks in the visual artefacts present in the slide images. Researchers are continuously striving to develop automatic machine-learning algorithms for processing medical images to assist in tumour detection. Nowadays, computer-based automated systems play an important role in cancer image analysis and help healthcare experts make rapid and correct inferences about the type of cancer. This study proposes an effective convolutional neural network-based (CNN-based) model that exploits the transfer-learning technique for automatic image classification between malignant and benign tumour, using histopathology images. Resnet50 architecture has been trained on new dataset for feature extraction, and fully connected layers have been fine-tuned for achieving highest training, validation and test accuracies. The result illustrated state-of-the-art performance of the proposed model with highest training, validation and test accuracies as 99.70%, 99.24% and 99.24%, respectively. Classification accuracy is increased by 0.66% and 0.2% when compared with similar recent studies on training and test data results. Average precision and F1 score have also improved, and receiver operating characteristic (RoC) area has been achieved to 99.1%. Thus, a reliable, accurate and consistent CNN model based on pre-built Resnet50 architecture has been developed.

KEYWORDS

Classification; histopathology images; convolutional neural network; breast cancer; breakHis

1 Introduction

The International Agency for Research (IARC), a part of the World Health Organization (WHO), coordinates and conducts both epidemiological and laboratory research into the causes of human cancer; it reported that breast cancer is the most commonly occurring cancer worldwide (11.7% of the total new cases). In the projected burden of cancer in 2040, it was estimated that in 2.3 million new cancer cases worldwide, one in every 8 cancers diagnosed in the year 2020 will turn out to be breast cancer; it is the fifth leading cause of cancer mortality worldwide, with 685,000 deaths in the year 2020. In women, breast cancer ranks first in terms of incidence and mortality in most countries around the world [1]. This high mortality rate can be reduced through timely



detection of the disease, in its early stages. The manual detection method, fine needle aspiration (FNA), is a simple and quick cytology procedure to discover cysts or cancer, by examining the fluid characteristics extracted from the suspected area; however, the diagnosis processes are subject to expert inferences on visual anomalies. This pathological challenge, associated with FNA, encourages researchers to discover a new, automatic, quick and consistent method to analyse and detect cancerous cells. The last decade has witnessed steady progress of the convolutional neural network (CNN), which represented significant contribution towards automated detection of tumour images and provides additional support to healthcare experts, allowing them to make accurate decisions. Therefore, research developments are focused on computer-aided diagnostic (CAD) methods for disease prognosis and diagnosis purposes. In CNN, breast cancer diagnosis is carried out by examining the images acquired through different medical instruments like mammograph, breast ultrasound, magnetic resonance imaging (MRI) or the open biopsy process. A typical CAD system has four common steps: (a) input images, (b) pre-processing, (c) feature extraction and (d) classification; after these steps, results are produced. A specialized artificial neural network (ANN) model—CNN—includes all the above-mentioned steps.

ANN techniques have become popular among researchers due to their excellent learning ability and built-in parameter tuning mechanism. As a result, ANN is widely used to examine theoretical and applied science models. To study and determine the behavior of any model, correct mapping between inputs and the target parameters of the model is an essential step. ANN methods play a key role in modifying crucial internal parameters and providing an optimized network weight. Some recent studies have described the use of an ANN method alone or in conjunction with genetic algorithm (GA), particle swarm optimisation (PSO), whale optimization algorithm and sequential quadratic programming (SQP), to simulate well-known multi-disciplinary equations. In the studies [2–6], Emden-Fowler model for its second and third order systems has been solved successfully. In other studies [7,8], Lane-Emden delay differential model was analysed. Several authors have used conventional mathematical simulations like fractional dynamics behaviours using Chebyshev polynomials converted into a linear equations [9], masses on the vertices of equilateral triangle [10] and modelling of human liver with the Caputo-Fabrizio fractional derivative [11] for their experimental models. Furthermore, researchers [12–16] presented mathematical modelling in some different areas and have claimed that their proposed algorithms can be suitably applied in diverse fields. Moreover, ANN has evolved as a promising technique in the healthcare sector for diagnoses and processing medical images.

In our present study, an ANN technique was implemented for adjusting network weights recursively for minimizing detection errors. In the prior reported work [17], researchers have proposed a dataset and applied a two-class computer-based classification; they obtained accuracy within the ranges of 80% to 85%. With this background, this study aims to propose an effective classifier with high classification rate and accurate detection of malignant and benign tumours from breast histopathology images. Some images of benign and malignant tumours are shown in Fig. 1.

The prime objective of this study is to provide an effective model, with a high accuracy rate, for the classification of cancerous and non-cancerous images. The remaining sections are organised as follows: Section 2 is dedicated to the literature review, Section 3 introduces the CNN concepts, Section 4 describes research methodology with experimental outcomes and Section 5 concludes the study and unfolds further research opportunities.

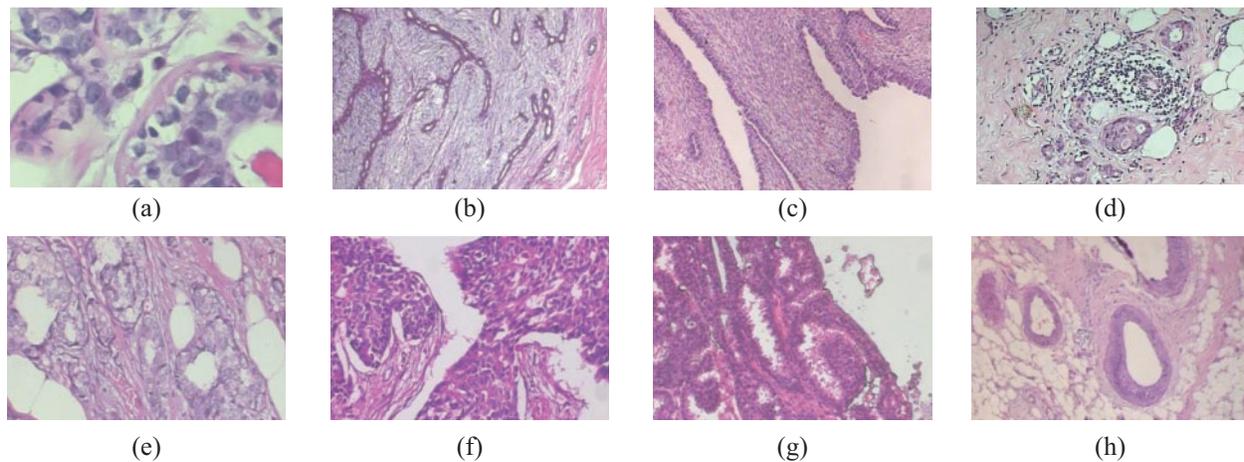


Figure 1: Different breast cancer biopsy images [17]. (a–d) Benign tumour images. (e–h) Malignant tumour images

2 Literature Review

This literature review includes the research papers that used pre-trained systems for breast cancer classification. Some substantial and relevant studies have been discussed in this section.

In a study, researcher have developed a deep feature-based model (DeCAF) [18], for the classification of breast histopathology images, and the pre-trained BVLC CaffeNet model, which is available for free in the café library, was implemented as a feature extractor. Various outputs obtained from different fully connected layers were analysed individually and in different combinations. Logistic regression was used as a base classifier for the evolution of these deep features. The experimental outcome has exhibited a high accuracy rate and better performance than hand-made features for image recognition. In an experimental work [19], a transfer-learning method was employed to classify histopathological images, and a pre-trained CNN was used to encode the local features into Fisher vector (FV). Authors also introduced a new adaptation layer that enhances discriminant characteristics of features and improves classification accuracy. VGG-VD was implemented for accomplishing transfer learning. The study [20] presented a classification scheme to generate and assign confidence score to images of each class. The classification was based on the Kernel principal component analysis (KPCA) ensemble method, and the proposed scheme was successfully tested on two different datasets: (a) breast cancer and (b) optical coherence tomography. This study has also imposed a research challenge to find an optimiser, to reduce the fixed size parameters applied in this work. In the reported work [21], 1.2 million high-resolution images have been categorised in 1,000 different classes through a large, deeper and focused neural network in the ImageNet (LSVRC-2010) competition. This novel model consisted of 60 million parameters, 650,000 neurons and five convolution stages. The fully connected layers applied the SoftMax function. Experiment outcome reported the lowest error rates of 37.5% and 17.0%, respectively. The research work [22] provides an automatic two-step classifier for slide investigation. In the first step, coarse regions were analysed in the whole slide to obtain diverse spatially localised features, after which the detailed analysis of the selected tiled region was completed. This experiment used elastic net classifier and weighted voting classifier for cancer detection on brain image dataset. The research paper [23] presented a probabilistic classifier that combines multiple-instance learning and relational learning. The instance-based learning was used

for the classification task, whereas relationship-learning map was used for the changes in the cell formation due to cancer. This approach had been evaluated on breast and Barrett's cancer tissue microarray datasets. In another work, the researcher had developed an accurate, reliable, and multistage method for measuring the neoplastic nuclei size for pleomorphism grading [24]. In the preliminary stage, image quality, staining value and tissue appearance were observed. In a later stage, machine-learning methods were used to score and remove abnormality that existed in nuclear contour, for better segmentation. A study has discussed difficulties associated with multiclass classification due to broad variability, high coherency, and inhomogeneity of colour distribution. In this paper, authors had resolved these issues and proposed a breast cancer multi-classification method, which used a structured deep-learning model [25]. This model includes a training stage, extracting features and optimises inter-class distances. The validation stage was used for parameter tuning, while the testing stage comprised model evaluation. The final accuracy was obtained as 93.2% for multi-class binary classification on the breast biopsy image dataset. A study has addressed an intra-embedding supervised classification method to classify histopathology images; in this innovative work [26], Fisher encodings are embedded with CNN architecture. The results of the work concluded that the proposed method can successfully resolve the issues of high dimensionality and busy visual components, which are associated with FV. This model was evaluated in lymphoma and breast cancer datasets, where highest classification accuracies on different magnification levels were reported on the breast cancer histopathology dataset. An experimental work [27] was focused on the use of colour-textural characteristics of the breast cancer histopathology images, for effective classification. In this work, different classifiers, namely support vector machine (SVM), decision tree, nearest neighbour and discriminant analysis, were included for creating an ensemble classifier. The proposed model was studied on various optical magnification levels, to acquire better discriminative features. A CNN-based approach was applied in [28], where patch-level classifications were carried out on breast cancer biopsy images. The maximum accuracies reported 77.8% for four classes and 83.3% for two classes, for instance, carcinoma/non-carcinoma through the CNN-SVM model. Another approach [29] had examined 74 features and their importance was inferred, based on pixel level, contrast and texture. Additional features were also studied like geometry and context for significant contribution, to achieve a better accuracy rate. In the novel work [30], researchers have applied CNN for correct detection of smoke in a region from video frames.

Based on earlier research work, it can be summarised that the use of machine-learning methods, especially deep learning algorithm, can be efficiently optimised for cancer cell detection, segmentation and classification. Review of literature also highlighted the importance of input image quality, feature extraction and classifier for development of an effective computer-based system. A lot of contributions have already been made in the recent years, but there is always scope for achieving better classification accuracy, which is the basis of our research work.

3 CNN Architecture

Automatic feature extraction is the prime benefit of using CNN. This concept was originated through experimental work done by Hubel et al. [31] in 1968. They had analysed the cortex responses on different light intensities for simple, complex and hypercomplex patterns. Based on this concept, Fukushima [32] had proposed a "Neocognitron" neural network process model to recognise visual patterns, using the "learning-without-a-teacher" concept. This model applies the novel S-layer and C-layer. The S-layer portrayed geometrical features, and the C-layer was applied to down-step the S-output, by responding selectively for a single-stimulus pattern. CNN

architecture is more complex due to multiple layers, which implement a convolution operation to extract local and global features, to determine the underlying pattern of an image. Extracted features are extremely useful; they guide the network towards the learning direction and lead to automatic decision-making. CNN model has another important layer—pooling layer—which reduces the dimension received from the previous layer. In a single CNN architecture, there can be many combinations of convolution and pooling layers. These intermediate layers are used to identify spatial visual patterns from images, as shown in Fig. 2. These convolution and pooling layers are discussed in Subsections 3.1 and 3.2, respectively.

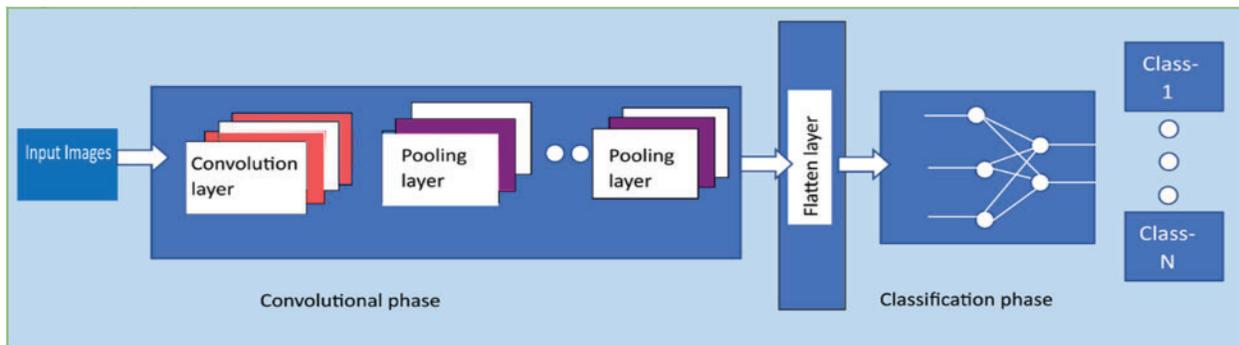


Figure 2: Schematic diagram of typical CNN architecture

3.1 Convolution Layer

The key function of the CNN model is executed by its convolve layer, which significantly identifies the different shapes of the input image, using different filters (or kernels). More specifically, one filter can only detect one pattern; therefore, multiple filters are generally used to determine complete image components. To understand the convolution process, for example, an $n \times n$ size image and an $m \times m$ size filter (such as $m = 1, 3, 5 \dots$ and $m < n$) are fed to a convolution layer such that the filter slides over the image to form an $m \times m$ image region. For this image region, a dot operation takes place between corresponding pixels of the filter and the specific image region. The result of this operation provides most valuable pixel information, which expresses a particular pattern. The mean of all $m \times m$ pixels is calculated, which is stored in a single-cell output matrix (i.e., feature map). This process is repeated until the filter slides on every possible location over the image. Once the whole process is complete, the feature map is generated. Two important methods, stride and padding, also affect the feature map result during this convolution process. A stride 's' determines the value of jump or movement of the filter from current location to next location over the image; generally, the values are $s = 1, 2$ or 3 , whereas padding adds some additional rows and columns in an image with the pixel value = 0. Padding ensures the proper representation of corner and edge pixel values in the feature map. This feature maps are fed to the next layer, as illustrated in Fig. 3.

Eq. (1) is expressing the mathematical operation performed during the convolution process. Suppose $X_{i,j}$ is the height and width of an image region and F_{v_1, v_2} is filter used in a convolution process, where $v_1 = v_2$, filter size is $F_{v_1, v_2} < X_{i,j}$ and $*$ is a dot operator, then each movement of the filter over the particular region of image (i.e., $m \times n$) involves two actions: (a) dot operation: the pixel-wise dot operations are performed between the specific pixels of the selected image region and the corresponding pixel of the applied filter and (b) summation: all resultants are summed together to form a single input value, to contribute a cell value in the feature matrix. Following

expression with bias, provides the complete feature matrix that can be referred [33].

$$\text{Input image}(X_{i,j}) * \text{Filter}(F_{v_1,v_2}) = \left(\sum_{m=-\frac{v_1+1}{2}}^{\frac{v_1-1}{2}} \sum_{n=-\frac{v_2+1}{2}}^{\frac{v_2-1}{2}} X_{i-m,j-n} * F_{m,n} \right) \quad (1)$$

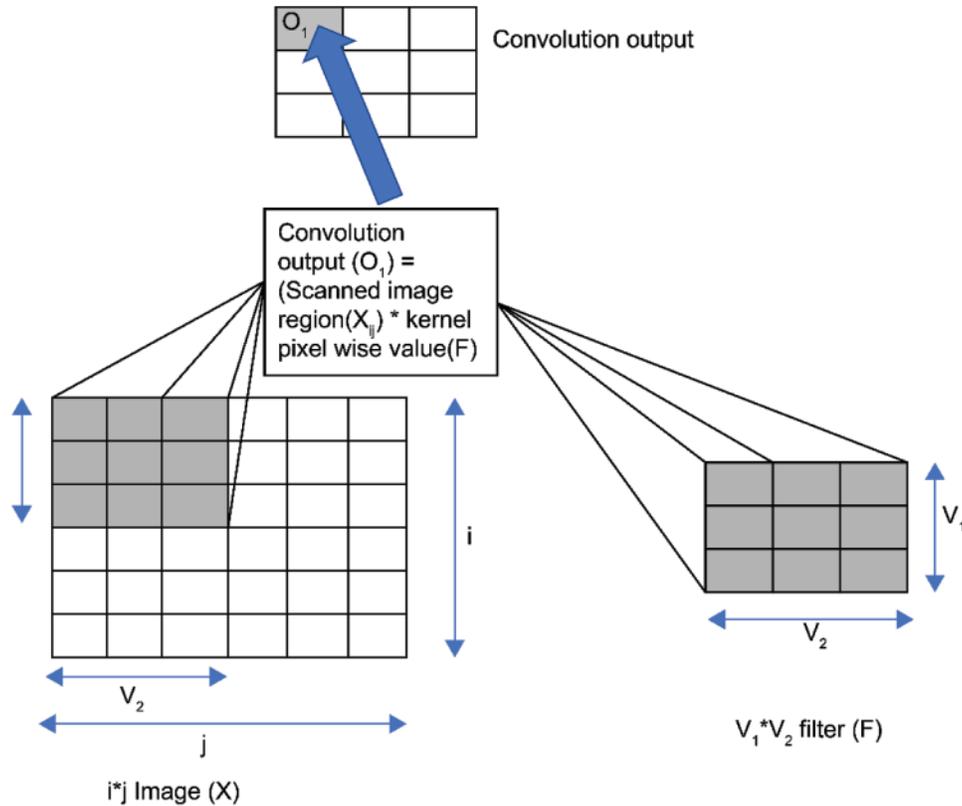


Figure 3: Illustration of convolve process

The size of convolutional layer can be defined as Eq. (2), where input size (R), Filter size (F), padding (p) and stride (s) are used in a particular layer:

$$\text{Feature map output size } (o) = \frac{(R - F + 2p)}{s} + 1 \quad (2)$$

3.2 Pooling Layers

Pooling layer is the crucial next stage; it comprises decreasing the matrix size of the convolution layer. Pooling layers are arranged just after one or more convolution layers so that it can retain the most significant values from feature map. Two popular pooling methods are max pooling and average pooling. The max pooling method contains only the maximum value from a

specific sub matrix (i.e., $m * n$) of the feature map; average pooling collects the mean value of the feature submatrix, which can be referred to in Eq. (3) [34].

$$P_{ijq} = \max_{(m,n) \in F_{i,j}} P_{mnq} \tag{3}$$

Here, $F_{i,j}$ is a complete feature matrix of a particular layer ‘q’, and E_{mnq} are matrix elements on the feature map submatrix ($m*n$). Once this process is complete for all possible subregions of the feature map, the complete downsized output matrix can be obtained for further processing.

3.3 Classification Layer

3.3.1 The Classifier

Classification refers to assigning a single label to an unknown sample in the pool of different classes, whereas a classifier is a mathematical function applied to learn the underlying hidden patterns from the training dataset such that the input data can be mapped with their actual label. Once learning is complete, the classifier is used to predict unlabelled data. A fully connected (FC) ANN is usually applied as a classifier, followed by the flatten layer. This FC layer contains collection of nodes, like human brain neurons, and they are arranged in multiple layers of input, hidden and output. All nodes of one layer are linked with every node of the next layer. The final feature vectors are forwarded to this FC network, and input layers of the FC network are fed forward to the subsequent hidden layer. Every node of each layer has some specific weight; it is multiplied with the prior layer output and added with the bias values of its own layer. The last layer determines the most suitable class for the input image, and each node has its own summation and activation function. The activation function is a mathematical process which decides whether a node value needs to be activated (or triggered) or not for the prediction purpose, based on some predefined threshold value. The most popular activation functions are binary step (B), sigmoid (S), relu (R), tanH (Th) and SoftMax (S_m). The following Table 1 summarises some of the functions, with Eqs. (4)–(8).

Table 1: Various popular activation functions

Activation functions	Formula
Sigmoid	$S(\theta) = \frac{1}{1+e^{-\theta}}$ (4)
Relu	$R = \max(0, \theta)$ (5)
Binary step	$B(\theta) = \begin{cases} 0 & \text{when } \theta < 0 \\ 1 & \text{when } \theta \geq 0 \end{cases}$ (6)
TanH	$Th(\theta) = \left[\frac{2}{1+e^{-2\theta}} \right] - 1$ (7)
Softmax	$S_m(\theta)_i = \frac{e^{\theta_i}}{\sum_{j=1}^k e^{\theta_j}}$ (8)

3.3.2 The Cost Function

The cost function represents the learning ability of a classifier. It represents the gap between predicted and actual class of input dataset; this gap should be minimum. A counter feedback (i.e., backpropagation) process is employed to re-adjust neuron weights of all hidden layers in the opposite direction. After tuning the network weight, the updated cost value is checked, and gaps are analysed again. This learning process is repeated until and unless the optimum value, in the form of global minima, is achieved and the network classifies the maximum correct labels for

given dataset. The cost function can be mathematically expressed as Eq. (9)

$$\text{Cost (C)} = \frac{1}{M} \sum_{i=1}^M (\hat{Y}_i - Y_i)^2 \quad (9)$$

where X is the set of input samples $\{X_1, X_2 \dots X_M\}$ and Y is the corresponding true labels $\{Y_1, Y_2 \dots Y_M\}$ of X . The total number of samples are M and the predicted output label (\hat{Y}_i) for sample 'i' is such that $\hat{Y}_i = f(X_i)$.

4 Research Methodology

4.1 Breast Cancer Dataset

In this study, we have used a BreakHis open dataset, which contains different breast cancer classes in two main categories, benign and malignant. A total of 7,909 images of $40\times$, $100\times$, $200\times$ and $400\times$ magnification levels were obtained from 82 patients. Out of these, 2,480 belonged to the benign and 5,429 to the malignant class. Each image has been represented by an 8-bit RGB channel, with the size of 700×460 pixels. Total number of images, based on different magnification levels, are summarised in the Table 2.

Table 2: Dataset description according to each magnification level

Magnification	Total
$40\times$	1,995
$100\times$	2,081
$200\times$	2,013
$400\times$	1,820
Total of images	7,909

The benign tumour is further sub-classified into adenosis, fibroadenoma, phyllodes and tubular adenoma. Malignant tumour type is sub-categorised into ductal carcinoma, lobular carcinoma, mucinous carcinoma and papillary carcinoma. In this experimental work, all images have been manually partitioned into benign or malignant; this allows the problem to be formulated into binary classification. Data augmentation has also been applied due to lesser number of training data samples. The complete dataset was divided into training (6327 images), validation (790 images) and test subsets (792).

4.2 Proposed Model

In this experimental work, a multi-layered deep CNN framework Resnet50 was tuned for transferring the learning process. ResNet-50 is a prebuilt model which has been trained on the ImageNet dataset for identifying different images of 1,000 classes. ImageNet pre-trained weights were supplied as initial weights for the proposed deep neural network. The residual layers present in ResNet50 plays an important role, to transfer large gradient values to its prior adjacent layers. Due to this layer, the model can effectively extract complex and relevant patterns and resolve the vanishing gradient problem [35,36]. In our experimental setup, all pre-train model layers are kept open to learn new features from biopsy images. The feature matrices, acquired from CNN layers, were supplied to the fine-tuned FC layer, where the sigmoid function was used in the output layer.

Further, the Adam optimiser was applied, with the learning rate 0.00001, for achieving better accuracy. The loss function was set to binary cross entropy. Loss function indicates the difference between the actual and predicted value. Epoch was set to 80. The available data were highly imbalanced in their sample number, so for the data balancing, proper weights for each class were assigned; please refer to Fig. 4. The whole dataset was divided into different subsets as follows: training dataset with 6,327 images to train the model, validation dataset has 790 images for tuning network parameters and test subsets have 792 images to evaluate model performance.

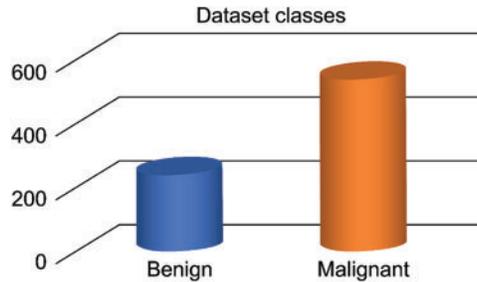


Figure 4: Indicating imbalanced dataset

The data augmentation has been applied additionally to generate flipped, zoom, shearing and scaling training datasets, whereas validation and test datasets have been only scaled. Images have been randomly selected from every category. The proposed architecture delivers an adequate performance with excellent accuracy and other model measures rates. Fig. 5 shows the flow chart of the proposed model.

4.3 Evaluation Parameters and Metrics

The utility of the model is mainly measured by accuracy rate of its classification. The confusion matrix is used to analyse the additional quality parameters like precision, recall and F1 score. A 2 * 2 confusion matrix represents performance measures for binary classification; it is inferred in Fig. 6. Here, a different cell element contains the true positive (TP), false negative (FN), false positive (FP) and true negative (TN) values. The TP refers to correct prediction of true positive class, FP refers to the cases where there is an actual negative class—but it is determined as a positive class—TN is referred as the negative class correctly predicted and FN shows a positive class incorrectly as a negative class.

The accuracy outcome is an important measure for any CNN model. It represents the ratio of total correct prediction (TP + TN) over all predictions (i.e., TN + TN + FP + FN) generated by the model. Accuracy, precision, recall and the F1 score can be calculated from the equations, Eqs. (10)–(13).

$$\text{Accuracy (A)} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \quad (10)$$

$$\text{Precision (P)} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (11)$$

$$\text{Recall or true positive rate (TPR)} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (12)$$

$$F1 \text{ score} = \frac{2 * (\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{13}$$

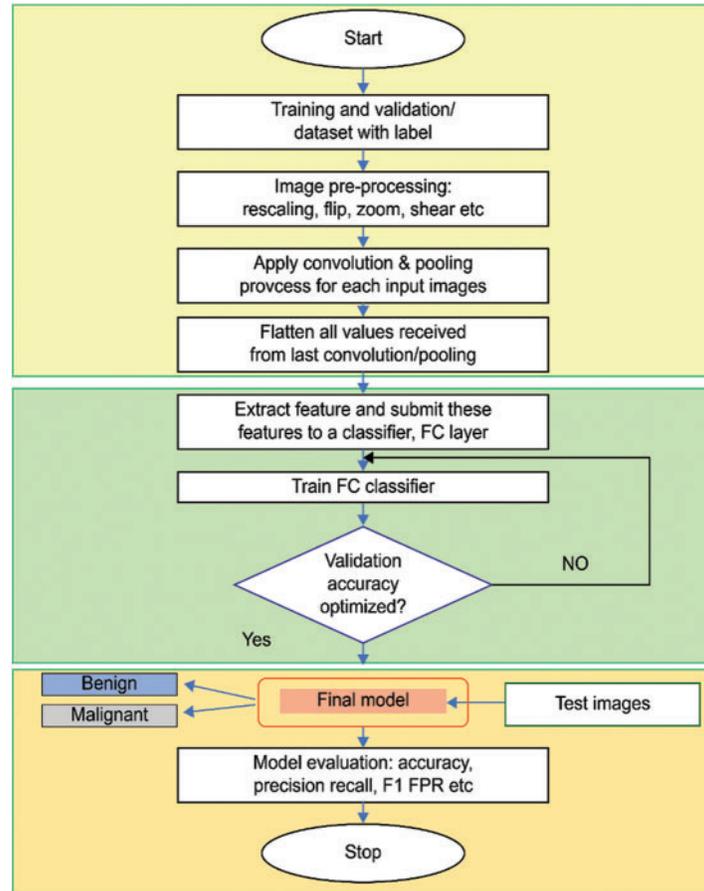


Figure 5: Step-by-step illustration of proposed model process

		Predicted Class	
		Positive Class (0)	Negative Class (1)
Actual classes	Positive Class(0)	True Positive	False Negative
	Negative Class(1)	False Positive	True Negative

Figure 6: Representing 2 * 2 confusion matrix for the binary classification

4.4 Experimental Results

This study was setup in Colab Notebook, offered by the google cloud platform, with a Tesla T4 GPU. Python and Keras, were used for the programming and model-tuning purposes. For performance evaluation, the Sklearn library [37] was employed, and conclusion was drawn from confusion matrices. This study consists of six major steps: (a) image pre-processing and data augmentation, (b) feature extraction, (c) network training and FC layer fine-tuning, (d) model accuracy evaluation using validation dataset, (e) finalisation of model and (f) final performance evaluation on accuracies and other parameters using unseen test data set, as presented in Fig. 5.

Image processing is an important step; it enhances the quality of image and prepares input images for network learning. Data augmentation was necessary due to lesser number of available images, which were further divided into train, validation and test datasets. Feature extraction has been carried out through the prebuilt model, Resnet50. The pre-trained weights were loaded initially, and the network layers have been kept trainable, allowing the model to learn new patterns from the provided unseen medical dataset. Fine-tuning was carried out through redefining the FC layer, and hyper parameters have been tuned by multiple runs for the optimum accuracy on training and validation dataset. Model training epochs were set as 80. Once the training was completed, the final accuracy on the test dataset was obtained. The experimental result projected exemplary training accuracy, 99.70%, and validation accuracy as 99.24%. Fig. 7a illustrates a graphical representation of the achieved accuracies for training and test datasets during 80 epochs. Fig. 7b determines model perfection, through the loss ratios obtained for training and validation as 0.190 and 0.0694, respectively.

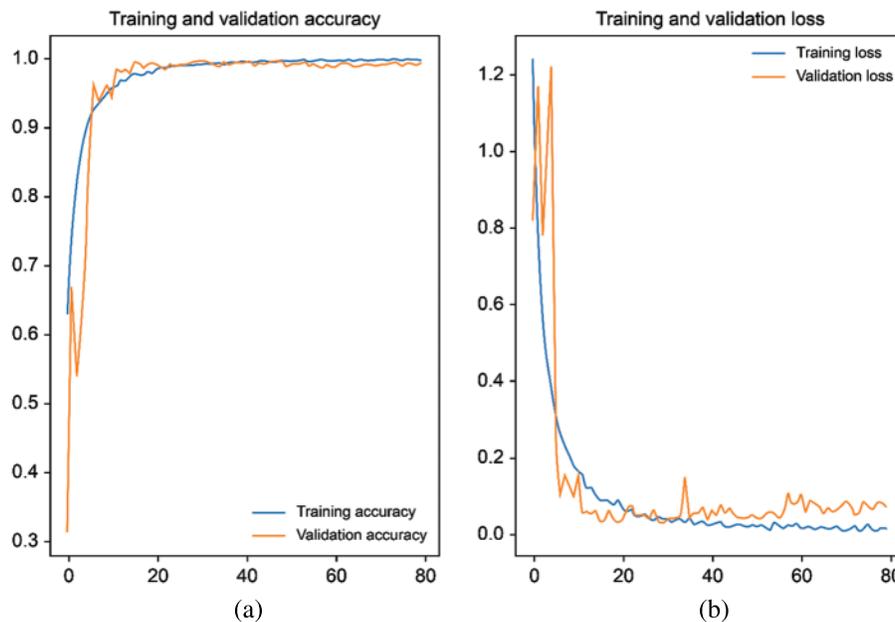


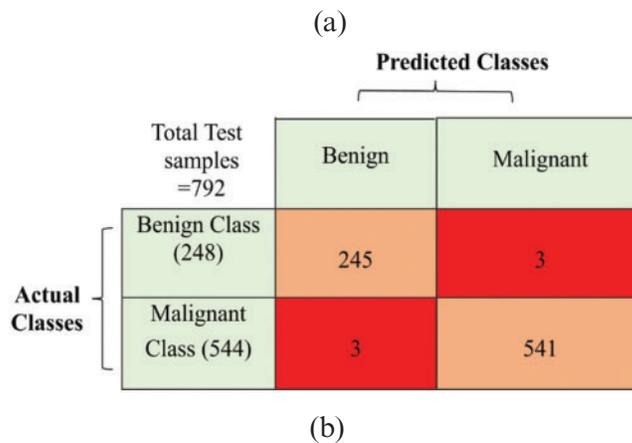
Figure 7: Demonstration of training and validation samples (a) accuracies and (b) the loss values

When the model was evaluated through untouched test dataset, encouraging results were found, with 99.24% accuracy and a loss value of 0.0491. Further, for deeper analysis of the model, confusion matrix and other evaluation parameters, such as precisions, recall, F1 score, were also

carefully observed; this concluded that the model is correctly classifying each sample of different data subsets; refer to [Tables 3a](#) and [3b](#).

Table 3: (a) Classification accuracies and sample data used (b) confusion matrix attributes

Evaluation method	Data-subset	Performance (in %)	Sample size (number of images)
Model accuracies	Training	99.70	6327
	Validation	99.24	790
	Test	99.24	792
ROC	Area	99.1	—



Receiver operating characteristics (RoC) analysis is a well-defined graphical representation to show the trade-off between true positive rate and false positive rate for binary classification tasks [38]. The area under the curve (AUC) is also important to determine model's capability, to distinguish between one class and another. In this experiment, RoC area was found as 99.1% as depicted in [Fig. 8](#). This clearly indicates that the proposed model can classify input images into their correct class.

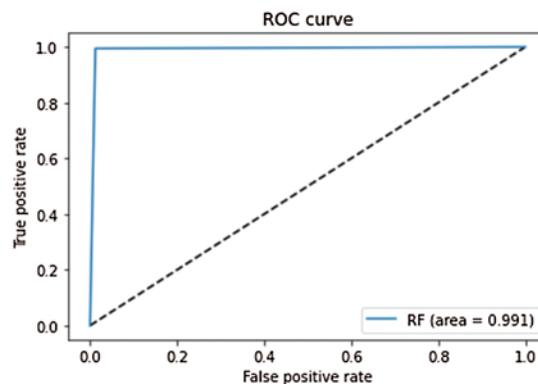


Figure 8: RoC curve and AUC evolution performed on test dataset

The individual class performance with precision, recall and F1-score is shown in [Table 4](#). In this table, benign class is encoded as ‘0’ and malignant class is represented as ‘1’. Parameter ‘Precision’ depicts the proportion of TP values over total samples predicted for positive inclusion of FP—please refer to [Eq. \(11\)](#). Precision value is an important parameter, as it indicates unbiased and precise model performance for positive class predictions. ‘Recall’ parameter also represents the proportion of positive class over sum of predicted TP and FN samples—please refer to [Eq. \(12\)](#). F1-score represents harmonic mean of precision and recall value—please refer [Eq. \(13\)](#). The term ‘Support’ defines the class samples involved in this calculation. The higher values of accuracy, precision, recall and F1 indicate superior model performance.

Table 4: Individual class and average accuracy scores generated on 792 data samples

	Precision	Recall	F1-score	Support
0	99	99	99	248
1	99	99	99	544
Accuracy			99	792
Macro average	99	99	99	792
Weighted average	99	99	99	792

The following [Table 5](#) illustrates the comparative performance of the model in similar studies. In this comparison, only two-class classification on the same BreakHis dataset have been considered. The exemplary accuracy, precision and F1 score rate of our proposed model can be noticed. A comparative account of accuracy of model is depicted in [Fig. 9](#).

Table 5: Comparative result of our proposed model with similar recent studies [39]

Paper	Model used	Accuracy	Precision	Recall	F1 score
[39]	CNN (image levels)	89.6	–	–	–
[18]	CNN with DeCAF features (image Level)	83	–	–	–
[40]	CNN-CT histogram (200X)	97.19	98	98.2	98
[41]	Dense net features and SVM	97.96	–	–	–
[42]	Fine-tuned Resnet-18 (magnification independent)	98.42	98.75	99.01	98.88
[43]	CNN with ensemble	99.04	98.94	99.65	–
Proposed model	Training dataset	99.70			
	Validation dataset	99.24	–	–	–
	Test dataset	99.24	99.12	99.12	99.12

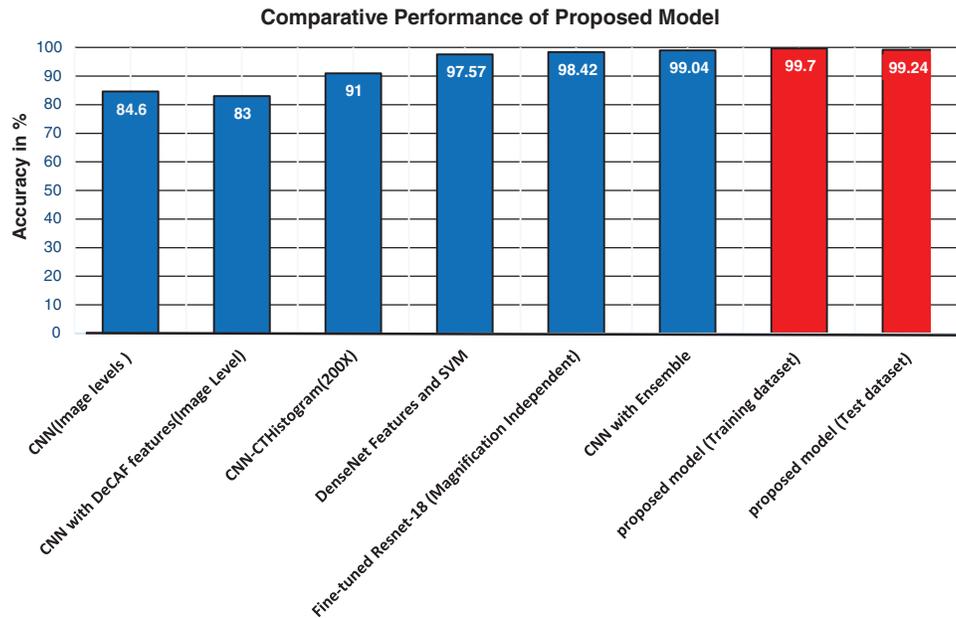


Figure 9: Illustration of accuracy performance of proposed model with latest similar work

5 Conclusion and Future Work

This study discussed the general functioning of the CNN model and proposed a model that was effective in the classification of cancerous and non-cancerous histopathology images taken from the Breakhis dataset. It was quite a challenging task to extract the prominent features due to variability in H&E stain, colour intensities, irregular shapes, high resolution and large sizes of images. Manual analysis of these images are always hard and time-consuming and are subject to the expertise level of specialists. Highly unbalanced and smaller sample images were another associated problem in the study; however, necessary steps, such as data augmentation and proper class weights, were taken to resolve these issues. The strength of our proposed model is based on the network's learning ability and accurate classification capabilities of the BreakHis dataset. The extracted feature also adequately determines the complex local and global features of histopathology images, and supports the redesigned FC layer learning. For the evaluation of the proposed model, only calculating the accuracy parameter is not enough. Therefore, other performance measures—RoC, precision, recall and F1 score—have also been calculated; this allows true and reliable detection of image classification. The results were encouraging when compared with similar recent works. The regularisation and dropout values can be tuned further for better results; however, the testing results adequately prove that the proposed model is outperforming previous models, and it is appropriate for providing preliminary assistance to the healthcare experts for quick and accurate decision-making. Thus, a reliable and consistent CNN model based on pre-built Resnet50 architecture, which exploits the transfer learning technique, has been developed.

In future, this work will be extended to study one or more model parameters and their optimisation, like network complexity, image channels, computational time and space complexity. Furthermore, experiments for the multiclass classification on same dataset will/can also be carried out.

Acknowledgement: The authors would like to acknowledge, VRI Lab (Spanhol et al.) for providing the Breast Cancer Histopathological Image Dataset on request, for this academic and research purpose.

Funding Statement: The authors received no funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. WHO (2020). Latest global cancer data: Cancer burden rises to 19.3 million new cases. <https://www.iarc.who.int/news-events/latest-global-cancer-data-cancer-burden-rises-to-19-3-million-new-cases-and-10-0-million-cancer-deaths-in-2020/>.
2. Sabir, Z., Amin, F., Pohl, D., Guirao, J. L. G. (2020). Intelligence computing approach for solving second order system of emden–Fowler model. *Journal of Intelligent & Fuzzy Systems*, 38(6), 7391–7406. DOI 10.3233/JIFS-179813.
3. Guirao, J. L. G., Sabir, Z., Saeed, T. (2020). Design and numerical solutions of a novel third-order nonlinear emden–Fowler delay differential model. *Mathematical Problems in Engineering*, 2020, 1–9. DOI 10.1155/2020/7359242.
4. Zhang, Y., Lin, J., Hu, Z., Khan, N. A., Sulaiman, M. (2021). Analysis of third-order nonlinear multi-singular emden–Fowler equation by using the LeNN-WOA-NM algorithm. *IEEE Access*, 9, 72111–72138. DOI 10.1109/ACCESS.2021.3078750.
5. Moaaz, O., Ramos, H., Awrejcewicz, J. (2021). Second-order Emden–Fowler neutral differential equations: A new precise criterion for oscillation. *Applied Mathematics Letters*, 118, 107172. DOI 10.1016/j.aml.2021.107172.
6. Omid, M., Arab, B., Rasanan, A. H. H., Rad, J. A., Parand, K. (2021). Learning nonlinear dynamics with behavior ordinary/partial/system of the differential equations: Looking through the lens of orthogonal neural networks. *Engineering with Computers*. DOI 10.1007/s00366-021-01297-8.
7. Abdelkawy, M. A., Sabir, Z., Guirao, J. L. G., Saeed, T. (2020). Numerical investigations of a new singular second-order nonlinear coupled functional Lane–Emden model. *Open Physics*, 18(1), 770–778. DOI 10.1515/phys-2020-0185.
8. Sabir, Z., Guirao, J. L. G., Saeed, T. (2021). Solving a novel designed second order nonlinear Lane–Emden delay differential model using the heuristic techniques. *Applied Soft Computing*, 102, 107105. DOI 10.1016/j.asoc.2021.107105.
9. Singh, H. (2021). Numerical simulation for fractional delay differential equations. *International Journal of Dynamics and Control*, 9(2), 463–474. DOI 10.1007/s40435-020-00671-6.
10. Baleanu, D., Ghanbari, B., Asad, H., Jajarmi, J., Mohammadi Pirouz, A. H. (2020). Planar system-masses in an equilateral triangle: Numerical study within Fractional calculus. *Computer Modeling in Engineering & Sciences*, 124, 953–968. DOI 10.32604/cmescs.2020.010236.
11. Baleanu, D., Jajarmi, A., Mohammadi, H., Rezapour, S. (2020). A new study on the mathematical modelling of human liver with caputo–Fabrizio fractional derivative. *Chaos, Solitons & Fractals*, 134, 109705. DOI 10.1016/j.chaos.2020.109705.
12. Yakut, O. (2021). Implementation of hydraulically driven barrel shooting control by utilizing artificial neural networks. *Mathematics and Computers in Simulation*, 190, 1206–1223. DOI 10.1016/j.matcom.2021.03.025.
13. Raja, M. A. Z., Manzar, M. A., Shah, S. M., Chen, Y. (2020). Integrated intelligence of fractional neural networks and sequential quadratic programming for Bagley–Torvik systems arising in fluid mechanics. *Journal of Computational and Nonlinear Dynamics*, 15(5), 051003. DOI 10.1115/1.4046496.
14. Petrosov, D. A., Lomazov, V. A., Petrosova, N. V. (2021). Model of an artificial neural network for solving the problem of controlling a genetic algorithm using the mathematical apparatus of the theory of petri nets. *Applied Sciences*, 11(9), 3899. DOI 10.3390/app11093899.

15. Zhao, Y., Dong, S., Jiang, F. (2021). Reliability analysis of mooring lines for floating structures using ANN-BN inference. *Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, 235(1), 236–254. DOI 10.1177/1475090220925200.
16. Chakraverty, S., Mall, S. (2020). Single layer chebyshev neural network model with regression-based weights for solving nonlinear ordinary differential equations. *Evolutionary Intelligence*, 13, 687–694. DOI 10.1007/s12065-020-00383-y.
17. Spanhol, F. A., Oliveira, L. S., Petitjean, C., Heutte, L. (2016). A dataset for breast cancer histopathological image classification. *IEEE Transactions on Biomedical Engineering*, 63(7), 1455–1462. DOI 10.1109/TBME.2015.2496264.
18. Spanhol, F. A., Oliveira, L. S., Cavalin, P. R., Petitjean, C., Heutte, L. (2017). Deep features for breast cancer histopathological image classification. *IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1868–1873. IEEE, Banff, AB. DOI 10.1109/SMC.2017.8122889.
19. Song, Y., Zou, J. J., Chang, H., Cai, W. (2017). Adapting fisher vectors for histopathology image classification. *IEEE 14th International Symposium on Biomedical Imaging*, pp. 600–603. IEEE, Melbourne, Australia. DOI 10.1109/ISBI.2017.7950592.
20. Zhang, Y., Zhang, B., Coenen, F., Xiao, J., Lu, W. (2014). One-class kernel subspace ensemble for medical image classification. *EURASIP Journal on Advances in Signal Processing*, 2014(1), 1–13. DOI 10.1186/1687-6180-2014-17.
21. Krizhevsky, A., Sutskever, I., Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. DOI 10.1145/3065386.
22. Barker, J., Hoogi, A., Depeursinge, A., Rubin, D. L. (2016). Automated classification of brain tumor type in whole-slide digital pathology images using local representative tiles. *Medical Image Analysis*, 30, 60–71. DOI 10.1016/j.media.2015.12.002.
23. Kandemir, M., Zhang, C., Hamprecht, F. A. (2014). Empowering multiple instance histopathology cancer diagnosis by cell graphs. In: Golland, P., Hata, N., Barillot, C., Hornegger, J. and Howe, R. (Eds.), *Medical image computing and computer-assisted intervention—MICCAI 2014*, pp. 228–235. Springer International Publishing: Cham. DOI 10.1007/978-3-319-10470-6_29.
24. Cosatto, E., Miller, M., Graf, H. P., Meyer, J. S. (2008). Grading nuclear pleomorphism on histological micrographs. *19th International Conference on Pattern Recognition*, pp. 1–4. Tampa, FL, USA. DOI 10.1109/ICPR.2008.4761112.
25. Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K. et al. (2017). Breast cancer multi-classification from histopathological images with structured deep learning model. *Scientific Reports*, 7(1), 1–10. DOI 10.1038/s41598-017-04075-z.
26. Song, Y., Chang, H., Huang, H., Cai, W. (2017). Supervised intra-embedding of fisher vectors for histopathology image classification. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D. L. and Duchesne, S. (Eds.), *Medical image computing and computer assisted intervention—MICCAI 2017*, pp. 99–106. Springer International Publishing: Cham. DOI 10.1007/978-3-319-66179-7_12.
27. Gupta, V., Bhavsar, A. (2017). Breast cancer histopathological image classification: Is magnification important? *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 17–24. IEEE, Honolulu, HI, USA. DOI 10.1109/CVPRW.2017.107.
28. Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P. et al. (2017). Classification of breast cancer histology images using convolutional neural networks. *PLoS One*, 12(6), e0177544. DOI 10.1371/journal.pone.0177544.
29. Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C. I. et al. (2017). Large scale deep learning for computer aided detection of mammographic lesions. *Medical Image Analysis*, 35, 303–312. DOI 10.1016/j.media.2016.07.007.
30. Matlani, P., Shrivastava, M. (2019). Hybrid deep VGG-NET convolutional classifier for video smoke detection. *Computer Modeling in Engineering & Sciences*, 119(3), 427–458. DOI 10.32604/cmcs.2019.04985.
31. Hubel, D. H., Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195(1), 215–243. DOI 10.1113/jphysiol.1968.sp008455.

32. Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36, 193–202. DOI 10.1007/BF00344251.
33. Nahid, A. A., Mehrabi, M. A., Kong, Y. (2018). Histopathological breast cancer image classification by deep neural network techniques guided by local clustering. *BioMed Research International*, 2018, 1–20. DOI 10.1155/2018/2362108.
34. Williams, T., Li, R. (2018). Wavelet pooling for convolutional neural networks. *International Conference on Learning Representations*, pp. 12. BC, Canada.
35. Habibzadeh Motlagh, M., Jannesari, M., Rezaei, Z., Totonchi, M., Baharvand, H. (2018). Automatic white blood cell classification using pre-trained deep learning models: ResNet and inception. *Tenth International Conference on Machine Vision*, pp. 105. SPIE: Vienna, Austria. DOI 10.1117/12.2311282.
36. Ioffe, S., Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *Proceedings of the 32nd International Conference on International Conference on Machine Learning*, pp. 448–456. Lille, France.
37. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B. et al. (2011). Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*, 12, 2825–2830.
38. He, X., Metz, C. E., Tsui, B. M. W., Links, J. M., Frey, E. C. (2006). Three-class ROC analysis—a decision theoretic approach under the ideal observer framework. *IEEE Transactions on Medical Imaging*, 25(5), 571–581. DOI 10.1109/TMI.2006.871416.
39. Spanhol, F. A., Oliveira, L. S., Petitjean, C., Heutte, L. (2016). Breast cancer histopathological image classification using convolutional neural networks. *International Joint Conference on Neural Networks*, pp. 2560–2567. IEEE, Vancouver, Canada. DOI 10.1109/IJCNN.2016.7727519.
40. Nahid, A. A., Kong, Y. (2018). Histopathological breast-image classification using local and frequency domains by convolutional neural network. *Information*, 9(1), 19. DOI 10.3390/info9010019.
41. Alinsaif, S., Lang, J. (2020). Histological image classification using deep features and transfer learning. *17th Conference on Computer and Robot Vision*, pp. 101–108. IEEE, Ottawa, Canada. DOI 10.1109/CRV50864.2020.00022.
42. Boumaraf, S., Liu, X., Zheng, Z., Ma, X., Ferkous, C. (2021). A new transfer learning based approach to magnification dependent and independent classification of breast cancer in histopathological images. *Biomedical Signal Processing and Control*, 63, 102192. DOI 10.1016/j.bspc.2020.102192.
43. Bardou, D., Zhang, K., Ahmad, S. M. (2018). Classification of breast cancer based on histology images using convolutional neural networks. *IEEE Access*, 6, 24680–93. DOI 10.1109/ACCESS.2018.2831280.