Technique for breast cancer classification using semi-supervised deep convolutional neural networks with transfer learning models

R. K. Chandana Mani and J. Kamalakannan*

School of Information Technology and Engineering, Vellore Institute of Technology, Vellore 632 014, India

Breast cancer affects several women worldwide every year. The survival rate of breast cancer depends on numerous factors. Early diagnosis and treatment are the most practical approaches to managing this disease. Deep learning-assisted cancer diagnosis is an effective technology for doctors to detect breast cancer quickly. Here, we propose a novel deep convolutional neural network-based transfer learning model for the accurate and most effective classification of breast cancer among women. This model is built using the pre-trained model. Inception-V3Net. First, the model is built with binary classification and then utilized to classify breast cancer histopathological images on a multi-class basis. The highest average accuracy attained by the proposed model is 94.8% when assessed under various magnifying factors. The final outcome of the proposed approach proves it to be more reliable and robust than the existing models. The proposed semi-supervised model of breast cancer classification is validated using the BreakHis public dataset. The result of the proposed CNN-built deep transfer learning approach is found to be better than the existing methods.

Keywords: Breast cancer, histopathological images, neural networks, semi-supervised model, transfer learning.

AT present, breast cancer is the major cause of cancer-related deaths among women. Over the years, new cases of breast cancer are increasing progressively. The usual medical screening process involves X-ray mammograms along with ultrasound imaging. However, the process of accurately identifying and classifying the type of breast cancer is a critical task. At the initial stage, cells in the breast grow in an uncontrolled manner. After a particular period, they form lumps and become tumours. The growth is found to be harmful if the cells invade surrounding tissues and spread across certain areas of the body[1]. The diagnosis based on ultrasound images is highly dependent on the radiologists. With the revolution in deep learning and data science, computer-aided diagnosis has become more prominent in the medical field, specifically for breast cancer diagnosis and treatment[2]. At an advanced level, with the help of deep-learning approaches, the complex patterns associated with cancer data can be observed and analysed efficiently. Further, accurately identifying the breast cancer subtype and prescribing the required treatment is also a challenging objective[3].

In general, computer-aided diagnosis of breast cancer deals with automated recognition of the disease with reduced mortality rate. The convolutional neural network (CNN) has been effective in image classification-related tasks. The core idea is to extract features and properly classify images. CNN contains many hidden layers through which it extracts features in an automated manner. Overall, it contains three layers, namely the input layer, hidden layer and output layers. Once the algorithm learns the features across layers, it performs the classification process using the fully connected output layer. Here, the key factor in training a CNN model for accurate prediction is the large amount of labelled data[4,5]. However, regarding medical imaging, especially in breast cancer classification, the amount of labelled data is scarce. Also, experts suggest that this process is often expensive and time-consuming. Another challenging factor is that training a larger model with a small amount of data often leads to overfitting. Such models will perform well with the training data but provide poor results on newly presented data. An alternative model is unsupervised learning that deals with unlabelled data. When used separately, neither approach is suitable for breast cancer diagnosis and treatment. However, an effective method exists to use large and powerful models for smaller datasets. One such model is transfer learning[6,7].

In transfer learning, the same model is trained with millions of images and can be successfully used for numerous applications. Through this, a new model is trained to make predictions about each image. The proposed approach in the present study uses the CNN-assisted transfer learning model along with the semi-supervised model. The semi-supervised model works on the basis of both labelled and unlabelled data. This approach is used to attain the desired results with limited data. When this approach is applied to breast cancer research, it provides more exciting results. Based on the features of the tumour cells, the proposed approach classifies the subtype of cancer at an earlier

*For correspondence. (e-mail: jkamalakannan@vit.ac.in)

stage with minimal data. The major goals of this study are as follows:

- A deep CNN-assisted semi-supervised transfer learning model is proposed that accurately classifies breast cancer patients at an early stage.
- The transfer learning model is developed with deep neural network models such as Inception-V3Net that are pre-trained to perform multiclass breast cancer classification on the BreakHis dataset of histopathological images.
- An efficient data augmentation algorithm is defined to improve the imbalanced class issue.
- The overall performance of this scheme is evaluated using various metrics such as F1 score, accuracy, precision and sensitivity.

Related works

We emphasize the comprehensive literature review of major concepts associated with the proposed approach. We provide a brief analysis of various kinds of deep-learning approaches that assist in efficiently classifying breast cancer at an early stage.

The conventional method of breast cancer classification includes standard classification algorithms like random forest, support vector machine (SVM), decision tree, etc.[8–10]. They make use of features such as colour and texture for the efficient classification of breast cancer. One major drawback of this approach is that it performs only the binary class classification.

Numerous deep learning models have been proposed and successfully used for breast cancer classification[11–13]. Ting *et al.*[14] proposed a competent CNN algorithm for classifying mammogram images. They have implemented the technique of feature-based pre-processing to improve the performance measure. This approach has attained a classification accuracy of 95% and a specificity of 90.70%. However, it fails to meet the evolving needs of breast cancer classification with increasing complexity.

Yan *et al.*[15] have proposed a hybrid neural network model for efficient breast cancer classification from histopathological images. They have used multilevel feature representation to build a network with patches. This approach integrates the benefits of both the recurrent and CNN models. It effectively preserves the long-term and short-term spatial correlation between the patches. It also provides an average accuracy of around 90.01% and shows less accuracy with benign images.

Spanhol *et al.*[16] have presented an efficient method for breast cancer classification using the CNN model. They have used a feed-forward approach for feature extraction and applied the CNN algorithm to classify histopathological images. Next, they trained the model with patches, and the final results were obtained with patch-based classification. Though this approach provides a better performance measure, it functions poorly when appropriately classifying various subtypes of cancerous images.

Deniz *et al.*[17] have made use of a transfer learning-based approach for efficient classification of histopathological images in the context of breast cancer detection. The authors stress on problem of fine-tuning with CNN-based classifiers when dealing with complex histopathological images. They have adapted the transfer learning-based deep feature extraction method to advance the accuracy of breast cancer classification. SVM was used for the classification process.

De Matos *et al.*[18] proposed a novel and efficient model for breast cancer classification using double-transfer learning. The process of feature extraction is done with the pre-trained Inception-V3 model. Classification of the images is done using the SVM classifier. The overall classification accuracy measure of this scheme was found to be comparatively better, but it lacked performance measures when tested with various magnification factors. Pang *et al.*[19] have proposed a semi-supervised GAN-grounded radiomics model for data augmentation in breast cancer classification. They performed the data augmentation process based on the semi-supervised model over ultrasound images. They have developed the GAN model to assess the ultrasound images effectively. This approach effectively generates high-quality ultrasound masses. The experimental results reveal that this approach provides more advanced breast mass classification.

An unsupervised learning model with generative adversarial neural networks has been proposed to classify breast cancer efficiently[20]. The authors have made use of the patch-based classification model to attain this objective. The DenseNet model was used for the classification process. This approach works well with coarse-grained high-resolution images. Another efficient model for the classification of histopathological images has been proposed[21]. The authors have extracted the features in multi-sized and discriminative patches along the overall tissue structure and morphology. A patch screening method using a clustering algorithm and CNN has been proposed. The experimental results of this approach were found to be comparatively better than the conventional methods.

To sum up, numerous techniques have been developed so far to detect breast cancer at an early stage[22–24]. The performance and accuracy of each method vary based on its overall architecture. However, many existing approaches face challenges when dealing with histopathological images with complex morphology and structure[25,26].

From Table 1 it is clear that though there has been outstanding research in the field of breast histopathological image classification, it is observed that the developed models still experience issues of deep learning, such as smaller datasets, class imbalance and lack of annotation. The approach proposed in this study aims to resolve these challenges more efficiently. Since we use the concept of transfer learning, it minimizes the disadvantages of small-scale datasets,

Table 1. Summary of related works

Year	Reference	Dataset	Methodology
2021	8	BreakHis	• Pre-trained ResNet architecture was used for magnification-dependent classification. • VGG-19 was used for breast image classification employing block-wise fine tuning strategy.
2019	11	BreakHis	• Feature extraction was done using three pre-trained architecture: GoogleNet, VGG-16 and ResNet. • Combined features from different architecture use average pooling classification.
2020	15	ICIAR	• Use CNN for multi-level feature extraction of patches. • RNN fuses the patches and performs multi-class classification.
2018	17	BreakHis	• AlexNet and VGG-16 are used for feature extraction and transfer learning. • SVM classifier is used for classifying concatenated features.
2020	20	BreakHis	• Screening patches are performed with the help of anomaly detection GAN. • Image classification is done by DenseNet121.
2019	21	ICIAR	• ResNet-50 is used for feature extraction. • SVM is used for multi-class classification.
2016	16	BreakHis	• Pre-trained models are used for binary classification. • AlexNet and LeNet architectures are utilized.
2019	25	BreakHis	• Utilized multiple instance learning. • CNN classification for patient-level and image-level tasks.

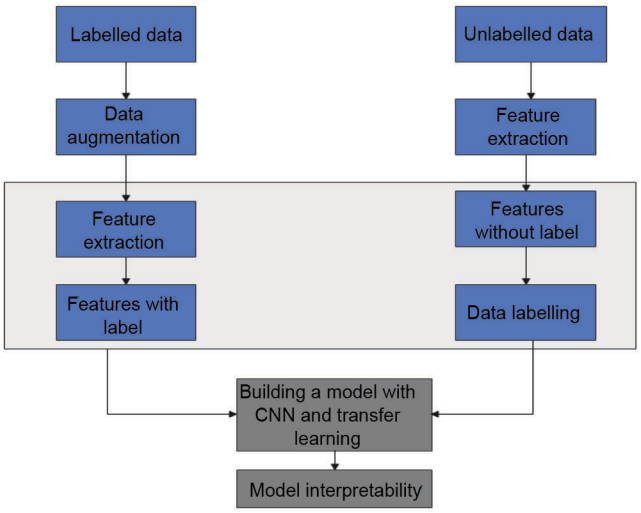

Figure 1. Overall work process of the proposed approach.

as it uses pre-trained weights. Also, with an advanced data augmentation algorithm, we have successfully resolved the issue of class imbalance. Our proposed method utilizes the process of data labelling to overcome the problem of annotation in medical images. However, the methods of patch-based feature extraction are challenging. We propose a novel architecture to perform patch-based feature extraction that obtains multi-class classification more efficiently and accurately.

Proposed deep convolutional neural network with transfer learning model

We present the major contributions of the proposed model that can efficiently classify histopathological images using a semi-supervised learning model with multi-class classification. Figure 1 illustrates the overall workflow of the proposed architecture.

Patch-based feature extraction

Let us consider a three-dimensional (3D) model of histopathological images. The first step is to generate a two-dimensional greyscale image $P_s = \{p_1, p_2, \ldots, p_n\}$, where n denotes the number of images. First, a grading value is estimated for every patch. In general, the patch library consists of thousands of patches, and the association between the target patch and patches found in the training library is determined using the weighting function. The weighting function for each training patch is computed as follows

$$W_i(j) = \mathrm{e}^{\frac{-\| P_i - P_L(j) \|_2^2}{h^2}}, \quad (1)$$

where the likeness between the training patch $P_L(j)$ and the target patch p_i is computed using the weight function $W_i(j)$. In this scenario, $P_L(j)$ represent the jth training patch associated with the library. The weight of the patch is close to 1 if the training patch is comparable to the target patch. Otherwise, the weight is assigned closer to zero. The grading values for the patches are estimated using the weight fusion method. Based on this method, grading values for all the N patches can be easily computed. Here, every training patch is used to compute the grading value based on its resemblance to the target patch. However, in a few cases, not all training patches need a similar pathological status corresponding to the target patch. When using misleading training patches, the final grading estimation result can be undermined. To overcome this issue, we use a sparse representation, patch-based grading method. The idea is to get the best representation of the target patch p_i from the

patch library and then compute the grading value. This is done with the help of the selected training patches. The sparse representation associated with the target patch is computed as follows

$$\alpha_i^{\wedge} = \alpha_i^{\min} \frac{1}{2}\left\| P_i - P_L \alpha_i \right\|_2^2 + \lambda_1 \left\| b_i \right\|_1 + \frac{\lambda^2}{2} \| b_i \left| a_i^{\wedge} \right|_2^2, \quad (2)$$

where $p_i^{\wedge}$ represents the coding coefficient of target patch p_i. However, the coefficient $b_i^{\wedge}$ is zero because of the sparsity constraint. The training patch is selected to provide pathological information to the target patch if the coefficient value of $a_i^{\wedge}$ is non-zero. Once we obtain the sparse solution, the value of the target patch is dependent on the coding coefficient $b_i^{\wedge}$ and also on the pathological status of the training patch. The grading value of the target patch is computed as follows

$$f_i = \frac{\sum_{j=1}^{N} b_i^{\wedge}(j) P_j}{\sum_{j=1}^{N} b_i^{\wedge}(j)}, \quad (3)$$

where N represents the sum of training patches in the patch library S_L and $b_j^{\wedge}(j)$ denotes the coding coefficient corresponding to the training patch $S_L(i)$.

Patch grounded network features: Each subject with breast cancer is identified with H patches. The corresponding relationship between the H patches within similar subjects is modelled using the network. In such a network developed, each edge signifies the correlation among a pair of patches, and each node represents a patch. The correlation factor between two patches is computed using various measures, including the mutual information n and the correlation coefficient. Here, we use correlation coefficient to compute the network features. The intensity of the patches is affected by the subjects because of various factors, including atrophyNotably, this network is a dense matrix, and it might not be similar to those developed using conventional approaches. Further, every correlation between patches reflects the structural changes for classification. To extract patch features, we have implemented the ResNet 34 model pretrained on ImageNet as the backbone. The last average pooling layer of the image classification model is eliminated here. ResNet 34 is used to balance the training samples efficiently. Every image from the dataset is first scaled to 224 × 334 and then fed into the backbone model to extract the patch features. Here, every patch feature indicates information from a particular region in the image. Each patch feature holds a dimension D, and it is denoted $P_j \in R^M$.

Data augmentation

In general, CNN needs substantial data to learn its parameters. One of the standard techniques to attain this objective is data augmentation. It helps enhance the system performance and avoid the possibility of overfitting and data imbalance[27–29]. Here, data augmentation is done only for the labelled data, especially with the training dataset. The concept of data augmentation during training is to extend the training data, which is made by creating instances. This data are derived from the training dataset, since it is possible to possess the ground truth label similar to its original image. This helps in encoding invariances that occur in the learning process. The training dataset is improved as follows

$$p(x) = P(B_b(x) \cdot b \sim P_\theta(b)). \quad (4)$$

Usually, data augmentation of images during the training phase is done through various processes, such as geometric distortion and photometric distortion. Geometric distortion includes rotation and can be used to a certain degree. Arbitrary rotations are used sometimes. The other option is mirroring. Photometric distortion includes contrast and brightness change of an image. In the proposed scheme the contrast of the image is changed to subject = (x(subject – $128 \cdot c > 0$). In case of randomized image augmentation, the contrast value is calculated for every image and random distribution is computed as $C \sim \text{Uniform}[p, q]$.

The parameters p, q should be selected on the validation dataset and not on the training set. The gamma change value is found as

$$\text{subject} = 255 * \left(\frac{\text{subject}}{255} \right)^C, \quad C > 0.$$

The brightness change is found as subject = C + subject. The generalization of the data augmentation scheme is as follows

$$x = \mu + \frac{\sqrt{\alpha \Sigma^{1/2}} \rho}{\sqrt{r}}, \quad P \sim N_e(0, J), \quad r \sim \frac{\alpha y_v^2}{v}, \quad P \perp r. \quad (5)$$

If we draw from the posterior $q(\theta | Y_{\text{obs}})$, where $\theta = (\mu, \Sigma)$, $Y_{\text{obs}} = \{Y_{n,n=1,\ldots,j}\}$, $Y_{\text{aug}} = ((Y_n, r_n), n = \{1, \ldots, i\})$. Since, $r_n = \alpha r_n^{\sim}$, the marginalization strategy could be accomplished. The standard augmentation scheme is given as

$$x_{\text{aug}}^{\wedge} = \{(x_n, r_n^{\sim}), n = 1, ..., i\}. \quad (6)$$

The joint posterior density value of θ, α and $r = \{r_1, \ldots, r_i\}$ is computed as

$$P(\theta, \alpha, r \mid Y_{\text{obs}}, \gamma, \beta) \alpha \alpha^{-\left[\frac{r+i(u+V)^2}{2}\right] \prod_{n=1}^{i} r_n \frac{u+v}{2} - 1 |\Sigma|^{\frac{-i+u+1}{2}}}$$

$$x \exp\left\{ \frac{-\sum_{n=1}^{i} r_i ((y_n - \mu)^T \Sigma^{-1} (y_n - \mu) + v) + B}{2\alpha} \right\},$$

which is followed as

$$r_i \mid \mu\Sigma, y_{\text{obs}}, \alpha \sim \frac{\alpha y^2 u + V}{(x_i - \mu)^T \Sigma^{-1} (x_i - \mu)\gamma}.$$

For every $n = 1, \ldots, i$,

$$\mu \mid \Sigma, x_{\text{aug}}, \alpha \sim I_d \left(\mu^{\wedge}, \frac{\alpha\Sigma}{\sum_{n=1}^{i} r_i} \right),$$

where

$$\mu^{\wedge} = \frac{\sum_{n=1}^{i} r_i y_i}{\sum_{n=1}^{i} r_i},$$

$$\Sigma^{-1} \mid x_{\text{aug}}, \alpha \sim \alpha W_{i-1} \left[\left(\sum_{n=1}^{i} r_n (x_n - \mu^n)(x_i - \mu^{\wedge})^T \right)^{-1} \right],$$

and

$$\alpha \mid x_{\text{aug}} \sim \frac{\beta + \gamma \sum_{n=1}^{i} r_i}{x^2_{r+iv}}.$$

$N_{\text{aug}}^{(w)}$ can be estimated through the second derivatives $\log p\ (\theta, \alpha \mid Y_{\text{aug}}, w)$ as a function of $\{\theta,\ \alpha\}$. Here r_i is replaced with

$$r_i^*(\theta) = E(r_i \mid x_i, \theta, \alpha) = \alpha E(r_i^{\sim} \mid x_1, \theta)$$

$$= \frac{\alpha(u+v)}{v + (x_i - \mu)^T \Sigma^{-1} (x_i - \mu)}.$$

The resultant expression $\theta = \theta^*$ is evaluated with the observed data posterior mode corresponding to θ and $\alpha = \alpha^{\wedge} = \beta/(r+2)$. The mode of the prior is found using $x(\alpha/\omega)$.

$$\frac{\partial \log p(\theta, \alpha \mid x_{\text{aug}}, w)}{\partial \alpha} = \frac{-r + i(u+v)^2}{2\alpha}$$

$$+ \frac{\sum_{n=1}^{i} r_n [(x_i - \mu)^T \Sigma^{-1} (X_i - \mu) + v] + \beta}{2\alpha^2}.$$

Here v is a non-zero vector of length $l(l+3)/2$ that is free of w.

$$I_{\alpha\alpha}(w) = \frac{-\gamma + i(u+v) + 2}{2\alpha^{\wedge 2}}$$

$$+ \sum_{n=1}^{i} \frac{r_i^*(\theta^*)[(x_i - \mu^*)^T \Sigma^{*-1} (x_i - u^y) + v] + B}{\alpha^{\wedge 3}}$$

$$= \frac{r + i(u+v) + 2}{2\alpha^{\wedge 2}}. \tag{7}$$

Generally, data augmentation is made through the transformation of input $x, f(x)$. Often, the value of brightness to be added is $F_f^{(x)}$. This parameter is used during training through random distribution $f \sim P_\theta(b)$. These are hyperparameters which are validated based on the validation set. The mappings are made using the following steps

Step 1: Take an independent value of $y^2u + v$. It is represented as $\{y^2u + v,\ n,\ n = 1, \ldots, i\}$. Then independently take the value $y_{iv}^2, y_v^2, R \sim I_d(o, N)$ and $W \sim W_{n-1}(l)$.

Step 2: $r_i^{\sim} = \dfrac{y^2u + v, n}{v + (x_i - \mu^{(t)})^T [\Sigma(t)]^{-1} (x_i - \mu^{(t)})}$

for $n = 1, \ldots, i$.

$$A = \left(\sum_{n=1}^{i} r_n^{\sim} (x_i - \mu^{\wedge(++)})\ (x_i - \mu^{\wedge(++1)})^T \right) \text{ with}$$

$$\mu^{\wedge(t+1)} = \frac{\sum_{n=1}^{i} r_i^{\sim} x_i}{\sum_{n=1}^{i} r_i^{\sim}}, \text{ and}$$

$$\mu^{(t+)} = \mu^{n(t+1)} + \frac{1}{\sqrt{\sum_{n=1}^{i} r_i^{\sim}}} (AW^{-1}B^T)_p. \tag{8}$$

Step 3: Then compute

$$\Sigma^{(t+1)} = \frac{x^2uv + x^2\gamma}{x_\gamma^{\sim 2} + v\sum_{n=1}^{1} r_i^{\sim}} AW^{-1}A^T,$$

$$\alpha^{(++)} = \frac{\beta + V\alpha^{(t)} \sum_{n=1}^{i} r_n^{\sim}}{x_{iv}^u + x_r^2},$$

$$\Sigma^{(++)} = \frac{x_{iv}^2 + x_v^2}{(\beta/\alpha^{(t)}) + v\sum_{n=1}^{i} r_i^{\sim}} AW^{-1}A^T,$$

$$\Sigma^{(++)} = \frac{x_{iv}^2}{v\sum_{n=1}^{i} r_i^{\sim}} BW^{-1}B^T. \quad (9)$$

From the above parameters, training is done in an efficient manner. Once all the parameters associated with the proposed scheme are selected, a final evaluation test data are used to check the overfitting problem through optimization parameters by determining the validation error. Table 2 depicts the transformation of the training dataset. Figure 2 shows the results of data augmentation.

Feature selection

Basically, training a network needs huge amounts of data because training with minimal data leads to overfitting. Here, we use fine-tuning as part of the pre-trained neural network to fit with the new dataset. A standard model of VGG = 16 network is used for feature extraction. The same procedure is followed for both the labelled and unlabelled data. The network contains of three shallow CNNs and VGG-VD (i.e. VGG-F, VGG-M and VGG-S). The deep neural network layer consists of 13 convolutional layers and three FC layers, where VGG-19 includes 16 convolutional layers and three FC layers. The corresponding VGG network rapidly approximates the target function then abstract relative feature representation. The fully connected layers from the VGG network are removed, and we make use of the Global average pooling operation consisting of four convolutional layers with 128, 256, 512, 512 channels. The process of feature selection is done through the evaluation of all patch features that contain only one input attribute. For instance, $\{A_1\}, \{A_2\}, ..., \{A_N\}$, where N is the input dimensionality, such that the best discrete feature is computed as $A(1)$. The process of forward selection identified the best subset with two components, viz. $A(1)$ and the other feature from the residual $N-1$ input attributes described. Here, we have a total of $N-1$ pairs. In order to search for the inclusive best input feature set, we use the exhaustive search.

$$NC(1) + \left(\frac{N}{2}\right)\left((2) + \ldots + \binom{N}{n}\right)C(n). \quad (10)$$

The overall process of feature selection is depicted in the following algorithm:

1. Assemble the training data from patches.
2. Shuffle the data patches.
3. Make it into n partitions (say $n = 20$).
4. For every partition ($i = 0, 1, \ldots, n-1$).
   a. Outer_Trainset(i) = every partition but not I.
   b. Outer_Testset(i) = 'I'th partition.
   c. Inner_Train(i) = to randomly select 70% of the outer train-set (i).
   d. Inner_TrainTest(i) = the next 30% of outer trainset(i).
   e. For $q = 0, 1, \ldots, n$, find the optimal feature set with j components PS_{ij}. This is made using leave-one-out on inner training (i). Let InnTest score$_{ij}$ = remaining score of P_{sj} on innerTest(j); end loop of (i).
   f. Select PS_{ij} (feature with best inner test score).
   g. Let outer score$_j$ = remaining score of selected features on outer_Test set (j).
   End loop (j);
   Return full procedure to evaluate feature selection of up to n attributes.

Data labelling

Now, let us consider two independent attribute sets. The unlabelled data are mapped to the labelled data once the two classifiers satisfy the agreement. Here, the process of independent assumption is relaxed, and co-training is made with weaker independent assumptions. Since we use the semi-supervised learning model, the proposed work uses a co-training procedure for confident labelling.

Step 1: Shuffle and partition the actual labelled dataset into two equal-sized subgroups (includes 10k patches), i.e. the size of every subgroup is half the size of the original labelled dataset.
Step 2: Train the gradient-boost classifier using the subgroup of labelled data.
Step 3: Append the common confident data to the labelled dataset.
Step 4: Set common confident instance by setting a threshold of min classifier (it should agree with the predicted label of unlabelled instances).
Step 5: For instance, set the threshold to 5.
Step 6: Repeat until we acquire confident data.
Step 7: Calculate the prediction accuracy of the classifier for every iteration.

Table 2. Transformation associated with the proposed approach for data augmentation

Parameters	Value
Total range of rotation	90°, 180°
Total random reflection	True
Total range of vertical shear	True
Total range of horizontal shear	True
Total range scale	True

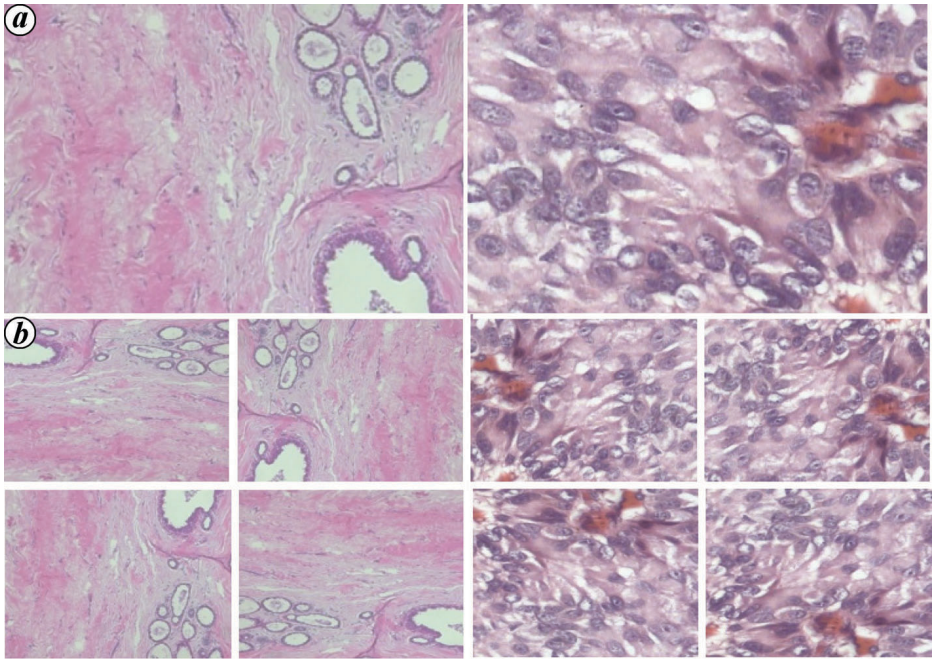

Figure 2. Depiction of data augmentation: (***a***) original image and (***b***) rotated image.

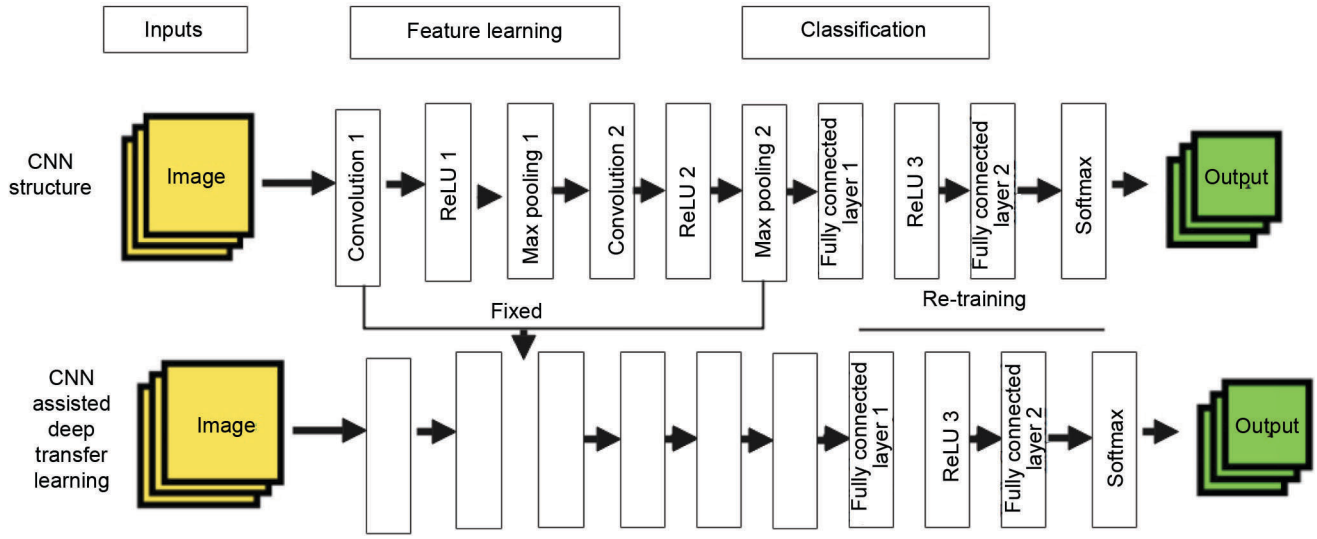

Figure 3. Architecture of the proposed CNN-based transfer learning model.

Step 8: Reshuffle the expanded labelled dataset if (accuracy of classifier < previous iteration).

Step 9: Obtain the final labelled dataset from the confidential instances.

Here, we add only the common confident instances to the labelled dataset. This does not imply that all the unlabelled data are added to the labelled data that are noisy. Also, the synthetic patches produced by deep convolutional generative adversarial neural networks are only used for the labelling process and to train the final network.

Proposed deep convolutional neural network architecture with transfer learning

CNN has acquired significant importance in recent years. In general, the CNN architecture consists of the input (i/p) layer, output (o/p) layer and multiple hidden layers. Whereas the fully connected layers, normalization layers (ReLU) and pooling layers are part of the hidden layers. In case of a complex model, the additional layers are used. Each layer in a CNN model possesses various characteristics of an image set. For example, if we feed a breast cancer medical

Table 3. Overview of image distribution in BreakHis dataset

Category	Subtype	Total no. of images	Magnification factor			
			40×	100×	200×	400×
Benign	Tabular adenoma	569	149	150	140	130
Benign	Phyllodes tumor	453	109	121	108	115
Benign	Adenosis	444	114	113	111	106
Benign	Fibroadenoma	1014	253	260	264	237
Malignant	Lobular carcinoma	626	156	170	163	137
Malignant	Ductal carcinoma	3451	864	903	896	788
Malignant	Papillary carcinoma	560	145	142	135	138
Malignant	Mucinous carcinoma	792	205	222	196	169

image as input to CNN, the network will learn its basic features in the initial layers. The subsequent set of layers will include shapes and objects related to breast cancer detection. Patch-based CNN works efficiently when compared to CNN training on entire image.

The image classification process is broken down into small parts called patches. In the proposed model, a 3×3 grid is used to represent the feature extraction process. The features are lined with the image patches. Every pixel is multiplied by an equivalent feature pixel, and once this process is over, the entire value is summated and divided by the total number of pixels in the feature space. The ultimate value of the feature forms the feature patch. The same procedure is repeated for all the patches, and this is known as convolution. The next layer deals with shrinking the image stock. This layer is mainly used for dimensionality reduction without the loss of important information. The normalization layer is a rectified linear unit (ReLU). It converts the negative values within the filter to 0. In the next step, we stack the convolution, ReLU layers. Thus, the output (o/p) of one layer becomes the input (i/p) of the next layer and achieve better results in decision-making. Figure 3 provides an overview of the proposed architecture.

The standards used for the convolution layer and weight of the patches are computed using backpropagation. In transfer learning, the model is trained and developed for one particular task and can be reused for a related task. The proposed model uses the Inception-V3 model that is trained on ImageNet and can be repurposed to train on a new dataset. This is done with the help of random weight initialization. The proposed approach uses tensor flow to facilitate the transfer learning-assisted pre-trained CNN model. The proposed CNN model permits image classification through transfer learning. In this model, CNN is run in a source domain, and the hidden layers of CNN are reused in the target domain. The fully connected layers 1 and 2 found in the deep layers are retrained with the dataset of the target domain.

The core idea behind the proposed model is feedforward neural networks. This is achieved with the help of a linear combination of neurons. The model has a global linear structure, which is attained using a training layer, where a single network layer is mapped to meaningful user-defined concepts.

Results and discussion

In this section, we present the results of experiments on breast cancer classification from histopathological images using the proposed approach. The experiment is conducted using the BreakHis dataset, with 7909 samples collected from 82 patients. The overall dataset is divided into two major categories, benign and malignant samples of 2480 and 5429. All the experiments in this paper are completed on NVIDIA Geforce GTX GPUs with Intel i7 9th gen processor using Tensor-flow framework. During network training, the proposed model uses the Adam optimisation algorithm with learning rate, beta1 and beta 2 are tuned to 0.0001, 0.6 and 0.8 respectively. The size of mini batch was arranged as 16, and the total number of epochs used for the proposed model is 500. Further, observations from the experiment are described in detail in the following sections.

Dataset description

Here, we have utilized the BreakHis dataset from the database of vision robotics and imaging, Federal University of Parana, Brazil. The dataset consists of two major groups: benign and malignant tumours. Each involves a distinct magnification of 40, 100, 200 and 400×. The images in the dataset have a resolution of 700×460 pixels in three-channels and are RGB images. Overall, the dataset contains information about 82 patients with 2480 benign and 5429 malignant samples. The benign images include four classes, namely tabular adenoma (TA), phyllodes tumour (PT), adenosis (A) and fibroadenoma (F). The malignant images also have four classes, namely lobular carcinoma (LC), ductal carcinoma (DC), papillary carcinoma (PC) and mucinous carcinoma (MC). Table 3 describes the images in the BreakHis dataset with its classes. Figure 4 provides an overview of the dataset under various magnifications.

Data pre-processing

Images must be processed before they can be used for model training and inference. This includes, but is not limited to, changes in colour, size and direction. Pre-processing is done to enhance the image quality so that we can analyse

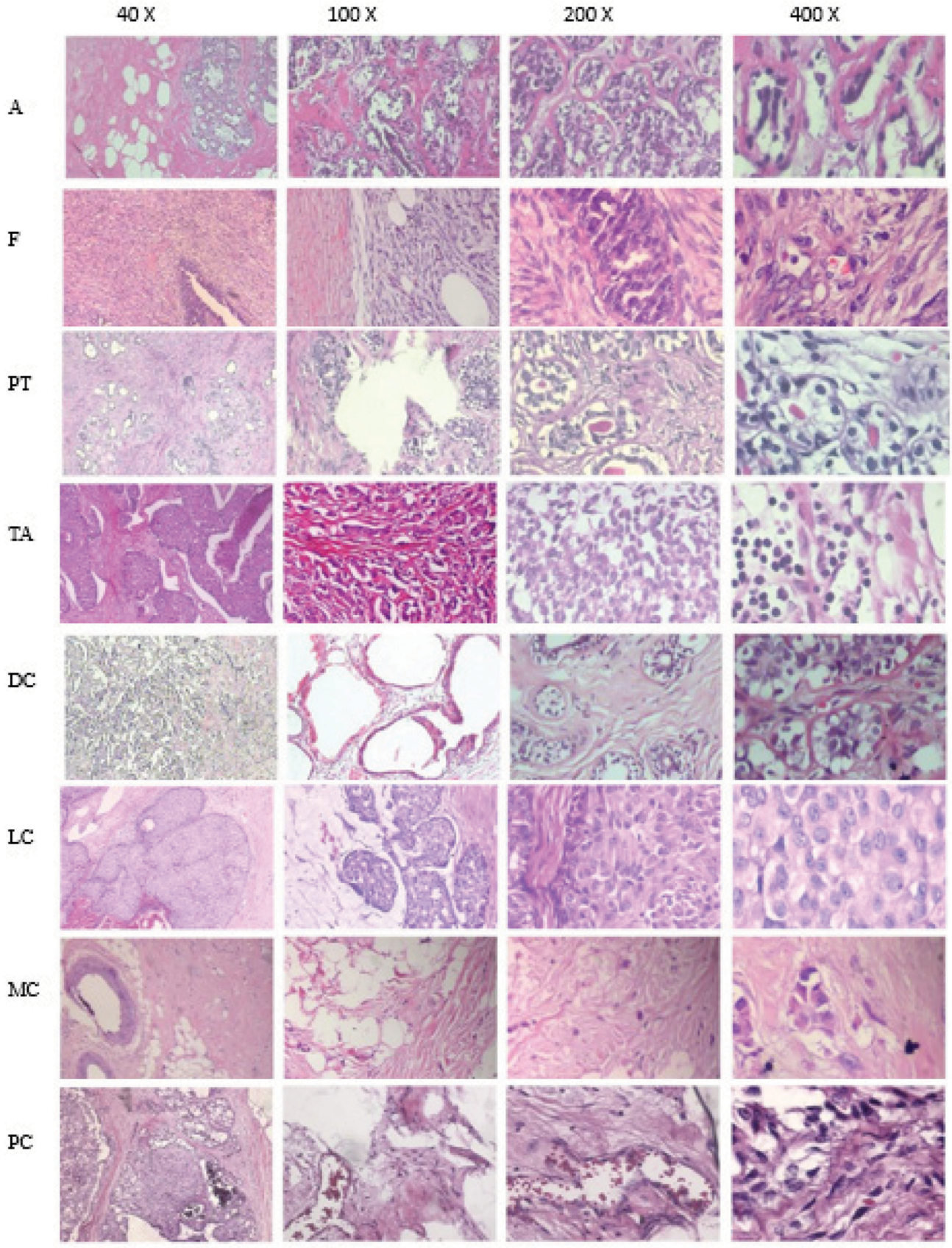

Figure 4. Sample images from BreakHis dataset. A, Adenosis; F, Fibroadenoma; TA, Tubular adenoma; PT, Phillodes tumor are benign; DC, Ductal carcinoma; LC, Lobular carcinoma; MC, Mucinous carcinoma; PC, Papillary carcinoma are malignant.

Table 4. Overall accuracy of the proposed multi-class classification

Accuracy (%)	Magnification factor
96.08	40×
95.12	100×
94.57	200×
93.67	400×

Table 5. Overall accuracy of the proposed binary class classification

Accuracy (%)	Magnification factor
98.23	40×
97.38	100×
96.45	200×
95.64	400×

Table 6. The performance statistics of the proposed approach for each sub-type

Target_Class name	Magnification factor	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Ductal carcinoma	40×	97.11	99	92	93
	100×	98.01	94	92	93
	200×	99.93	98	95	96
	400×	97.55	92	92	92
Adenosis	40×	99.2	95	100	97.2
	100×	99.16	97	96	97.2
	200×	98.39	93	97	94
	400×	98.99	95	97	94
Lobular carcinoma	40×	97.59	94	89	88
	100×	98.9	94	95	95
	200×	98.9	95	92	93
	400×	98.44	98	95	96
Fibroadenoma	40×	99.2	98	95	96
	100×	98.5	95	94	96
	200×	99.44	95	94	95
	400×	98.54	92	92	92
Papillary carcinoma	40×	97.59	96	92	94
	100×	98.9	91	98	94
	200×	98.03	95	98	96
	400×	98.56	92	98	95
Tubular adenoma	40×	99.48	95	96	95
	100×	99.53	95	95	95
	200×	99.81	99.02	98	97
	400×	98.41	95	95	95
Mucinous carcinoma	40×	97.97	97	92	91
	100×	98.07	95	95	95
	200×	98.66	92	92	92
	400×	98.01	93	92	91
Phyllodes tumor	40×	99.34	98	96.7	98.1
	100×	99.34	93	94	94
	200×	99.83	93	95	96
	400×	98.13	94	94	94
Overall accuracy	40×	96.08%		–	
	100×	98.12			
	200×	94.57		–	
	400×	93.67			

it more effectively. Through pre-processing, we can eliminate undesired distortions and enhance certain qualities crucial for developing the application. The qualities could change based on the application. An image must be pre-processed for the software to work correctly and deliver the required results.

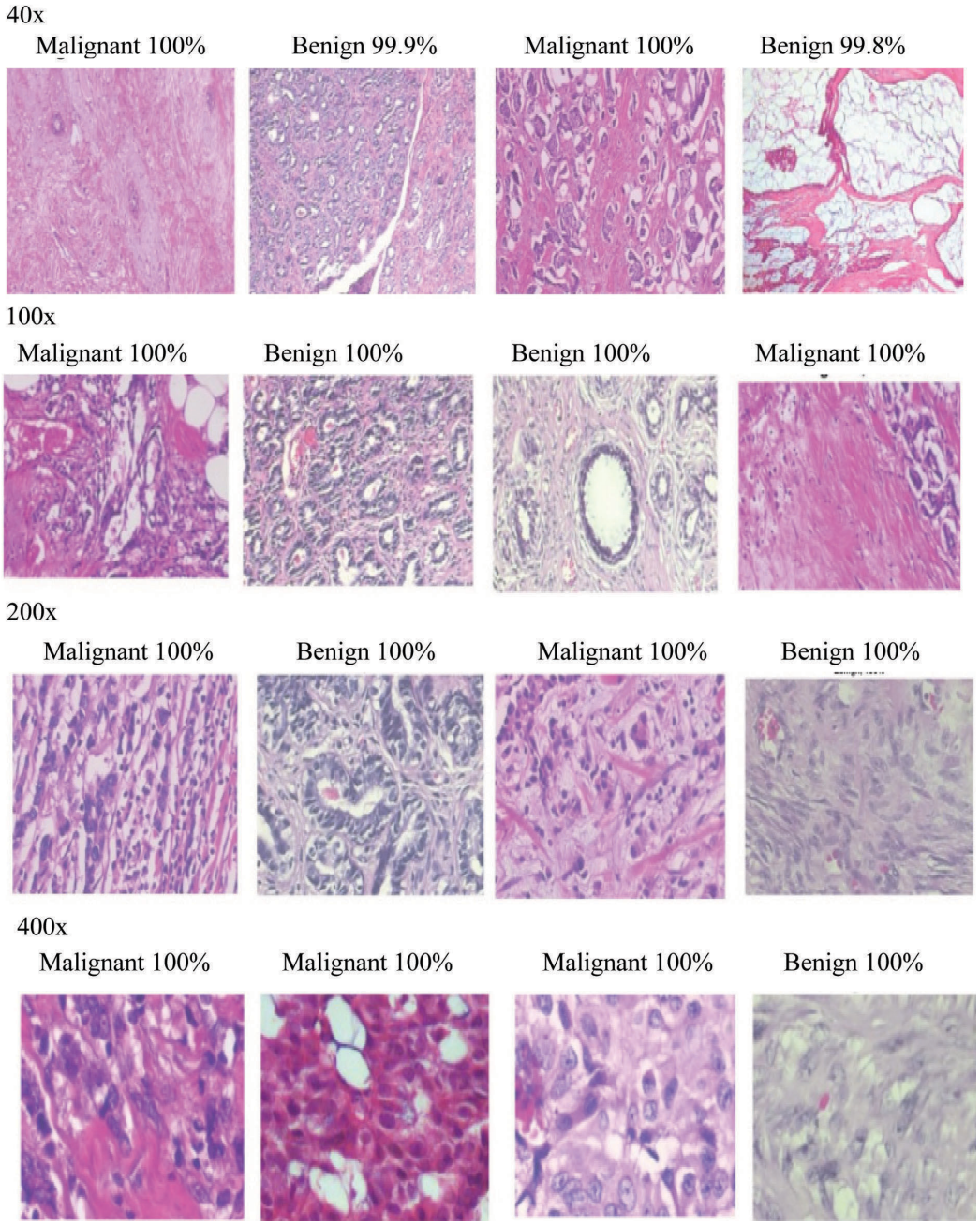

Figure 5. Proposed model prediction on sample test images.

First, the dataset is normalized to meet the requirements of various networks. This can be accomplished through data rescaling and cropping, so images from the dataset fit with the pre-trained deep neural network. The images are pre-processed before being fed into the network using standard methods. Here, we use the image-based approach, where the dataset is randomly divided without replacement. The training set consists of 70% of the images, and the remaining 30% forms the test dataset. The test and training data are trained under various magnifications separately. To avoid the problem of overfitting, the proposed approach has implemented data augmentation.

Results of the proposed multi-class classification model

The performance of the proposed model was evaluated using various metrics such as accuracy, precision, sensitivity, specificity and F1 score, and based on the pre-trained neural network model. The results were based on various magnification factors, viz. 40, 100, 200 and 400×. Table 4 shows the accuracy of the multiclass classification and Table 5 shows the overall accuracy of the binary class classification. Table 6 shows the overall performance measure.

Figure 5 depicts the model prediction on the sample test image of the proposed approach. It can be observed from the experiment that the proposed method has attained a higher accuracy rate of breast cancer classification under various magnifications such as 40, 100, 200 and 400×. The overall average accuracy of the proposed method is found to be 94.8%, and it gives the best performance measures. Thus, we can conclude that the proposed method is ideal for breast cancer classification under various magnifications. So far, the proposed model has been tested to assess its performance without any comparison with the existing approaches. To further assess the performance of the proposed semi-supervised classification model, we conducted three more experiments. In the first experiment, the total training and validation dataset contained 112, 325 labelled patches and the test dataset with 50,674 patches was fed across the network. It had taken over 100 runs. For every performance metric, the values were assessed and compared with the other approaches. The proposed approach provided an accuracy of 0.8795 and F-measure of 0.8123, which is comparatively higher than the existing approaches. The overall performance measure of the proposed approach was observed to be significantly better than the existing approaches (Tables 7–9).

In the second experiment, we assessed the performance of the proposed approach by adding some unlabelled data to the training dataset. The performance measures were evaluated using a combination of labelled and unlabelled data. The unlabelled data were labelled using the proposed data-labelling model. We assessed the performance measure for around 10,000 labelled instances, and the existing approaches achieved an accuracy of 0.8529, i.e. accuracy = 82.24% and F-score = 73.81%. The proposed network achieved an accuracy of 0.8735, i.e. accuracy = 86.75% and F-score = 79.29, using the same amount of labelled and unlabelled data. Figure 6 shows a comparison of the performance of labelled and unlabelled data.

In the third experiment, we evaluated the performance of the proposed model with a supervised approach. That is, a comparison was made between the semi-supervised and supervised models. The supervised model was trained with labelled data containing around 10,000 patches and a single generative network of semi-supervised learning. The proposed semi-supervised model uses a generative model trained with labelled and unlabelled data. The model is trained up to 12,000 patches out of 102,135 patches. To compare the performance measures, the test data used for both models was the same.

Table 7. The performance comparison between the proposed and other approaches for experiment 1

Scheme	F measure	Accuracy
Ref. 27	0.728	0.853
Ref. 28	0.743	0.8421
Proposed	0.8123	0.8795

Table 8. The performance comparison between the proposed and other approaches for experiment 2

Scheme	F measure	Accuracy
Ref. 26	0.7629	0.8529
Proposed	0.7923	0.8735

Table 9. The performance comparison between the proposed and other approaches for experiment 3

Scheme	F measure	Accuracy
Supervised model (ref. 27)	0.701	0.7829
Proposed	0.8135	0.8749

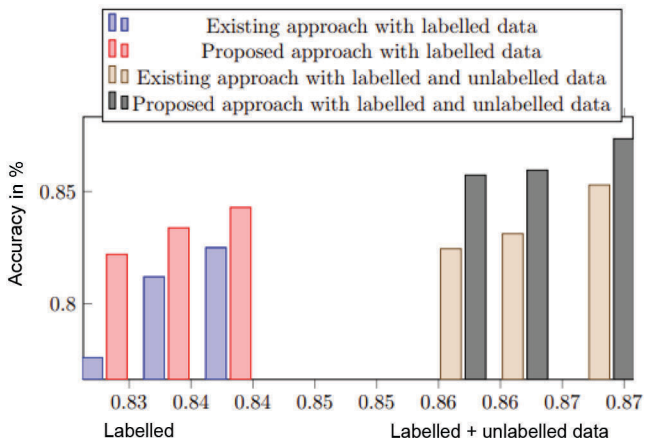

Figure 6. Performance of the proposed and existing approaches with various labelled and unlabelled data. L = 2500, L = 5000 and L = 10,000.

Conclusion

In this study, we propose a CNN-based transfer learning approach for multi-class classification of breast cancer images. We have extracted data from histopathological images in the BreakHis public dataset. Data augmentation techniques were used to improve the performance of the proposed approach. The results were assessed using various performance matrices, such as accuracy, precision, recall and sensitivity measures. The transfer learning model was implemented by replacing the last three layers in CNN. The proposed approach offers good accuracy of 94.8% with multi-class classification. In the future, this work can be extended to improve measures and histopathological images such as liver, prostrate and kidney.

Conflict of interest: The authors declare that there is no conflict of interest.

Ethical approval: This study did not include any human participants or animals.

1. Waks, A. G. and Winer, E. P., Breast cancer treatment: a review. *JAMA*, 2019, **321**(3), 288–300.

2. Sun, Y.-S. *et al*., Risk factors and preventions of breast cancer. *Int. J. Biol. Sci.*, 2017, **13**(11), 1387.
3. Reis-Filho, J. S. and Pusztai, L., Gene expression profiling in breast cancer: classification, prognostication, and prediction. *Lancet*, 2011, **378**(9805), 1812–1823.
4. Roslidar, R., Saddami, K., Arnia, F., Syukri, M. and Munadi, K., A study of fine-tuning CNN models based on thermal imaging for breast cancer classification. In IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), IEEE, 22 August 2019, pp. 77–81.
5. Gao, F. *et al*., SD-CNN: a shallow-deep CNN for improved breast cancer diagnosis. *Comput. Med. Imag. Graph.*, 2018, **70**, 53–62.
6. Zemmal, N., Azizi, N., Dey, N. and Sellami, M., Adaptive semi supervised support vector machine semi supervised learning with features cooperation for breast cancer classification. *J. Med. Imag. Health Informat.*, 2016, **6**(1), 53–62.
7. Peng, L. *et al*., An immune-inspired semi-supervised algorithm for breast cancer diagnosis. *Comput. Methods Programs Biomed.*, 2016, **134**, 259–265.
8. Boumaraf, S. *et al.*, Conventional machine learning versus deep learning for magnification dependent histopathological breast cancer image classification: a comparative study with visual explanation. *Diagnostics*, 2021, **11**(3), 528.
9. Vidić, I. *et al*., Support vector machine for breast cancer classification using diffusion-weighted MRI histogram features: preliminary study. *J. Magn. Reson. Imaging*, 2018, **47**(5), 1205–1216.
10. Costa, B. *et al*., Understanding breast cancer: from conventional therapies to repurposed drugs. *Eur. J. Pharm. Sci.*, 2020, **151**, 105401.
11. Khan, S. U. *et al*., A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recogn. Lett.*, 2019, **125**, 1–6.
12. Abdel-Zaher, A. M. and Eldeib, A. M., Breast cancer classification using deep belief networks. *Expert Syst. Appl.*, 2016, **46**, 139–144.
13. Golatkar, A., Anand, D. and Sethi, A., Classification of breast cancer histology using deep learning. In *International Conference Image Analysis and Recognition*, Springer, Cham, 2018.
14. Ting, F. F., Tan, Y. J. and Sim, K. S., Convolutional neural network improvement for breast cancer classification. *Expert Syst. Appl*., 2019, **120**, 103–115.
15. Yan, R. *et al*., Breast cancer histopathological image classification using a hybrid deep neural network. *Methods*, 2020, **173**, 52–60.
16. Spanhol, F. A., Oliveira, L. S., Petitjean, C. and Heutte, L., Breast cancer histopathological image classification using convolutional neural networks. In International Joint Conference on Neural Networks (IJCNN), IEEE, 24 July 2016, pp. 2560–2567.
17. Deniz, E., Şengür, A., Kadiroğlu, Z., Guo, Y., Bajaj, V. and Budak, Ü., Transfer learning based histopathologic image classification for breast cancer detection. *Health Inform. Sci. Syst.*, 2018, **6**(1), 1–7.
18. De Matos, J., Britto, A. D., Oliveira, L. E. and Koerich, A. L., Double transfer learning for breast cancer histopathologic image classification. In International Joint Conference on Neural Networks (IJCNN), IEEE, 14 July 2019, pp. 1–8.
19. Pang, T. *et al*., Semi-supervised GAN-based radiomics model for data augmentation in breast ultrasound mass classification. *Comput. Method. Program. Biomed*., 2021, **203**, 106018.
20. Man, R., Yang, P. and Xu, B., Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks. *IEEE Access*, 2020, **8**, 155362–155377.
21. Li, Y., Wu, J. and Wu, Q., Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning. *IEEE Access*, 2019, **7**, 21400–21408.
22. Awan, R., Koohbanani, N. A., Shaban, M., Lisowska, A. and Rajpoot, N., Context-aware learning using transferable features for classification of breast cancer histology images. In Image Analysis and Recognition: 15th International Conference, ICIAR, Springer International Publishing, Portugal, 27–29 June 2018, pp. 788–795.
23. Wang, D. *et al*., Deep learning for identifying metastatic breast cancer. 2016; arXiv:1606.05718.
24. Yang, Z. *et al*., EMS-Net: ensemble of multiscale convolutional neural networks for classification of breast cancer histology images. *Neurocomputing*, 2019, **366**, 46–53.
25. Sudharshan, P. J., Petitjean, C., Spanhol, F., Oliveira, L. E., Heutte, L. and Honeine, P., Multiple instance learning for histopathological breast cancer image classification. *Expert Syst. Appl*., 2019, **117**, 103–111.
26. Elbashir, M. K. *et al*., Lightweight convolutional neural network for breast cancer classification using RNA-seq gene expression data. *IEEE Access*, 2019, **7**, 185338–185348.
27. Gobinath, C. and Gopinath. M. P., Design and development of SER-UNet model for glaucoma image analysis. *J. Electron. Imag*., 2022, **31**(4), 041215.
28. Gobinath, C. and Gopinath, M. P., Attention aware fully convolutional deep learning model for retinal blood vessel segmentation. *J. Intell. Fuzzy Syst.*, (preprint), 1–11.
29. Murugesan, M., Kaliannan, K., Balraj, S., Singaram, K., Kaliannan, T. and Albert, J. R., A hybrid deep learning model for effective segmentation and classification of lung nodules from CT images. *J. Intell. Fuzzy Syst.*, 2021, **42**(3), 2667–2679; doi:10.3233/JIFS-212189.
30. Zhou, J. *et al*., Weakly supervised 3D deep learning for breast cancer classification and localization of the lesions in MR images. *J. Magn. Reson. Imaging*, 2019, **50**(4), 1144–1151.
31. Bhuiyan, M. N., Shamsujjoha, M., Ripon, S. H., Proma, F. H. and Khan, F., Transfer learning and supervised classifier based prediction model for breast cancer. In *Big Data Analytics for Intelligent Healthcare Management*, Academic Press, 2019, pp. 59–86.

Received 15 December 2022; revised accepted 8 August 2023

doi: 10.18520/cs/v125/i9/970-982