Tech Science Press

Optimal Deep Transfer Learning Model for Histopathological Breast Cancer Classification

Mahmoud Ragab[1,2,3,*] and Alaa F. Nahhas[4]

[1]Information Technology Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia
[2]Center for Artificial Intelligence in Precision Medicines, King Abdulaziz University, Jeddah, 21589, Saudi Arabia
[3]Mathematics Department, Faculty of Science, Al-Azhar University, Naser City, 11884, Cairo, Egypt
[4]Biochemistry Department, Faculty of Science, King Abdulaziz University, Jeddah, 21589, Saudi Arabia
*Corresponding Author: Mahmoud Ragab. Email: mragab@kau.edu.sa

Abstract: Earlier recognition of breast cancer is crucial to decrease the severity and optimize the survival rate. One of the commonly utilized imaging modalities for breast cancer is histopathological images. Since manual inspection of histopathological images is a challenging task, automated tools using deep learning (DL) and artificial intelligence (AI) approaches need to be designed. The latest advances of DL models help in accomplishing maximum image classification performance in several application areas. In this view, this study develops a Deep Transfer Learning with Rider Optimization Algorithm for Histopathological Classification of Breast Cancer (DTLRO-HCBC) technique. The proposed DTLRO-HCBC technique aims to categorize the existence of breast cancer using histopathological images. To accomplish this, the DTLRO-HCBC technique undergoes pre-processing and data augmentation to increase quantitative analysis. Then, optimal SqueezeNet model is employed for feature extractor and the hyperparameter tuning process is carried out using the Adadelta optimizer. Finally, rider optimization with deep feed forward neural network (RO-DFFNN) technique was utilized employed for breast cancer classification. The RO algorithm is applied for optimally adjusting the weight and bias values of the DFFNN technique. For demonstrating the greater performance of the DTLRO-HCBC approach, a sequence of simulations were carried out and the outcomes reported its promising performance over the current state of art approaches.

Keywords: Breast cancer; histopathological images; machine learning; biomedical analysis; deep learning; computer vision

1 Introduction

Cancer is a crucial healthcare problem around the world. Amongst the cancer kind, breast cancer occurrence rate is the 2nd maximum for women, apart from lung cancer. Additionally, the death

rate of breast tumors is advanced than other kinds of tumors [1]. Even with the fast development in medication, the study of histopathological image (HI) remains the most commonly utilized technique for breast cancer diagnoses. In each HI analysis task, the predominant is the classifier process. Since the automated and accurate classifier of higher-resolution HI is the bottleneck and cornerstone of other researches, like gland segmentation, nuclei localization, and mitosis detection [2]. At present, histopathological imaging in medical settings is mostly depending on the manual qualitative analysis of pathologists. But a minimum three problem arises from this analysis technique [3]. Firstly, there is a deficiency of pathologists, particularly in small hospitals and lesser developed areas. This unbalanced distribution and resource shortage is a crucial issue to be resolved. Next, whether the histopathological diagnoses are accurate or incorrect depends fully on the long-term accumulated diagnostic experience and pathologist philosophical professional knowledge [4]. This pathologist subjectiveness has resulted in propagation of diagnosis inconsistency. Then, the complication of the HI makes pathologists prone to inattention and fatigue. Facing this problem, it is crucial to design precise and automatic HI analytical methods, particularly classification method, to mitigate this problem [5].

But the study of HI is a time-consuming and difficult process that needs professional knowledge. Moreover, the result might be affected by the experience level of the pathologist included [6]. Thus, computer-aided analysis of HI plays a considerable part in the diagnoses of breast cancer and its prognoses. But, the procedure of designing tool for carrying out this study is obstructed by the succeeding problem. Firstly, HI of breast cancer is fine-grained, higher-resolution image that depicts complex texture and rich geometric structure [7]. The consistency between classes and the variability with a class can create classifier very hard, particularly while handling various classes. The next problem is the limitation of feature extraction method for HI of breast cancer. Deep learning (DL) techniques has the ability to learn abstract illustrations of information, extract features and retrieve data from data automatically [8]. They could resolve the problem of conventional feature extraction and be effectively employed in biomedical science, computer vision, etc.

Wei et al. [9] estimate the efficiency of Auxiliary Classifier Generative Adversarial Network (ACGAN) augmentation approach by implementing breast cancer HI classifier trained on improved dataset. For this classification, we utilize a transfer learning (TL) method wherein the convolution feature is extracted from a pertained system and next given into extreme gradient boosting (XGBoost) classifier. Gheshlaghi et al. [10] presented a hybrid convolution and recurrent deep neural network (DNN) for breast cancer HI classification. According to the rich multi-level feature illustration of the HI, the technique incorporates the advantage of convolution neural network (CNN), recurrent neural network (RNN), and long short term memory (LSTM) spatial relations among patches are conserved. Yan et al. [11] propose an attention high-order deep network (AHoNet) by instantaneously embedding attention model and higher-order statistical illustration as to residual convolution system. Firstly, AHoNet uses an effective network attention model using local cross-channel communication and non-dimension reduction to accomplish local salient deep features of breast cancer HI. Zou et al. [12], proposed a HI classification technique based on cross-domain transfer learning and multistage feature fusion (Cd-dtffNet). Firstly, a system using RL is developed that could extract features from various stages. Next, the capability to extract global features are learned by traveling from the bridge field to the target field, and the capability to extract local features are learned through travelling from the source field to bridge field.

Wang et al. [13] introduce a breast tumor HI classification by assembling compact CNN. Firstly, a hybrid CNN method is developed that comprises global and local model branch. With local voting and two-branch data integration, the hybrid method attains strong representative capability. Next, by embedded the presented Squeeze-Excitation-Pruning (SEP) as to hybrid method, the network

significance is learned and the unwanted channel is eliminated. Zhu et al. [14] propose a CNN that involves a convolution layer, smaller (SE)-ResNet model, and fully connected layer. Then, presented a smaller SE-ResNet model that is a development on the integration of residual model and SE block, and accomplishes the similar efficiency with less parameter. Additionally, presented a learning rate scheduler that could attain outstanding results without finetuning the learning rate.

This study develops a Deep Transfer Learning with Rider Optimization Algorithm for Histopathological Classification of Breast Cancer (DTLRO-HCBC) technique. The proposed DTLRO-HCBC technique undergoes pre-processing and data augmentation to increase quantitative analysis. In addition, optimal SqueezeNet model is employed for feature extractor and the hyperparameter tuning process is carried out using the Adadelta optimizer. Besides, rider optimization with deep feed forward neural network (RO-DFFNN) technique was utilized for breast cancer classification (BCC). For demonstrating the greater performance of the DTLRO-HCBC system, a sequence of simulations was carried out.

2 The Proposed Model

This study has developed a new DTLRO-HCBC system for BCC and detection. The proposed DTLRO-HCBC technique aims to categorize the existence of breast cancer using histopathological images. The DTLRO-HCBC technique comprises a sequence of processes namely pre-processing, SqueezeNet based feature extraction, Adadelta hyperparameter optimizer, DFFNN classifier, and RO parameter optimization.

2.1 Stage I: Feature Extraction

Next to image pre-processing, the optimal SQNet model is employed as feature extractor. SQNet is a convolutional system that outperformed AlexNet when utilizing 50x parameterss [15,16]. It has 18 layers, involving max-pooling (MXP), Global Average Pooling (GAP), convolutional layer, softmax output layer, and fire levels. The input has 227×227-dimension images and contains RGB channels. Convolution is utilized for generalizing the max pooling and input image. Since the positive module, all the convolutional layer implements element-wise activation. Amongst the convolutional layers, SQNet utilizess the fire layers that aree composed of squeeze and expansion stages. The squeezing procedure decreases the depth, while the expanding phase rises when preserving a similar feature size. At last, with the merging process, the expanding output is arranged in the dimension of the input.

Lastly, the resulting f(y) of the squeezing procedure with the feature maps (FM), kernel (W), and C specifies channel of tensor correspondingly:

$$f(y) = \sum_{fm=1}^{FM} \sum_{c=1}^{C} W_c^f * x_c^{fm} \qquad (1)$$

whereas f(y) denotes an output $\in R^N$. Assume X_i indicate input with size of $(W_i, H_i, C_i) \in RN$ of layer i. W represents the weight, H determines the height, and C indicates the channel correspondingly. The MXP layer and spatial dimension execute downsampling in the system and the GAP, that convert the feature map class as to individual value. Finally, Softmax activation function produces multiclass probability distribution.

The hyperparameters involved in the SQNet model can be modified by the use of Adadelta optimizer. Adadelta [17] is an extension of Adagrad to observe reducing learning rate. The running average $E[g^2]t$ at time step t later only depending on the preceding average and the existing gradient.

$$E[g^2] = \gamma E[g^2] t - 1 + (1 - \gamma) g_t^2 \tag{2}$$

Next set γ to same value as momentum, about 0.9. For clearness, rewrite vanilla stochastic gradient descent (SGD) upgrade interms of the variable upgraded vector $\triangle \theta_t$:

$$\triangle \theta_t = - \alpha . g_{t,i} \tag{3}$$

$$\triangle \theta_{t+1} = \theta_t + \delta \theta_t \tag{4}$$

The variable upgrade vector of Adagrad is previously acquired from the following:

$$\triangle \theta_t = - \frac{\alpha}{\sqrt{G_t + \varepsilon}} \odot g_t \tag{5}$$

While the denominator represents the root mean squared (RMS) error condition of the gradient, then substitute with the standard short-hand:

$$\triangle \theta_t = - \frac{\alpha}{RMS[g]_t} \tag{6}$$

The author noted that the unit in this upgrade (SGD, Momentum, or Adagrad) must have a similar hypothetical unit as the variable. To understand this, firstly determine exponentially decaying average, squared gradient, and parameter upgrades.

$$E[\triangle \theta^2]_t = \gamma E \triangle \theta_{t-1}^2 + (1 - \gamma) \triangle \theta_t^2 \tag{7}$$

The root means squared error of variable upgrades as follows:

$$RMS[\triangle \theta]_t = \sqrt{E[\triangle \theta^2]_t + \varepsilon} \tag{8}$$

Replace the learning rate in the preceding upgrade rule with the RMS of variable upgrade. lastly produces the Adadelta upgrade rules:

$$\triangle \theta_t = - \frac{RMS[\triangle \theta]}{RMS[g]_t} \tag{9}$$

$$\theta_{t+1} = \theta_t + \triangle \theta_t \tag{10}$$

Using Adadelta, they don't want to set default learning rate, since it was removed from the upgrade rules.

2.2 Stage II: Image Classification

At the time of BCC, the features are passed into the DFFNN model to categorize the images. Assume the DFFNN encompasses of single input layer, m hidden layers, and single output layer [18]. The input layer states input parameter $x \in \mathcal{X}$ and feeds a vector. Fig. 1 depicts the framework of DFFNN.

$$f^{(0)}(x) = [h_1^{(0)}(x), \ \ldots, \ h_{k_0}^{(0)}(x)]^T \in \mathbb{R}^{k_0} \tag{11}$$

to the primary hidden layer, whereas k_0 denotes the amount of layer in the input neuron and $h_j^{(0)} : \mathcal{X} \mapsto \mathbb{R}$ indicates the activation function of j-th layer in the input layer. For each $i = 1, \ldots, m$, the i-th hidden neuron declare $f^{(i-1)}(x) \in \mathbb{R}^{k_{i-1}}$ from the preceding layer and constructed a vector $f^{(i)}(x) \in \mathbb{R}^{k_i}$, widely named a feature, as follows

$$f^{(i)}(x) = [h^{(i)}\left(w_1^{(i)T}f^{(i-1)}(x) + b_1^{(i)}\right), \ldots, h^{(i)}\left(w_{k_i}^{(i)T}f^{(i-1)}(x) + b_{k_i}^{(i)}\right)]^T, \tag{12}$$

whereas k_i indicates the amount of layers in the i-th hidden state, $h^{(i)} : \mathbb{R} \mapsto \mathbb{R}$ shows the activation function in i-th hidden state, $w_j^{(i)} \in \mathbb{R}^{k_{i-1}}$ and $b_j^{(i)} \in \mathbb{R}$ indicates the weight vector and bias of $j-th$ layer in the $i-th$ hidden state, correspondingly. Signify $W^{(i)} = \left[w_1^{(i)}, \ldots, w_{k_i}^{(i)}\right]$ and $b^{(i)} = \left[b_1^{(i)}, \ldots, b_{k_i}^{(i)}\right]^T$, (12) is formulated by

$$f^{(i)}(x) = h^{(i)}\left(W^{(i)T}f^{(i-1)}(x) + b^{(i)}\right), \tag{13}$$

whereas $h^{(i)} : \mathbb{R}^{k_i} \mapsto \mathbb{R}^{k_i}$ shows a vector-valued function. For notational simplicity, consider $k = k_m$ and $f(x) = f^{(m)}(x)$. The output neuron states $f(x) \in \mathbb{R}^k$ from the previous hidden layer and produced an output vector $a(x) \in \mathcal{H}^n$, named an action, as follows

$$a(x) = h\left(W^Tf(x) + b\right)$$
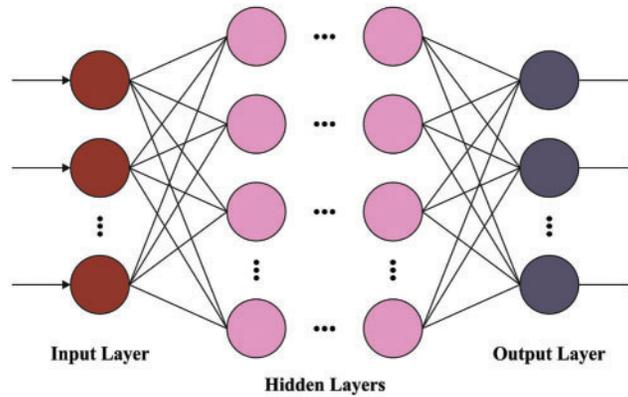$$= [h\left(w_1^Tf(x) + b_1\right), \cdots, h\left(w_n^Tf(x) + b_n\right)]^T, \tag{14}$$

Figure 1: DFFNN structure

Her n shows the amount of neurons, $h : \mathbb{R} \mapsto \mathcal{H}$ indicates the activation function in the output layer, $\mathcal{H}$ indicates the image set of h with $\mathcal{H} \subseteq \mathbb{R}$, $w_j \in \mathbb{R}^k$ and $b_j \in \mathbb{R}$ denotes the weight vector and bias of the j-th layer in the output layer, correspondingly, $W = [w_1, \ldots, w_n]$ and $b = [b_1, \ldots, b_n]^T$.

2.3 Stage III: Parameter Optimization

To effectually tune the parameters presented in the DFFNN technique, the RO approach has been utilized. The inspiration for emerging ROA [19] is attained from the race occurring amongst the collection of riders to accomplish the target. Now, attacker, bypass rider, follower, and overtaker are the 4 classes that are considered in ROA. The overall amount of riders is splitted into 4 classes of riders. All the groups of riders pose the winning approach to achieve the aim. By avoiding the bypass rider neglects the primary route to accomplish the target. Now, the followers follow the route of primary rider. According to the primary rider, the overtaker proceeds in their individual location to attain the target. The maximal speed is employed by the attackers to accomplish the target position by considering the rider's location. In the following, seven steps included in ROA are mentioned in the following.

1. Rider and group parameters Initialization: here, the follower, bypass rider, attacker, and overtaker are collectively represented as E and the position of rider is arbitrarily initialized.

The group initialization can be expressed by

$$S_t = \{S_t(u, v)\};\ 1 \le u \le D;\ 1 \le v \le R \tag{15}$$

Here, the overall amount of riders included in the competition is represented as D. The dimension or number of co-ordinates is denoted as R, and at time t the uth location of the rider is signified as $S_t uv$. The amount of riders contributing to the competition is defined by the overall amount of riders existing in all the groups as follows.

$$D = M + L + Q + T \tag{16}$$

The M, L, Q and T variables are utilized for representing the overall amount of bypass rider, follower, overtaker, and attacker.

Also, the rider variables namely gear, steering, brake, and accelerator are initialized afterward the group initialization [20]. Now, the parameter A is utilized for denoting steering direction. It can be expressed as follows

$$A_t = \left\{A_{u,v}^t\right\};\ 1 \le u \le D;\ 1 \le v \le R \tag{17}$$

Now, the steering direction of uth rider vehicles are represented as $A_{u,v}^t$. The steering initialization direction while the time is zero is shown below

$$A_{u,v} = \begin{cases} \theta_j; & if\ v = 1 \\ A_{u,v-1} + \phi; & if\ v \ne 1 \& L_{u_j v-1} + \phi \le 360 \\ A_{u,v-1} + \phi \le 360; & otherwise \end{cases} \tag{18}$$

The θ_i and ϕ are u^{th} rider vehicle location and coordinate angles.

2. To describe the success rate: All the riders should be described according to the distance when the group initialization and rider parameter is completed.

$$w_u = \frac{1}{||S_u - H_A||} \tag{19}$$

Here, the uth rider location is denoted as S_u and the location of target is shown as H_A.

3. To determine the primary rider: while determining the primary rider, the success rate served a critical role. The primary rider isn't constant since the rider location differs at all time limits. Thus, depending on the success rate, rider becomes the primary rider.

4. To upgrade the location of the rider: The rider position is upgraded for discovering the primary rider that is winner in all the sets.

   i) Upgrade location of bypass rider: while the bypass rider neglects the conventional route utilized by another rider, the location Upgrade can be performed in an arbitrary way for the bypass rider as in the following

$$S_{t+1}^M (u,\ v) = \delta[S_t(\eta,\ v)^\star \beta(v) + S_t(\xi, v)^\star [1 - \beta(v)]] \tag{20}$$

Here, the arbitrary number of δ ranging from 0 to 1, the arbitrary number of η ranging from 1 to D, the arbitrary number of ξ ranging from 1 to D and the arbitrary number of β ranging from 0 to 1.

   ii) Upgrade the location of follower: The arithmetical equation to upgrade the place of the follower is as follows

$$S_{;+1}^L (u,p) = S^H(H,p) + [cos(A_{u,p}^t)^\star S^H(H,p)^\star b_u^t] \tag{21}$$

Now, the coordinate selector is represented as p, S^H denotes the location of the primary rider, H indicates the index of primary rider, and b_u^t shows the distance covered by the uth rider.

iii) Upgrade the location of over taker: here, three significant factors should be taken into account namely the success rate, the direction indicator, and the coordinate selector.

$$S_{t+1}^Q (u, p) = S_t (u, p) + \left[Y_t^j (u) * S^H (H, p) \right] \tag{22}$$

iv) Upgrade the location of the attacker: Although the attacker goals to accomplish the location of the leader, the location Upgrade equation of the follower is utilized for upgrading the location of the attacker.

$$S_{t+1}^T (u, v) = S^H (H, v) + \left[cos \left(A_{u,v}^t \right)^* S^H (H, v) \right] + b_u^t \tag{23}$$

Now $S^H (H, v)$ signifies the location of the primary rider and the steering direction of the uth rider of the vth coordinate is denoted as $A_{u,v}^t$.

5. Describe the success rate again: Finally, rider finished the location upgrade method, the success rate can be evaluated by the rider's new location, and the existing primary rider could be recognized.
6. Upgrade the Rider constraint: It is important to upgrade the variable of the rider to obtain the optimum solution. Additionally, constraints such as activity, counter is also upgraded that facilitate upgrading the gear values and steering angle.
7. End procedure: The above-mentioned steps should be continuously iterated until the end time (T_{off}) and within that time, the primary rider should be identified that is taken into account as the winner of the bike competition. Fig. 2 illustrates the steps involved in ROA technique.
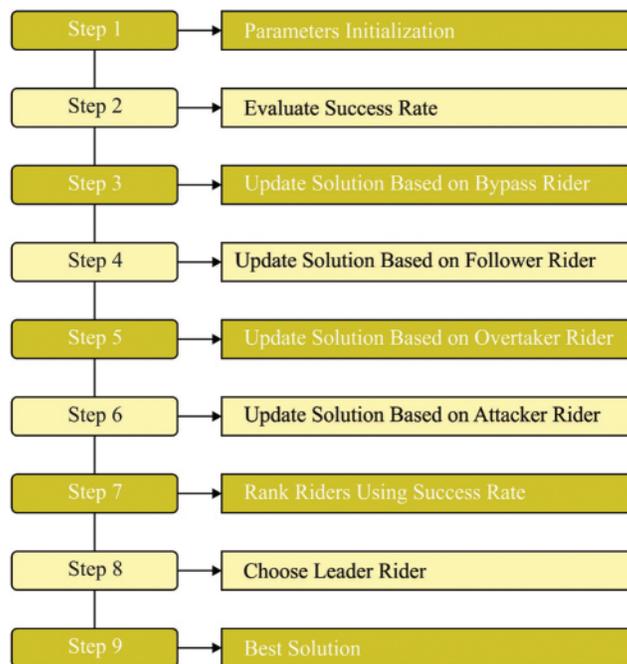

Figure 2: Steps involved in RO technique

3 Experimental Validation

In this section, the DLTRO-HCBC technique is simulated using a set of HIs collected from various sources. The experimental results of the DLTRO-HCBC technique are investigated under distinct sizes of training/testing data. A few sample images are demonstrated in Fig. 3.
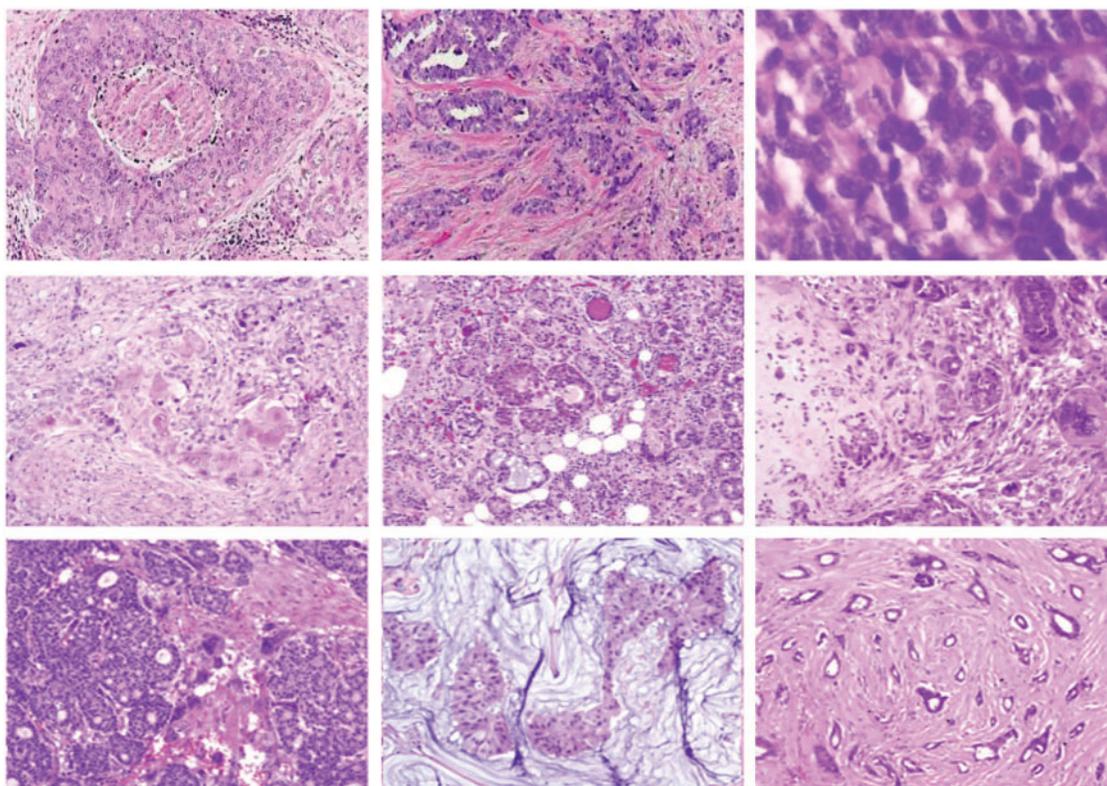

Figure 3: Sample images

Tab. 1 provides detailed breast cancer classifier results of the DLTRO-HCBC technique. Fig. 4 establishes the classification outcomes of the DLTRO-HCBC technique on testing or training data of 80:20. The figure illustrated that the DLTRO-HCBC technique has proficiently identified all three classes. For sample, the DLTRO-HCBC system has identified Normal class with $prec_n$ of 94.02%, $reca_l$ of 94.12%, $accu_y$ of 92.16%, F_{score} of 92.88%, and Mathew Correlation Coefficient (MCC) of 93.40%. Meanwhile, the DLTRO-HCBC technique has identified Benign class with $prec_n$ of 94.07%, $reca_l$ of 92.14%, $accu_y$ of 92.55%, F_{score} of 93.93%, and MCC of 93.74%. Eventually, the DLTRO-HCBC technique has identified Invasive Carcinoma class with $prec_n$ of 92.41%, $reca_l$ of 94.17%, $accu_y$ of 94.08%, F_{score} of 94.11%, and MCC of 92.14%. Fig. 5 proves the classification results of the DLTRO-HCBC method on training/testing data of 70:30. The figure illustrated that the DLTRO-HCBC technique has proficiently identified all three classes. For example, the DLTRO-HCBC system has identified Normal class with $prec_n$ of 92.42%, $reca_l$ of 94.17%, $accu_y$ of 93%, F_{score} of 93.48%, and MCC of 92.37%. Meanwhile, the DLTRO-HCBC technique has identified Benign class with $prec_n$ of 92.05%, $reca_l$ of 92.46%, $accu_y$ of 91.46%, F_{score} of 94.26%, and MCC of 93.34%. At last, the DLTRO-HCBC technique has identified Invasive Carcinoma class with $prec_n$ of 93.54%, $reca_l$ of 94.03%, $accu_y$ of 93.91%, F_{score} of 91.87%, and MCC of 93.35%.

Table 1: Result analysis of DLTRO-HCBC technique with different training/testing datasets

Class labels	$Prec_n$	$Reca_l$	Acc_y	F_{Score}	MCC
Training/Testing (80:20)					
Normal	94.02	94.12	92.16	92.88	93.40
Benign	94.07	92.14	92.55	93.93	93.74
In situ carcinoma	93.57	93.38	93.68	92.24	93.81
Invasive carcinoma	92.41	94.17	94.08	94.11	92.14
Average	93.52	93.45	93.12	93.29	93.27
Training/Testing (70:30)					
Normal	92.42	94.17	93.00	93.48	92.37
Benign	92.05	92.46	91.46	94.26	93.34
In situ carcinoma	93.73	92.80	93.87	93.37	93.78
Invasive carcinoma	93.54	94.03	93.91	91.87	93.35
Average	92.94	93.37	93.06	93.25	93.21
Training/Testing (60:40)					
Normal	91.63	91.92	92.32	91.98	93.35
Benign	93.20	92.39	93.98	91.67	92.37
In situ carcinoma	91.04	93.87	92.85	91.85	92.57
Invasive carcinoma	93.19	91.86	92.56	93.58	93.34
Average	92.27	92.51	92.93	92.27	92.91
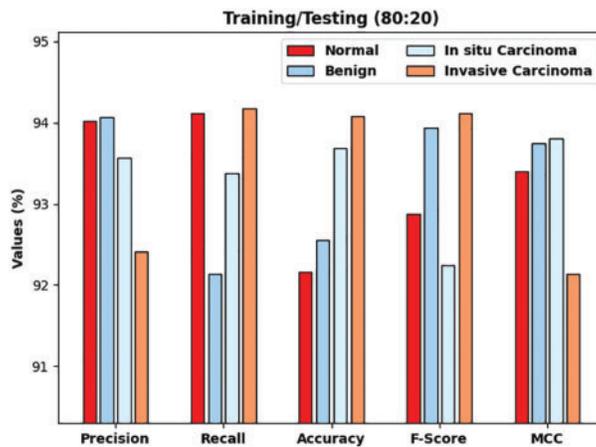
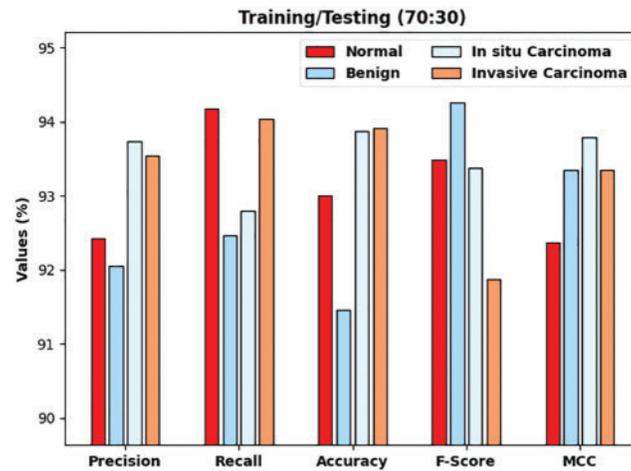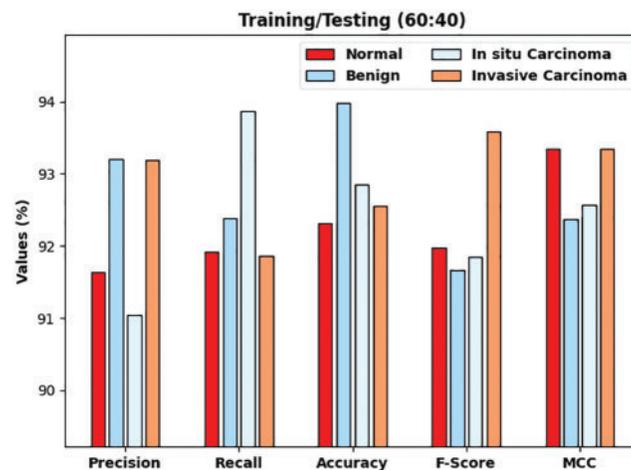**Figure 4:** Result analysis of DLTRO-HCBC technique under training/testing of 80:20 dataset

Figure 5: Result analysis of DLTRO-HCBC technique under training/testing of 70:30 dataset

Fig. 6 determines the classification results of the DLTRO-HCBC technique on training/testing data of 60:40. The figure illustrated that the DLTRO-HCBC technique has proficiently identified all three classes. For example, the DLTRO-HCBC system has identified Normal class with $prec_n$ of 91.63%, $reca_l$ of 91.92%, $accu_y$ of 92.32%, F_{score} of 91.98%, and MCC of 93.35%. Meanwhile, the DLTRO-HCBC technique has identified Benign class with $prec_n$ of 93.20%, $reca_l$ of 92.39%, $accu_y$ of 93.98%, F_{score} of 91.67%, and MCC of 92.37%. Finally, the DLTRO-HCBC technique has identified Invasive Carcinoma class with $prec_n$ of 93.19%, $reca_l$ of 91.86%, $accu_y$ of 92.56%, F_{score} of 93.58%, and MCC of 93.34%.

Figure 6: Result analysis of DLTRO-HCBC technique under training/testing of 60:40 dataset

Fig. 7 illustrates the experimental validation of the DLTRO-HCBC technique on training/testing data of 80:20. The figure portrayed that the DLTRO-HCBC technique has obtained precision of 93.52% and recall of 93.45%. At the same time, the DLTRO-HCBC technique has resulted in accuracy of 93.12% and F-score of 93.29%. Besides, the DLTRO-HCBC technique has obtained MCC of 93.27%.
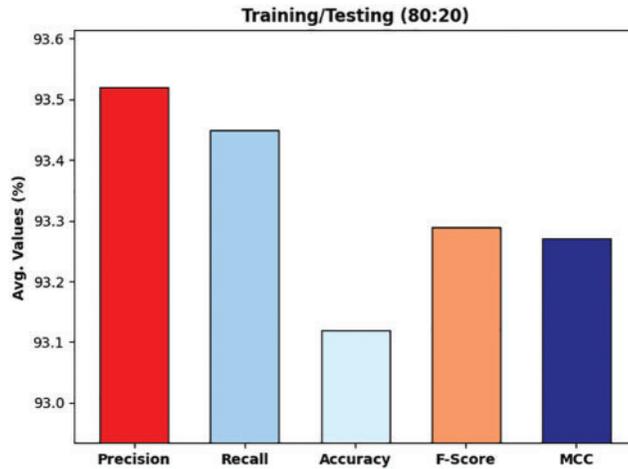
Figure 7: Overall analysis of DLTRO-HCBC technique under training/testing of 80:20 dataset

Fig. 8 validates the accuracy assessment of the DLTRO-HCBC model under training and testing of 80:20 dataset. The results described that the DLTRO-HCBC model has the aptitude of gaining improved values of training and validation accuracies. It is visible that the validation accuracy values are slightly higher than training accuracy.
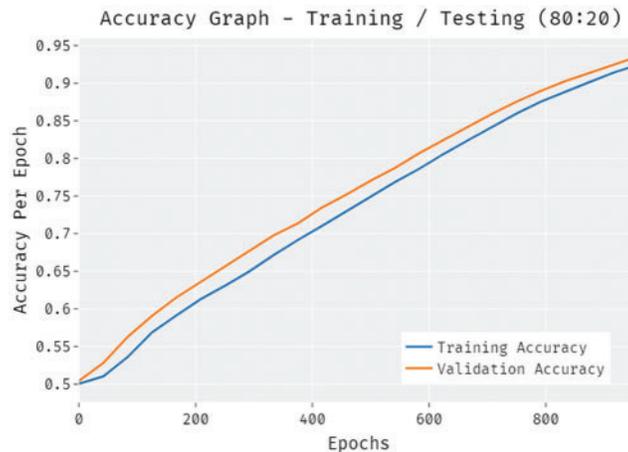

Figure 8: Accuracy analysis of DLTRO-HCBC technique under training/testing of 80:20 dataset

A brief training and validation loss offered by the DLTRO-HCBC model are reported in Fig. 9 under training and testing of 80:20 dataset. The results revealed that the DLTRO-HCBC model has accomplished minimum values of training and validation losses.

Fig. 10 illustrates the experimental validation of the DLTRO-HCBC technique on training/testing data of 70:30. The figure portrayed that the DLTRO-HCBC technique has obtained precision of 92.94% and recall of 93.37%. At the same time, the DLTRO-HCBC technique has resulted in accuracy of 93.06% and F-score of 93.25%. Besides, the DLTRO-HCBC technique has obtained MCC of 93.21%.
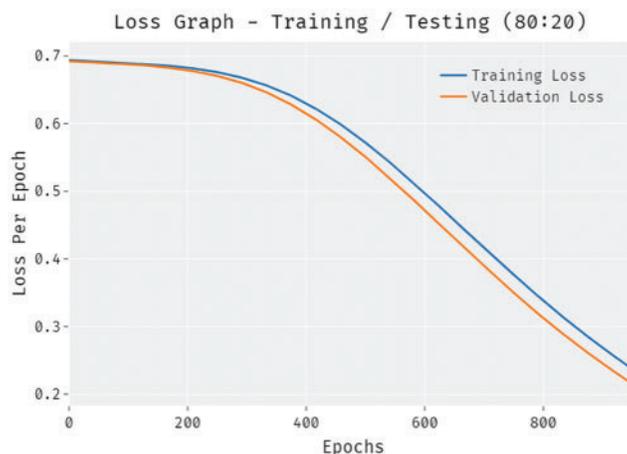
Figure 9: Loss analysis of DLTRO-HCBC technique under training/testing of 80:20 dataset
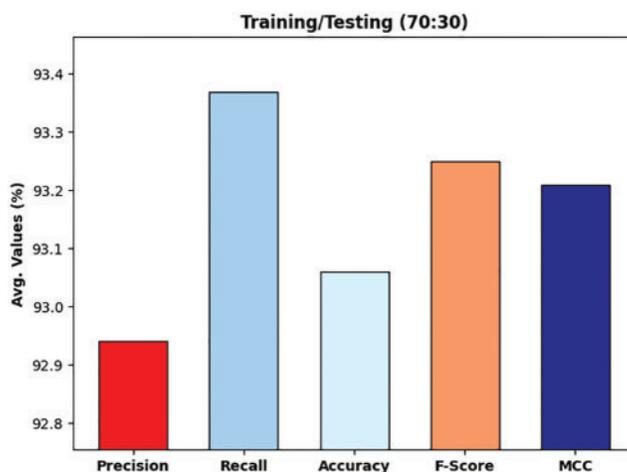

Figure 10: Overall analysis of DLTRO-HCBC technique under training/testing of 70:30 dataset

Fig. 11 validates the accuracy assessment of the DLTRO-HCBC model under training and testing of 70:30 dataset. The results described that the DLTRO-HCBC model has the aptitude of gaining improved values of training and validation accuracies. It is visible that the validation accuracy values are slightly higher than training accuracy.

A brief training and validation loss offered by the DLTRO-HCBC model are reported in Fig. 12 under training and testing of 70:30 dataset. The outcomes revealed that the DLTRO-HCBC model has gained minimum values of training and validation losses.

Fig. 13 illustrates the experimental validation of the DLTRO-HCBC technique on training/testing data of 60:40. The figure portrayed that the DLTRO-HCBC technique has obtained precision of 92.27% and recall of 92.51%. At the same time, the DLTRO-HCBC technique has resulted in accuracy of 92.93% and F-score of 92.27%. Besides, the DLTRO-HCBC technique has obtained MCC of 92.91%.
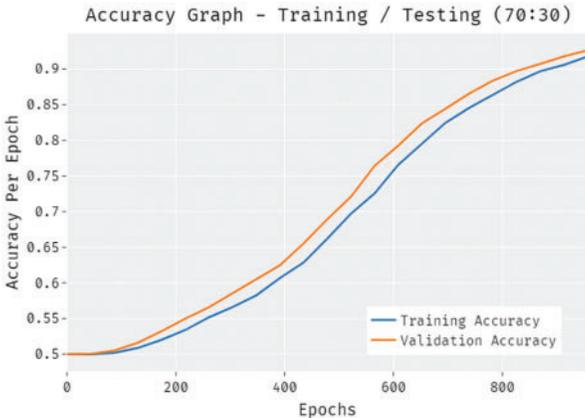
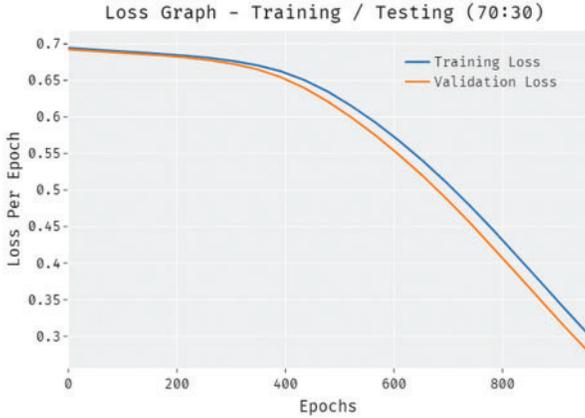**Figure 11:** Accuracy analysis of DLTRO-HCBC technique under training/testing of 70:30 dataset

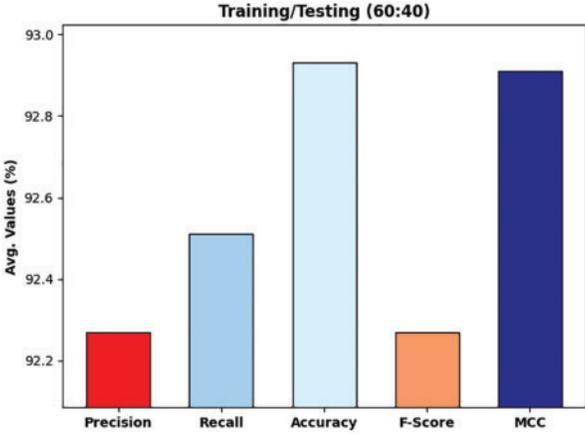**Figure 12:** Loss analysis of DLTRO-HCBC technique under training/testing of 70:30 dataset

Figure 13: Overall analysis of DLTRO-HCBC technique under training/testing of 60:40 dataset

Tab. 2 and Fig. 14 highlights the comparative accuracy analysis of the DLTRO-HCBC technique. The results indicated that the VGG16 system has demonstrated poor results with least accuracy of 80.35%. Simultaneously, the ResNet-50 and Inception v3 techniques have resulted in somewhat enhanced accuracy of 82.54% and 81.97% correspondingly. Also, the Incep. V3-LSTM model has accomplished reasonable results with an accuracy of 91.24%. Though the Incep. V3-bidirectional LSTM (BiLSTM model has resulted in competitive outcome with an accuracy of 92.45%, the DLTRO-HCBC technique has reached maximum accuracy of 93.12%. After examining the above-mentioned tables and figures, it can be clear that the DLTRO-HCBC method has gained maximum outcome on BCC.

Table 2: Comparative analysis of DLTRO-HCBC technique with recent approaches

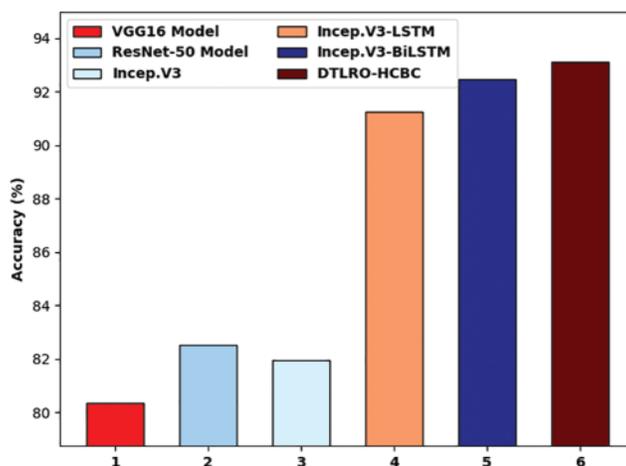| Methods	Accuracy
VGG16 model	80.35
ResNet-50 model	82.54
Incep.V3	81.97
Incep.V3-LSTM	91.24
Incep.V3-BiLSTM	92.45
DTLRO-HCBC	93.12

Figure 14: Comparative analysis of DLTRO-HCBC technique with recent approaches

4 Conclusion

The latest advances of DL models help in accomplishing maximum image classification performance in several application areas. This study has developed a new DTLRO-HCBC system for breast cancer detection and classification. The proposed DTLRO-HCBC technique aims to categorize the existence of breast cancer using histopathological images. The DTLRO-HCBC technique comprises a series of processes i.e., pre-processing, SqueezeNet based feature extraction, Adadelta hyperparameter optimizer, DFFNN classifier, and RO parameter optimization. The RO algorithm is applied for optimally adjusting the weight and bias values of the DFFNN model. The experimental results

of the DLTRO-HCBC technique are investigated under distinct sizes of training/testing data. For demonstrating the greater performance of DTLRO-HCBC process, a sequence of simulations was carried out and the outcomes reported its promising performance over the recent approaches. In future, hybrid metaheuristics can be employed for hyperparameter tuning process.

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

References

[1] F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *2016 Int. Joint Conf. on Neural Networks (IJCNN)*, Vancouver, BC, Canada, pp. 2560–2567, 2016.

[2] B. H. Aksebzeci, and Ö. Kayaalti, "Computer-aided classification of breast cancer histopathological images," *2017 Medical Technologies National Congress (TIPTEKNO)*, Trabzon, Turkey, pp. 1–4, 2017. https://doi.org/10.1109/TIPTEKNO.2017.8238076.

[3] A. Qu, J. Chen, L. Wang, J. Yuan, F. Yang *et al.,* "Two-step segmentation of hematoxylin-eosin stained histopathological images for prognosis of breast cancer," in *2014 IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM)*, Belfast, United Kingdom, pp. 218–223, 2014.

[4] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, Y. Ma *et al.,* "Histopathological whole slide image analysis using context-based CBIR," *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, pp. 1641–1652, 2018.

[5] W. Sun, G. Z. Dai, X. R. Zhang, X. Z. He and X. Chen, "TBE-Net: A three-branch embedding network with part-aware ability and feature complementary learning for vehicle re-identification," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2021. https://doi.org/10.1109/TITS.2021.3130403.

[6] W. Sun, L. Dai, X. R. Zhang, P. S. Chang and X. Z. He, "RSOD: Real-time small object detection algorithm in UAV-based traffic monitoring," *Applied Intelligence*, pp. 1–16, 2021. https://doi.org/10.1007/s10489-021-02893-3.

[7] S. K. Jafarbiglo, H. Danyali and M. S. Helfroush, "Nuclear atypia grading in histopathological images of breast cancer using convolutional neural networks," in *2018 4th Iranian Conf. on Signal Processing and Intelligent Systems (ICSPIS)*, Tehran, Iran, pp. 89–93, 2018.

[8] S. Angara, M. Robinson and P. Guillen-Rondon, "Convolutional neural networks for breast cancer histopathological image classification," in *2018 4th Int. Conf. on Big Data and Information Analytics (BigDIA)*, Houston, TX, USA, pp. 1–6, 2018.

[9] B. Wei, Z. Han, X. He and Y. Yin, "Deep learning model based breast cancer histopathological image classification," in *2017 IEEE 2nd Int. Conf. on Cloud Computing and Big Data Analysis (ICCCBDA)*, Chengdu, pp. 348–353, 2017.

[10] S. H. Gheshlaghi, C. N. E. Kan and D. H. Ye, "Breast cancer histopathological image classification with adversarial image synthesis," in *2021 43rd Annual Int. Conf. of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Mexico, pp. 3387–3390, 2021.

[11] R. Yan, F. Re, Z. Wang, L. Wang, T. Zhang *et al.,* "Breast cancer histopathological image classification using a hybrid deep neural network," *Methods*, vol. 173, pp. 52–60, 2020.

[12] Y. Zou, J. Zhang, S. Huang and B. Liu, "Breast cancer histopathological image classification using attention high-order deep network," *International Journal of Imaging Systems and Technology*, vol. 32, no. 1, pp. 266–279, 2022.

[13] P. Wang, P. Li, Y. Li, J. Wang and J. Xu, "Histopathological image classification based on cross-domain deep transferred feature fusion," *Biomedical Signal Processing and Control*, vol. 68, pp. 102705, 2021.

[14] C. Zhu, F. Song, Y. Wang, H. Dong, Y. Guo *et al.,* "Breast cancer histopathology image classification through assembling multiple compact CNNs," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 198, 2019.

[15] Y. Jiang, L. Chen, H. Zhang and X. Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module," *PLoS ONE*, vol. 14, no. 3, pp. e0214587, 2019.

[16] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally *et al.,* "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and<0.5 MB model size," arXiv preprint arXiv:1602.07360, 2016.

[17] S. Kumari and M. Bhatia, "A cognitive framework based on deep neural network for classification of coronavirus disease," *Journal of Ambient Intelligence and Humanized Computing*, vol. 2022, pp. 1–15, 2022, https://doi.org/10.1007/s12652-022-03756-6.

[18] T. Le, J. Kim and H. Kim, "An effective intrusion detection classifier using long short-term memory with gradient descent optimization," in *2017 Int. Conf. on Platform Technology and Service (PlatCon)*, Busan, Korea (South), pp. 1–6, 2017.

[19] T. K. Gupta and K. Raza, "Optimizing deep feedforward neural network architecture: A tabu search based approach," *Neural Processing Letters*, vol. 51, no. 3, pp. 2855–2870, 2020.

[20] D. Binu and B. S. Kariyappa, "RideNN: A new rider optimization algorithm-based neural network for fault diagnosis in analog circuits," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 1, pp. 2–26, 2019.