



BrC-MCDLM: breast Cancer detection using Multi-Channel deep learning model

Jitendra V. Tembhurne¹ · Anupama Hazarika² · Tausif Diwan¹

Received: 9 September 2020 / Revised: 19 May 2021 / Accepted: 28 June 2021 /
Published online: 16 July 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

Breast cancer (BrC) is a lethal form of cancer which causes numerous deaths in women across the world. Generally, mammograms and histopathology biopsy images are recommended for early detection of BrC as they enable a more reliable prediction than just using mammograms. However, research indicates that even the most experienced dermatologists can detect BrC in early stage with an average accuracy of less than 80%. Over the years, researchers have made significant progress in the development of automated tools and techniques to assist radiologists or medical practitioners in BrC detection. Various machine learning and deep learning based architectures are extensively experimented on different publicly available datasets to improve the performance measures. There is further scope of improvements by extracting better representative features with deep architectural variants or ensembles techniques to minimize the misclassifications. Learnt parameters of any pretrained deep models may provide a better starting point for any other architectures using transfer learning technique. In this work, we propose computer-aided transfer learning based deep model as a binary classifier for breast cancer detection. Generally, deep learning architectures are sequential, following only a single channel for features' extraction and further classification. However, fused features extracted from multiple channels may better represent features qualitatively. The novelty of our approach is the use of multi-channel merging techniques for devising a dual-architecture ensemble. The models are trained and tested on the BreakHis dataset and an improvement in comparison with the state-of-the-arts is observed in various performance metrics. Among several combinations for ensemble architectures by utilizing various pretrained models, the Xception + InceptionV3 combination achieved an average accuracy of 97.5% for multi-channelled architecture, setting benchmarking results for further research in this direction.

Keywords Breast cancer · Classification · Deep convolutional neural network · Multi-channel merging

✉ Jitendra V. Tembhurne
jtembhurne@iiitn.ac.in

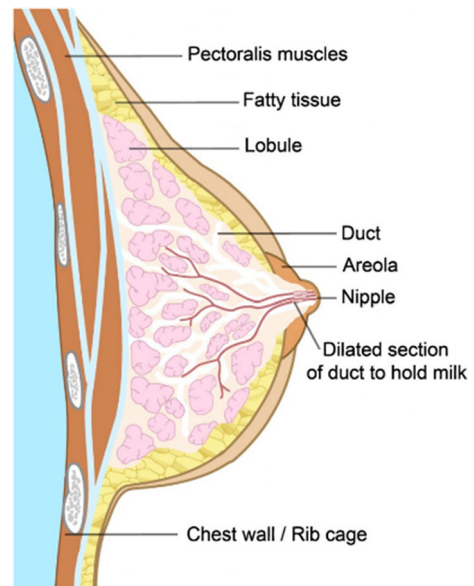
1 Introduction

Globally, the highest number of abnormal deaths are caused by Cancer which exceeds the number of deaths caused by heart diseases [27]. Breast cancer (BrC) is the most common form of cancer among women. It impacts 2.1 million women every year worldwide and is the third leading cause of cancer-related deaths [30]. In 2018, BrC emerged as the second commonly diagnosed cancer with 11.6% of the total cancer cases and 6.6% of the total female deaths due to cancer [7]. This number is perpetually on the rise globally [5] and it accounts for 14% of cancers in Indian women [11]. Studies suggest that approximately 1 in 8 women is diagnosed with invasive BrC in their lifetime and 1 in 39 women dies from BrC, if no corrective measures are taken.

The breast is the tissue in women overlying the chest muscles. Breasts are made up of fatty tissues and a specialized tissue producing milk called the glandular tissue. The milk-producing part of the breast is organized into 15 to 20 sections, called lobes as shown in Fig. 1. Within each lobe, smaller structures are present called as lobules, where milk is produced. The milk travels through a network of tiny tubes called ducts. The ducts get connected and come together into larger ducts, which eventually exit the skin in the nipple. The dark area of skin surrounding the nipple is called the areola. Connective tissue and ligaments provide support to the breast and give it its shape.

BrC is a group of diseases wherein cells in breast tissue begin to grow abnormally. These cells get divided rapidly than healthy cells and continue to accumulate, forming a mass or tumour. Most BrCs begin in the lobules or in the ducts that connect the lobules to the nipple. Broadly, cancers are divided in two categories viz. *Benign* and *Malignant*. The benign tumours are in-situ or non-invasive. They do not spread and are not harmful. However, the malignant tumours are invasive and harmful. BrC is caused by a complex interaction of environmental factors and genetic susceptibility. Survival of malignant BrC patients largely depends on early detection of tumours. The most common physical sign of BrC is a painless lump. Less

Fig. 1 Female Breast Anatomy



common signs and symptoms include breast pain or heaviness; persistent changes such as swelling, thickening, or redness of the skin; and nipple changes such as spontaneous discharge (especially if bloody), scaliness, or retraction.

A late detection of this disease is often due to the lack of early symptoms which makes the treatment challenging [18]. Since the lives of people are at stake, there is no scope for misclassification. To avoid misclassification in cancer detection, it is aimed to reduce *false negatives*, which can be reflected in *recall* performance measure. Hence, correct classification or identification of cancer is vital for early-stage treatment in cancer diagnosis.

Studies have shown that early diagnosis and suitable treatment can significantly reduce the mortality rate [17]. For an early detection of BrC, screening is extremely essential. It involves testing women to identify cancers before any physical symptoms appear. Some common clinical methods of BrC screening are mammography, breast ultrasound, biopsy and magnetic resonance imaging (MRI) [16]. Radiology images like mammograms can locate BrC lesions but cannot substantiate whether a highlighted region is cancerous or non-cancerous. In breast biopsy [34], small tissue (sample) is taken from the breast under suspicious region and examined by the microscope. This gives a more credible cancer diagnosis than radiology images. However, visual inspection of histopathological (Hp) stained tissue sections is time consuming and requires sophisticated tools. In addition, to investigate sample correctly, various factors or characteristics of the radiologist play a vital role such as field experience, expertise, domain knowledge, and regular workload [2].

To minimize diagnostic errors, digital pathology laboratories convert histopathological breast tissue slides into digital images known as histopathological images by using scanners with different zooming factors. Hp images allow the pathologist to distinguish between a normal tissue, benign tissue, and malignant lesions. Due to Hp images, computerized classification models are designed for the detection of BrC to support pathologists and assist as the basis of second opinion in BrC diagnosis, and overcome the issues of microscopic manual analysis of breast histopathological slides [3, 19].

There are numerous research works found in the literature focusing on the improvements on the various performance issues in the early-stage breast cancer detection and diagnosis. Though, existing approaches and techniques resolve the issues to a considerable extent, the aforementioned statistics are just due to the several misclassifications followed by late diagnosis. BrC is a classic example wherein we cannot afford any misclassifications due to the existing non optimal techniques or by the expert dermatologists. Due to limited accuracy in the contemporary Artificial Intelligence (AI)/ML techniques, the further improvements are possible by adopting the advancements in the DL technologies. We propose a novel approach for the breast cancer detection using multi-channel merging that demonstrates measurable improvements in comparison with the state-of-the-arts. Specifically, we propose a computer-aided transfer learning (TL) based binary classification deep model for BrC detection. Generally, deep learning architectures are sequential, following only a single channel for features' extraction and then perform the classification. Different instances of a dataset may pursue varying efficiency in features' extraction and representation against any pretrained convolutional neural network (CNN) features' extractor. Feature's fusion is a heavily adopted methodology in machine learning for various tasks. With the advent of deep learning tools and models, handcrafted features engineering is no longer needed. Using transfer learning mechanism, we can employ multiple pretrained CNNs for features' extraction and perform their fine tuning. In breast cancer classification, existing deep models adopt only single channel in the entire features' extraction and classification task. However, fused features extracted from

multiple channels may better represent features qualitatively. The motivation of the work lies in the anticipation that the meaningful combination of existing ensembles could yield improved results than their individual counterparts. To the best of our knowledge, this is the first attempt that employs multiple channels for efficient features' extraction and representation in breast cancer classification.

We performed a comparative study to analyse the performance of five pre-trained CNNs and three ensemble models to determine the best method for BrC classification on the BreakHis dataset [26]. We also performed extensive investigation in determining the best setup of hyper-parameters for five models pretrained on ImageNet namely ResNeXt101 [33], InceptionResNetV2 [34], DenseNet121 [35], InceptionV3 [36], and Xception [32] and their three ensembles Xception + InceptionV3, Xception + InceptionResNetV2, and InceptionV3 + InceptionResNetV2. These models are fine-tuned further on BreakHis dataset using TL to accurately classify the Hp images into benign or malignant cancer. Moreover, we have used the image augmentation techniques such as *rotation* and *flipping* to enhance the number of samples in model training.

The remaining sections of this paper are structured as follows. In Section 2, we present the relevant literature survey which demonstrate the BrC classification problem in medical imaging using various machine learning or deep learning models. In Section 3, we highlight the proposed methodology including the experimentation and implementation setup. In Section 4, we present the results and discuss the relevance of the results. Finally, conclusions and future scope are presented in the last section of the paper.

2 Related work

The two approaches generally used for BrC classification are machine learning (ML) and deep learning (DL). ML or DL based classification models mainly comprises of four steps: i) pre-processing of image, ii) extraction of features, iii) model raining, and iv) performance evaluation of the model. The success of traditional ML-oriented BrC classification models depends upon the discriminative hand-engineered features (HeF) extraction step [1]. They classify the mitosis candidates using different manually extracted features, such as morphological, textural, and statistical features. However, since the mitoses have varied shapes and textures, it is hard to manually define features that can effectively represent the mitotic cells [28]. The major limitations of hand-engineered feature extraction process include non-trivial tasks such as the requirement of domain knowledge and rework for each datasets publicly available. These are prone to lose important information of correlated neighbouring pixels [19]. ML system for BrC recognition based on Neural Networks (NN) and Support Vector Machine (SVM) was published in 2013 that reported 94% recognition accuracy on a dataset consisting of 92 samples only [11].

Since the arrival of DL based models especially, CNN, in the medical imaging field, researchers have developed several models based on CNN and its architectural variants to tackle different medical tasks. The last decade has seen significant efforts put forth for BrC detection using either mammograms or Hp images. Other medical modalities such as breast ultrasound and MRI are rarely used for classification. Study [15] employs both hand-crafted and convolutional features to make the mitosis detection system more effective. In DL, we do not require HeF extraction step which is a part in traditional ML approach. Moreover, DL techniques involve minimal data pre-processing tasks and identifies relevant information in a

self-taught manner. BrC classification using DL model are of two types: (i) the de novo models i.e., classification models based on DL, developed from scratch, and (ii) TL-based models comprises of models created and retrained after fine-tuning pretrained models (AlexNet) [19].

The de novo model sizes are often small whereas large dataset is needed for DL models. Moreover, the model performance can be improved by adding some customized layers, and to avoid overfitting of model large dataset with proper labels are required for training. However, we lack in the large sized labelled datasets in medical domain. In addition, deep model training on the large dataset requires more time as the parameters learning start from scratch and many computational resources are required such as high-end GPU, RAM, and storage devices, etc. The overfitting may be observed in TL-based models when the target dataset is very small as majority of pretrained models trained on large dataset. Thus, new features cannot be learned by this model when we consider target dataset with only few instances. Pretrained models are often larger or deeper (e.g. ResNet has 150 layers) and thus require a large amount of data, and consume more time and computational resources to train.

In [23], a deep model is utilized wherein 2-stage TL and evolutionary pruning approach is adopted where AlexNet is trained on ImageNet and deep CNN is utilized for training as well as TL for classification. Mammograms are segmented, and TL-based model is adopted to extract the features from Region-of-Interests (RoI). Random forest (RF) is employed for classification after features extraction for binary classification of BrC and the accuracy of 90% is achieved. Some limitations of the model include requirement of a GPU, since it was a large model and required more resources and time. The performance is measured based on sensitivity and precision. In addition, an exclusive dataset is utilized and hence the results obtained could not be compared with those results obtained by other researchers.

In [55], a transfer learning is applied in combination with diverse supervised ML algorithms to classify BrC intrinsic subtype. The Pan-cancer dataset was used to train three autoencoders (AEs) in an unsupervised manner on a heterogeneous dataset having different types of cancers. The resulting AEs were then fine-tuned in a supervised method on a reduced dataset of BrC labelled samples. The results indicated that the leveraging information from many cancer types does not contribute to solve a more complex and specific classification task of BrC intrinsic subtypes. Furthermore, an interpretation of the neural-like network focussed on the modern neuro-paradigms deficiencies and limitations elimination is performed by the geometric data transformations (GDT) method in [56]. The concept of information modelling based on a novel model of geometric transformations is presented by the authors. Geometric Transformations Model (GTM) uses a single methodological framework for various tasks and fast non-iterative study with pre-defined number of computational steps. It provides repeatability of the training outcomes and the possibility to obtain satisfactory solutions for large and small training samples.

For an extensive training process, there is a requirement for data including all the variations and characteristics in BrC. Also, it must be ensured that the images are clicked at different lighting conditions and backgrounds. Due to the huge amount of effort required, procuring a live dataset has always been a difficulty in developing models for BrC classification. For many years, research is being conducted using a very small number of samples from private datasets. In 2016 [26], a database for BrC classification problem was released, and one research group reported 85.1% accuracy utilizing SVM and Parameter Free Threshold Adjacency Statistics (PFTAS) features for patient-level analysis, which was the highest recognition accuracy at that time. In [22], MIAS and DDSM mammogram-based public datasets were used to classify BrC specifically into malignant and benign cases. Initially, ANN and CNN are applied for

mammogram segmentation. Thereafter, shape features, intensity, and texture extraction is performed on segmented images. Lastly, the performance is evaluated using Random Forest (RF), Naïve Bayesian (NB), SVM, and k-nearest neighbour (kNN) and an accuracy of 96.47% is obtained.

Most of the current studies use public datasets and very few studies have trained their CNN classifiers on large labelled datasets. To tackle this issue of small datasets, we used the image augmentation methods that includes image modifications by flipping, rotating, and scaling the image. This helps in improving the overall training process in computer vision and related tasks.

Over the years, CNNs have proven to be the most effective in image classification tasks. Long Short-Term Memory (LSTM) networks allow encapsulation of a wider sequence of images for prediction. It has the potential to memorize long-term dependencies. It has been observed in studies that LSTMs are able to complement the feature extraction ability of CNN when used in a layered order. LSTMs have the capacity to selectively remember patterns for a long duration of time and CNNs are able to extract the important features out of it. Thus, LSTM-CNN layered structure when used for image classification, has an edge over conventional CNN classifier. Study [20] classified a set of BrC images on BreakHis dataset using novel deep neural network (DNN) models guided by statistical and structural information obtained from the images. CNN, LSTM, and ensemble of CNN + LSTM were employed for the classification of BrC. Moreover, decision-making is performed by softmax and SVM layers after the feature extraction stage. The proposed method achieved the higher accuracy of 91% for BrC classification using highly accurate and efficient pre-trained models trained on the publicly available large BreakHis dataset.

The spectrum of deep learning applications has become so broad in the recent time that it has covered and performed tremendously well in various domains in comparison with its contemporary machine learning models. We can list some of them here such as natural language understanding and generation, smart city and homes, wind estimation and forecasting, and many more. The complete enumeration of these applications is not in the scope of this paper. However, we list several examples of deep models in the context of smart cities for the sake of understanding the broad application range of deep learning. Citywide crowd flow prediction and its accurate analysis is one of the most important tasks in smart cities traffic management system. Contemporary deep models such as CNN and LSTM perform comparatively well for the aforementioned task, however, neural networks incorporating the spatio-temporal features in the deep neural networks outperforms these traditional deep models in the performance improvement of crowd flow prediction [62, 63].

In Table 1, we present the comparative analysis of existing models for breast cancer detection based on the accuracy reported in the literature. Moreover, Table 2 shows the comparison of various models based on model utilized, drawbacks, and future direction. Thus, after inspecting the facts presented in Tables 1 and 2, we need robust deep model to detect the breast cancer with higher accuracy to better assist in diagnosis.

3 Materials and methods

The architecture of our proposed binary classification model for BrC detection is shown in Fig. 2. Firstly, we load the dataset and split it into training and testing set in the 80:20 ratio. We perform image pre-processing on the input images before training. To do so, we performed

Table 1 A comparative analysis of existing models for breast cancer detection based on accuracy

Ref.	Model	Dataset	Accuracy (%)
[21]	Nawaz W et al. (2018)	ICIAR 2018 Challenge	81.25
[6]	Awan R et al. (2018)	ICIAR 2018 Challenge	83.33
[28]	Vang YS et al. (2018)	ICIAR 2018 Challenge	87.50
[12]	Guo Y et al. (2018)	ICIAR 2018 Challenge	87.50
[24]	Sarkar MI et al. (2019)	ICIAR 2018 Challenge	89.00
[10]	Ferreira CA et al. (2018)	ICIAR 2018 Challenge	90.00
[14]	Kassami SH et al. (2019)	ICIAR 2018 Challenge	92.50
[29]	Wang Z et al. (2018)	ICIAR 2018 Challenge	93
[4]	Alzubaidi L et al. (2020)	ICIAR 2018 Challenge	96.1
[37]	Wang Y et al. (2020)	ICIAR 2018 Challenge	97.70
[38]	Ragab DA et al. (2019)	DDSM	87.2
[39]	Hadush S et al. (2020)	Mini-MIAS	91.86
	Kelas MK (2019)	Dataset obtained from the measurements of an antenna consisting of frequency bandwidth, dielectric constant of the antenna's substrate, electric field, and tumour information	90.0
[40]	Zheng J et al. (2020)	INbreast	97.2
[42]	Hamed G et al. (2020)	The common breast mammographic datasets created using a combination of DDSM, CBIS-DDSM, MIAS, INBreast, BCDR datasets	97
[43]	Ak MF (2020)	Wisconsin breast cancer	98.1
[44]	Vnigazova BP (2020)	Wisconsin breast cancer	99.6

Table 2 Critical review of existing models based on model used, drawbacks and future research directions

Ref. Model	Shortcomings / Drawbacks	Research direction
[4] Deep CNN	Complex model	To apply on microscopy images in same domain
[6] ResNet with contextual information and ML for classification	Designed for small dataset	NA
[10] Inception, ResNetV2 through transfer learning	Less accuracy	NA
[12] Modified GoogLeNetV3, Bagging and voting for classification	Less accuracy	Model modification to reduce the complexity
[14] Deep CNN with transfer learning	Less accuracy	To improve accuracy, use of ensemble of deep models
[21] Modified AlexNet	Less accuracy	To develop customized deep model for higher accuracy
[24] InceptionV4	Less accuracy	NA
[28] GoogLeNetV3 for heatmap probability, majority voting, gradient boosting machine, logistic regression and SVM for classification	Less accuracy	NA
[29] VGG16, VGG19	Less accuracy	To investigate the performance of deep model based on different scales
[37] AlexNet	Less accuracy	To use ensemble of VGG and ResNet to achieve higher performance
SVM as a classifier		
[38] R-CNN	Less accuracy	To investigate macro calcification abnormalities and increase dataset
	Small dataset	
[39] Classification algorithms	Less accuracy	NA
[40] ML and DL based model	Complex model	NA
[41] DenseNet, ResNet, InceptionV3, VGG16	Complex model	To apply model on clinical data, also investigate on different types of cancers
[42] Ensemble-SVM for classification		
AlexNet, GoogLeNet, YOLO, RetinaNet	Less accuracy	NA
[43] Data visualization based ML model	Complex model	NA
[44] ANOVA-BOOTSTRAP-SVM model	Complex model	To apply in other medical dataset

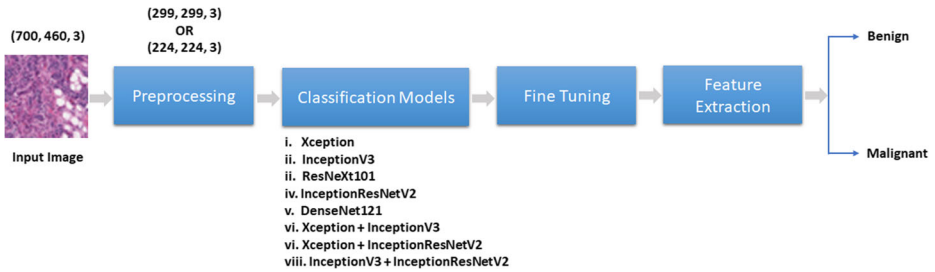


Fig. 2 Proposed Architecture for BrC Classification

image augmentation with random horizontal and vertical flips. We re-scale all the images to reconcile with the input dimensions of the pre-trained architecture used in this experimentation. Furthermore, we feed the processed images into the pre-trained architecture for features’ extraction and fine tuning. The output is then classified into benign or malignant cancer. We explored this proposition for five different pre-trained architectures namely; Xception, InceptionV3, ResNeXt101, InceptionResNetV2, and DenseNet121. To achieve more accuracy, we propose a multi-channel merging network, obtained by concatenation of two architectures from the set of five pre-trained architectures. The proposed ensemble models for the breast cancer classification includes, Xception + InceptionV3, Xception + InceptionResNetV2, and InceptionV3 + InceptionResNetV2. The architecture of one of the ensemble models, obtained using Xception + InceptionV3 is presented in Fig. 3.

3.1 Dataset

Our research focuses on identifying BrC in Hp images of the breast tissues. We used the publicly available BreakHis dataset [26] for the construction and evaluation of the proposed models. BreakHis dataset was developed jointly by Pathological Anatomy and Cytopathology, Brazil and P & D Laboratory. It comprises of 7909 microscopic images; classified into 2480 benign images and 5429 malignant images at different magnifications collected from 82 patients diagnosed with breast tumour tissues, as shown in Fig. 4 (ii). Each image is 700 × 460 pixels in size with 3-channel RGB, 8-bit depth in each colour, and stored in *png* format. Sample images of the BreakHis dataset at different magnification levels are shown in Fig. 4 (i). For this research, we split the dataset into training set (80%) and testing set (20%). We ensure

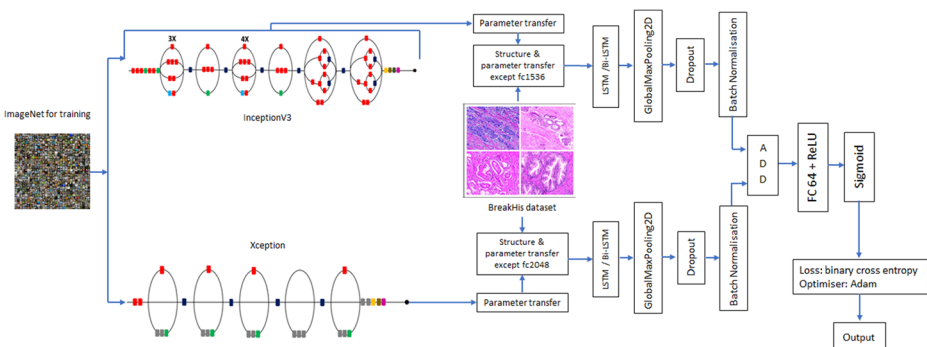


Fig. 3 Proposed Ensemble Architecture (using Xception + InceptionV3) for BrC Classification

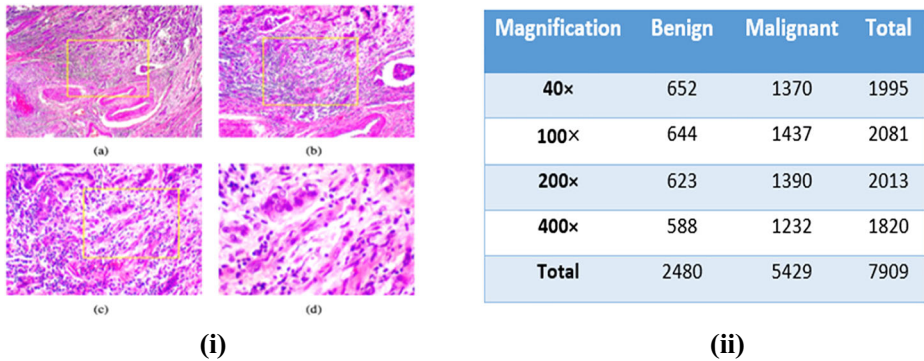


Fig. 4 BreakHis dataset: (i) Breast Tissue Samples at (a) 40x, (b) 100x, (c) 200x and (d) 400x, and (ii) Dataset Distribution into Different Magnifications

that the proportionate inclusion of samples from different magnifications in model training and testing. This helps in an unbiased evaluation of the model's performance.

3.2 Image pre-processing

This phase removes non-informative features and applies image augmentation to avoid CNN overfitting and class imbalance to improve the classification performance. CNNs achieve better performance with a larger amount of training data. Thus, we use image augmentation to synthetically create similar images by applying some image pre-processing techniques on the available images. We perform this using the built-in pre-processing function of Keras [8] ImageDataGenerator. We rotate the images up to 90 degrees and perform random horizontal and vertical flips. Figure 5 shows the resultant image. To avoid class imbalance, we select an equal number of augmented images from each class. The image size of the training set automatically changes on applying the mentioned augmentation techniques. By using bicubic interpolation (BI) method [54], we rescaled all the images to $299 \times 299 \times 3$ or $224 \times 224 \times 3$ so that the images reconcile with the input image dimension for the pre-trained models we have used. In BI, sixteen nearest neighbours of a pixel are considered. The intensity value assigned to a point (x, y) is obtained using the Eq. 1 where the 16 coefficients are determined from the sixteen equations in sixteen unknowns that can be written using the 16 nearest neighbours of point (x, y) .

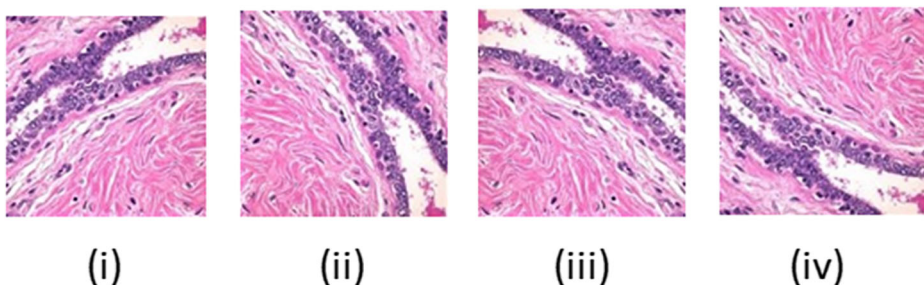


Fig. 5 Image Augmentation: (i) BreakHis sample image, (ii) 90° Rotation, (iii) Horizontal Flip, and (iv) Vertical Flip

$$v(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij}x^i y^j \tag{1}$$

where, $v(x, y)$ gives the value of intensity at the point (x, y) .

3.3 Transfer learning

In transfer learning, the knowledge or parameters of an already trained ML/DL model is applied to a different but related problem. The main advantages of transfer learning are: better initialization of model’s parameters, saving training time, better performance of neural networks, and not requiring a lots of data. In our study, we have utilized five pre-trained models namely Xception, InceptionV3, ResNeXt101, InceptionResNetV2, and DenseNet121. The architecture of each of the pre-trained models are shown in Fig. 6.

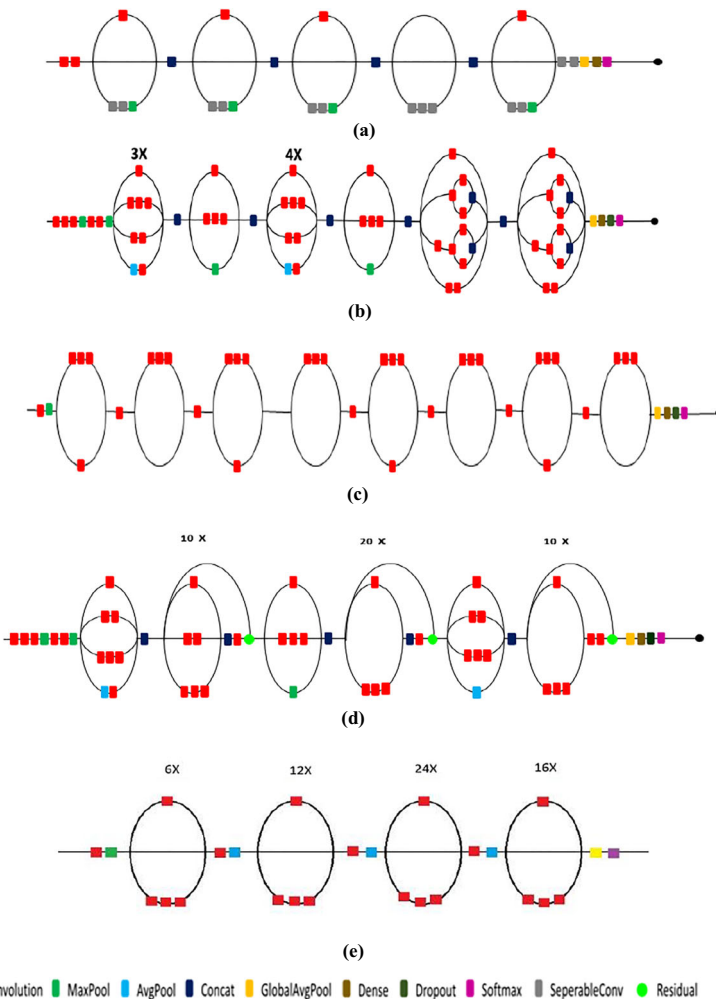


Fig. 6 Architecture of (a) Xception, (b) InceptionV3, (c) ResNeXt101, (d) InceptionResNetV2, (e) DenseNet121

3.4 Classification models: Convergence and fine tuning

We have performed some customizations on the DL architectures to make it better suited for our task of BrC detection. These customizations are described as follows:

3.4.1 Binary cross entropy as loss function

As our prediction model acts as a binary predictor, Logistic loss or Binary cross entropy loss is used for computing the model loss and can be described using Eqs. 2 and 3 for the i^{th} instance in a batch and for an entire batch respectively. y_i and $p(y_i)$ indicate the ground truth label and the predicted output by the model respectively. N indicates the number of samples in a mini batch in the gradient descent algorithm. As suggested by the following equations, predicted outcome should be close to unity for minimizing the loss if the ground truth label is true. On the other side, predicted outcome by the model should be close to zero for minimization of the loss if the ground truth label of that instance is false. In case of a batch, we take the average of log loss of all the samples in a batch.

$$H_p(q) = -y_i \log(p(y_i)) - (1-y_i) \log(1-p(y_i)) \quad (2)$$

$$H_p(q) = -\frac{1}{N} \sum_{i=1}^N y_i \times \log(p(y_i)) + (1-y_i) \times \log(1-p(y_i)) \quad (3)$$

3.4.2 Adam as optimizer

For better convergence of the gradient descent algorithm, Adam is used as a gradient descent optimizer that is a combination of momentum and RMSProp. Parameters, i.e., weight matrices in the subsequent iterations are getting updated using Eq. 4.

$$w_t = w_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \quad (4)$$

Where w is model weights, η is the learning rate, \hat{m}_t and \hat{v}_t represent momentum and RMSProp respectively and ϵ is a very small number which prevents a division by zero. Momentum and RMSProp are the exponentially moving averages of the partial derivatives of the cost function with respect to the weight matrices and biases, and can be represented using Eq. 5 and Eq. 6 respectively.

$$m_i = \beta_1 m_{i-1} + (1-\beta_1) [\nabla E \theta_i] \quad (5)$$

$$v_i = \beta_2 v_{i-1} + (1-\beta_2) [\nabla E(\theta_i)]^2 \quad (6)$$

where β_1 and β_2 are the hyperparameters, 0.9 and 0.999 are the values of these hyperparameters respectively considered in our experimentations.

3.4.3 CNN-LSTM/BiLSTM layer

We have added a CNN-LSTM/BiLSTM layer to our model. In bidirectional LSTM, the output at time t is not only dependent on the previous frames in the sequence, but also on the upcoming frames [14]. Bidirectional Recurrent Neural Networks (RNNs) have two RNNs stacked on top of each other. One RNN goes in the forward direction and another one goes in the backward direction. The combined output is then computed based on the hidden state of both RNNs. The CNN-LSTM/BiLSTM network architecture used in our work is specifically designed for sequence prediction problems with spatial inputs like images. It involves using CNN layers for feature extraction on input data combined with LSTM/BiLSTMs to support sequence prediction. This method is accurate for generating a textual description of a single image. CNN layers are added on the front end followed by LSTM/BiLSTM layers with a dense layer producing an output. The CNN model helps in feature extraction whereas the LSTM/BiLSTM model interprets the features across different time steps. The CNN-LSTM/BiLSTM is presented in the Fig. 7.

3.4.4 Multi-channel merging

Generally, deep learning architectures are sequential, following only a single channel for features' extraction and further classification. However, fused features extracted from multiple channels may better represent features qualitatively. This is the backbone of our entire experimentation. Here, instance specific characteristics plays a vital role in the feature extraction process by various pretrained CNNs, although, the instances belong to the same or similar datasets. It might be possible that one pretrained CNN model perform better features' extraction for an instance whereas it may not capture the similar efficient features for another instance. This may probably happen due to the inherent intra class variance among the different instances of the same dataset. This is the overall intuition behind employing two pretrained CNNs on two different channels, extracted features by these pretrained CNNs are merged just before the final classification. However, we have not explored more than two channels to maintain the model complexity to a significant level.

We have used two pre-trained blocks namely, Xception and InceptionV3 for feature extraction as shown in Fig. 3. We have transformed each block into sub-nets with distinct convolutional kernels and merge the two branches to get a combined convolutional network for super resolution by concatenating the different features extracted by the two sub-nets individually. Using these techniques, we developed three multi-channelled ensemble architectures namely Xception + InceptionV3, Xception + InceptionResNetV2, and InceptionV3 + InceptionResNetV2. Table 3 presents the complete architectural details of the proposed model to visualize the layer-by-layer composition, specifically for Xception and InceptionV3 based multi-channel deep model. Moreover, Fig. 8 also illustrates the detailed structure of multi-channel deep model proposed in this paper.



Fig. 7 CNN-LSTM/BiLSTM Network Architecture

Table 3 Proposed multi-channel model’s layer wise details

Layers	Input dimension	Output dimension
Input layer 1	(299, 299, 3)	(299, 299, 3)
Input layer 2	(299, 299, 3)	(299, 299, 3)
Xception layer	(299, 299, 3)	(10, 10, 2048)
InceptionV3 layer	(299, 299, 3)	(8, 8, 2048)
Reshape layer 1	(10, 10, 2048)	(100, 2048)
Reshape layer 2	(8, 8, 2048)	(64, 2048)
LSTM layer 1	(100, 2048)	(128)
LSTM layer 2	(64, 2048)	(128)
Reshape layer 3	(128)	(1, 2, 64)
Reshape layer 4	(128)	(1, 2, 64)
Max Pool layer 1	(1, 2, 64)	(64)
Max Pool layer 2	(1, 2, 64)	(64)
Dropout layer 1	(64)	(64)
Dropout layer 2	(64)	(64)
Batch Norm. layer 1	(64)	(64)
Batch Norm. layer 2	(64)	(64)
Concat layer	(64, 64)	(64)
Dense layer	(64)	(2)

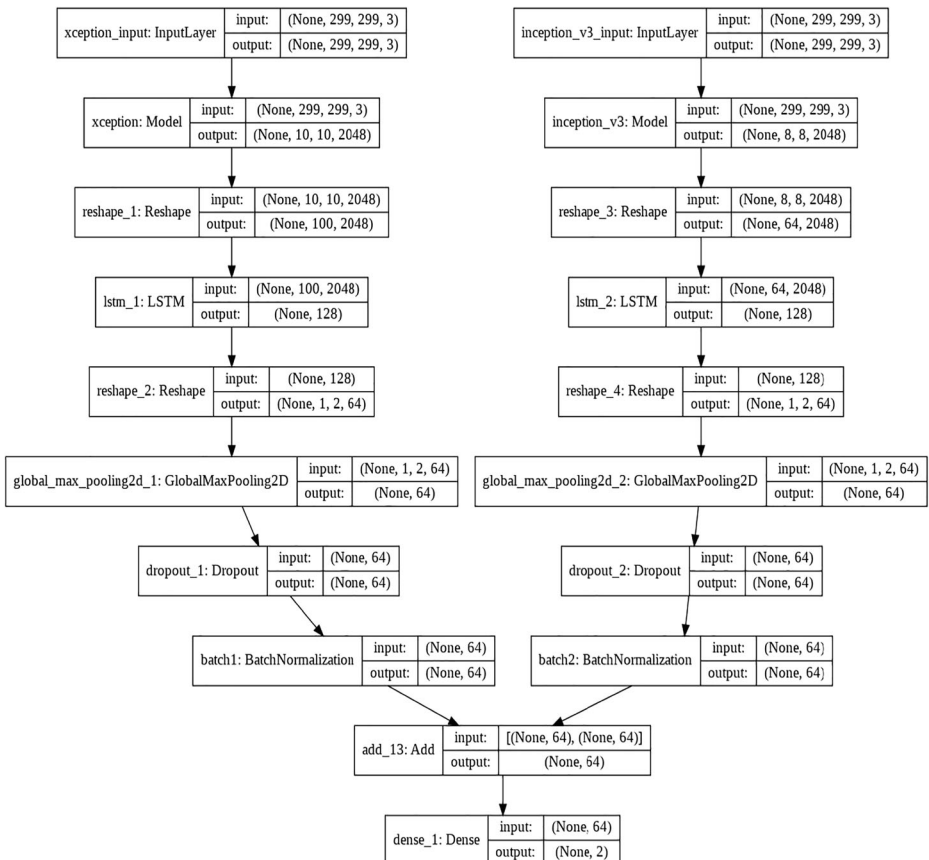


Fig. 8 Structure of multi-channel deep model for BrC Classification

3.4.5 Hyper-parameters tuning

The learning rate is the hyper-parameter that determines how fast we are adjusting the weights of our network or model towards the local minima. Higher learning rate can result in abrupt weight changes that it might result in overshooting the local minima. This causes the training or testing error to fluctuate drastically between consecutive epochs. Moreover, lower learning rate can result in taking longer time to train our network or model as little steps are adopted towards the local or global minima of the loss curve. Thus, the learning rate is one of the most important hyper-parameters that needs to be tuned intelligently when building the model. We have set the learning rate for our model to 0.0001.

Furthermore, an epoch indicates the number of iterations required of the entire training set processed by the ML/DL model. It is impossible to know the suitable number of epochs for a model training. Thus, we have added a call-back function to run epochs for a considerable number of times before there is no further improvements in accuracy. In addition, batch size also play a vital role in model training i.e. batch size refers to the number of training examples utilized in one iteration. We have trained the proposed model on a batch size of 8 for 20 epochs. Moreover, we monitor the value of testing accuracy for the proposed model. When there is no further improvement in the testing accuracy for 5 consecutive epochs, the learning rate is reduced by a factor of 0.2 up to a minimum of 0.0000001.

3.5 Performance metrics

The different performance measures are used for evaluating the various machine learning and deep learning models based on the problem and context. Accuracy, precision, recall, and F1-score are the most important performance measures in case of classification tasks. However, Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) are the performance measures, generally adopted in case of regression problems. As we are focusing on binary classification in our cancer detection task using TL-based deep models, we precisely define the performance measures used in our experimentation to evaluate the proposed model. Accuracy is a ratio of correct classification to the total number of instances, as indicated using Eq. 6. It does not take specific misclassification into account. Precision can be defined as how many samples are True Positives (TP) among all the samples predicted as positives and presented with the help of Eq. 7. However, to take care of False Negatives (FN), Recall is introduced and can be defined as the ratio of True positives to the Actual Positives samples. It is presented using Eq. 8. F1-score is the most important performance measures that takes care of both Precision as well as Recall and can be defined as a harmonic mean of both of these, as indicated using Eq. 9. Here, TP indicate True Positives, FP indicate False Positives, TN represent True Negatives, and FN represent False Negatives.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (6)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (7)$$

$$Recall = \frac{TP}{(TP + FN)} \quad (8)$$

$$F1\text{-score} = 2 \left(\frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (9)$$

3.6 Learning curves

The performance of model learning can be visualized by the learning curves over the time. Basically, learning curve is adopted to investigate the performance of proposed ML/DL model that how it is learning at the time of training from the targeted dataset incrementally. By observing learning curve for the targeted dataset, we can get an insight of learning behaviour of proposed model. It also gives an idea of how well the model is generalizing.

To measure the losses in the proposed model, we have plotted the loss curve. It gives an overview of the training process and the direction in which the model learns. In this work, we have generated loss curve across the number of epochs wherein the loss function calculates the loss across every data item for each epoch. This gives the quantitative loss measure at each epoch. To get a good insight on the model loss, we plotted the training and testing loss on the same graph. In the loss graphs, we observe a continuous decrease in both the training and testing loss. Towards the last few epochs, the training and testing loss decreases to a point of stability with a minimal gap between the two final loss values. This shows that our loss curves have obtained a good fit. The gap between the training and the testing loss in the loss graph is generally known as the generalization gap. We observe a very small gap between the two losses in our model, which implies that our model has achieved a good generalization.

Another performance measure is accuracy curve where an overview of the training process and the direction in which the model learns can be visualized and we can depict the progress of deep model. To visualize the accuracy progress over the time for the targeted dataset, we have generated the accuracy curve across the number of epochs i.e., accuracy is calculated across every data items. So, we plotted the training and testing accuracy on the same graph for the better understanding of model losses. The gap between training and testing accuracy is a clear indication of overfitting i.e., larger the gap implies higher the overfitting. In our case, we observe a minimal gap between the training and testing accuracies for all the proposed models. So, we conclude that our model does not overfit and we have also ensured this by applying dropout layers of 50% to our models. Dropout layers disable the neurons during training and reduces the complexity of the proposed models.

4 Result and discussion

This section reports and discusses the experimental results of proposed models in terms of overall predictive accuracy and other performance measures. All the experiments are accomplished using Deep Learning library [8] on Intel® Core™ i7-8550U CPU @ 1.80GHz processor with 8GB RAM enabled with NVIDIA GeForce 940MX graphic card. Moreover, python 3.7 is utilized for the modelling and programming purpose.

The performance is evaluated in terms of F1-score, recall, precision, and prediction accuracy for all the standalone and multi-channelled architectures. Table 4 shows the various measures obtained for individual architectures on BreakHis dataset. The categorical accuracy for Xception, InceptionV3, ResNetXt101, InceptionResNetV2, and DenseNet121 are found to

Table 4 Independent Model's Accuracy, Precision, Recall and F1-score for BreakHis dataset

Model	Accuracy (%)	Weighted Average		
		Precision (%)	Recall (%)	F1-score (%)
Xception	92.17	85	85	85
InceptionV3	91.75	85	83	84
ResNeXt101	90.43	84	84	84
InceptionResNetV2	90.45	81	81	81
DenseNet121	89.13	81	82	82

be 92.17%, 91.75%, 90.43%, 90.45%, and 89.13% respectively. Xception and InceptionV3 show the highest accuracy, followed by InceptionResNetV2. The accuracy of the ensemble models viz. Xception + InceptionV3, Xception + InceptionResNetV2, and InceptionV3 + InceptionResNetV2 are 97.5%, 96.25%, and 95.12% respectively and the results are shown in Table 5. It took approximately 2h to train each of the ensemble models.

We also evaluate the weighted averages of recall, precision, and F1-scores to check the performance of models with respect to the number of images for each class of testing data. We observed that the Xception (92.17%) and InceptionV3 (91.75%) shows the best accuracy among all the five architectures.

Moreover, amongst the ensemble models, Xception + InceptionV3 achieves the highest accuracy of 97.5%. Our proposed architecture for the Xception + InceptionV3 model outperforms several models that have been trained and evaluated on the BreakHis dataset. The comparison amongst proposed models and other existing models for the BrC classification on the BreakHis dataset is shown in Table 6. Thus, based on the accuracy, we propose the application of Xception + InceptionV3 model for the detection of BrC. The training-testing accuracy curves and training-testing loss curves for the five individual models and the three ensemble models are represented in Figs. 9 and 10 respectively.

From Fig. 9, we observe very good start for both the training and testing accuracy by using transfer learning. In the starting epochs itself, models are blessed with the considerable training and testing accuracies. If we would have started model training from scratch without any transfer learning, we would have observed very less training and testing accuracies, resulting more epochs to reach at the convergence points. However, with the usage of multichannel merging, features are fused from multiple channels to better represent the qualitative features. From Fig. 10, it is clear that the start of the training accuracy is not good but model evolved eventually to grasp the advantages of multi-channel merging and outperform the standalone models after certain epochs.

Our model (Xception + InceptionV3) classifies images with an accuracy of 97.50% and a loss of 0.1, and outperforms many state-of-the-art models. In addition, if we replace the CNN-

Table 5 Ensemble Model's Accuracy, Precision, Recall and F1-score for BreakHis dataset

Model	Accuracy (%)	Weighted Average		
		Precision (%)	Recall (%)	F1-score (%)
Xception + InceptionV3	97.50	89	89	89
Xception + InceptionResNetV2	96.25	87	88	88
InceptionV3 + InceptionResNetV2	95.12	86	86	86

Table 6 Comparison with existing models trained on the BreakHis dataset

Ref.	Model	Accuracy (%)
[13]	Gupta V et al. (2017)	88.89
[45]	Chan A et al. (2016)	96.4
[46]	Bardou D et al. (2018)	96.15
[47]	Kahya MA et al. (2017)	94.54
[48]	Alom MZ et al. (2019)	92.04
[49]	Arslan AK et al. (2019)	91.4
[50]	Bayramoglu N et al. (2016)	84.63
[51]	Spanhol FA et al. (2017)	84.2
[52]	Araújo et al. (2017)	83.3
[53]	Jiang Y et al. (2019)	98.87
[57]	Wenzhong at al. (2020)	96.43
[58]	Gupta at el. (2020)	93.27
[58]	Gupta at el. (2020)	86.73
[59]	Nawaz at el. (2018)	95.40
[60]	Patil at el. (2019)	86.56
[61]	Hameed at el. (2020)	95.29
Our	Xception	92.17
	InceptionV3	91.75
	ResNeXt101	90.43
	InceptionResNetV2	90.45
	DenseNet121	89.13
	Xception + InceptionV3	97.50
	Xception + InceptionResNetV2	96.25
	InceptionV3 + InceptionResNetV2	95.12

LSTM layer with a CNN-BiLSTM layer, then we achieve an accuracy of 96.25% and a loss of 0.11 for the Xception + InceptionV3 model, measured separately. The CNN-LSTM shows slightly better scores for precision, recall, and F1-score, as compared to the CNN-BiLSTM architecture. These results show that BiLSTM also performs better for the BrC prediction task as compared to the LSTM layer. Misclassification needs to be absolutely eradicated before this model can be utilized for second opinion for BrC patients.

From Table 6, we depict that [13, 48–52, 58, 60] has achieved the accuracy in the range of 84% to 92% which is lower as compared to the accuracy (97.50%) of our proposed model. Thus, an improved accuracy is achieved, an increment of around 5.5% to 13.5% by the proposed model on the BreakHis dataset. Nevertheless, we witness the higher/similar accuracy by our multi-channel model that matches the state-of-the-art [45, 46, 53, 57] on the BreakHis dataset.

5 Conclusions and future research directions

This research propose a new model for the detection of BrC. It is a binary classification model based on transfer learning that takes biopsy images as an input and classifies into benign or malignant cancer. Histopathological images from the publicly available BreakHis dataset are utilized for training and testing of the proposed models. The two major heuristics i.e., transfer learning from the pretrained models and multi-channel merging are experimented in the course of breast cancer detection and classification on the aforementioned dataset. The idea behind the

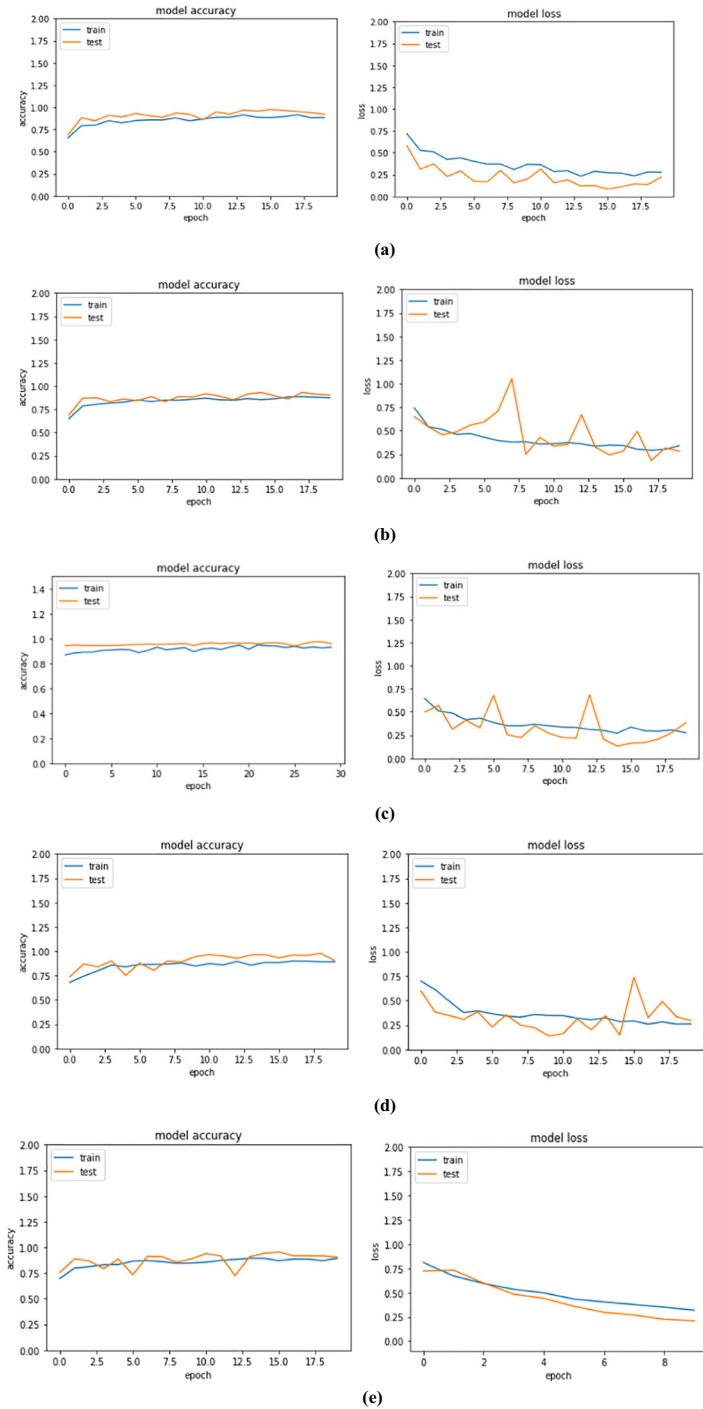


Fig. 9 Training-testing Accuracy and Loss Curves for (a) Xception, (b) InceptionV3, (c) ResNeXt101, (d) InceptionResNetV2, and (e) DenseNet121 models

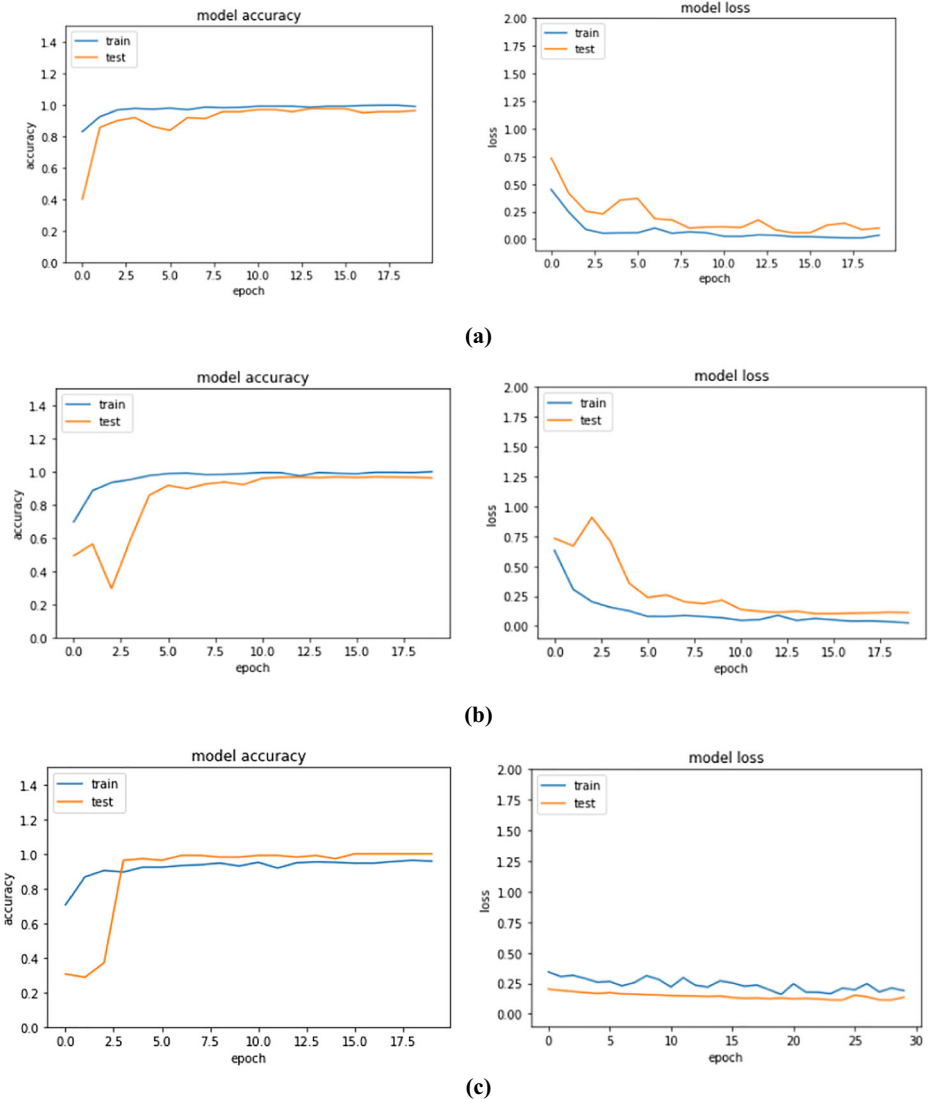


Fig. 10 Training-testing Accuracy and Loss Curves for (a) Xception + InceptionV3, (b) Xception + InceptionResNetV2, and (c) InceptionV3 + InceptionResNetV2

usage of transfer learning is quite intuitive and experimented a lot in the literature to provide better starting point in the model training. Generally, deep learning architectures are sequential, following only a single channel for features' extraction and further classification. However, fused features extracted from multiple channels may better represent features qualitatively. This was the intuition behind employing multichannel merged networks in breast cancer classification. To the best of our knowledge, multichannel merging along with transfer learning is utilized for the first time in the breast cancer classification task. The model not only performed well with the standalone architectures but ensembles using multichannel merging also outperform state-of-the-arts in the breast cancer classification. Moreover, the

higher accuracy is achieved for the proposed standalone and ensemble model is 92.17% and 97.50% respectively. Although, ensembles are used globally to increase the accuracy in classification tasks but it also significantly increases the architectural complexity, leading to much longer training time for the model. We also observe that our architecture for the Xception model also outperforms several models. Thus, keeping in mind the model complexity and increased training time, the Xception model can also serve as a reasonable alternative for the Xception + InceptionV3 model.

Future scope for this work includes using histopathological images for detection of other types of cancers. The binary classifier in the model can be extended to solve multi-class classification problems that can classify different types and stages of BrC like predicting the four subtypes of benign BrC viz. tubular adenoma (TA), fibroadenoma (F), phyllodes tumour (PT), adenosis (A) and the four subtypes of malignant BrC viz. mucinous carcinoma (MC), lobular carcinoma (LC), papillary carcinoma (PC), ductal carcinoma (DC)). It can also be experimented to classify other types of cancers such as liver, lung, prostate, bladder, colon, etc. using biopsy images. The architectural variants such as different types of convolutions (tiled-, transpose-, and dilated-convolution), other ensembles, faster prototyping, and various optimization techniques should be explored for further enhancements in the breast cancer classification.

References

1. Aghdam MH, Heidari S (2015) Feature selection using particle swarm optimization in text categorization. *J Artificial Intell Soft Comput Res* 5(4):231–238
2. Allison KH, Reisch LM, Carney PA, Weaver DL, Schnitt SJ, O' Malley FP, Elmore JG (2014) Understanding diagnostic variability in breast pathology: lessons learned from an expert consensus review panel. *Histopathology* 65(2):240–251. <https://doi.org/10.1111/his.12387>
3. Alom MZ, Yakopcic C, Taha TM, Asari VK (2019) Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. *J Digit Imaging* 32(4):605–617
4. Alzubaidi L, Al-Shamma O, Fadhel MA, Farhan L, Zhang J, Dyan Y (2020) Optimizing the performance of breast Cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model. *Electronics* 9(3):445. <https://doi.org/10.3390/electronics9030445>
5. American Cancer Society (2020) Breast Cancer Facts and Figures 2019–2020. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breastcancer-facts-and-figures-2019-2020.pdf>.
6. Awan R, Koohbanani N, Shaban M, Lisowska A, Rajpoot N (2018) Context-aware learning using transferable features for classification of breast cancer histology images. In: *International conference image analysis and recognition*. Springer, Cham, pp 788–795
7. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global Cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J Clin* 68(6):394–424. <https://doi.org/10.3322/caac.21492>
8. Chollet F (2015) Keras. <https://github.com/fchollet/keras>.
9. Cytecure Cancer Hospitals (2020) Statistics of breast cancer in India. <https://cytecure.com/blog/statistics-ofbreast-cancer/>. Accessed 10 June, 2020.
10. Ferreira C, Melo T, Sousa P, Meyer M, Shakibapour E, Costa P, Campilho A (2018) Classification of breast cancer histology images through transfer learning using a pre-trained inception ResNet v2. In: *International conference image analysis and recognition*. Springer, Cham, pp 763–770. <https://doi.org/10.1007/978-3-319-93000-8-86>
11. George YM, Zayed HH, Roushdy MI, Elbagoury BM (2014) Remote computer-aided breast Cancer detection and diagnosis system based on cytological images. *IEEE Syst J* 8(3):949–964. <https://doi.org/10.1109/JSYST.2013.2279415>

12. Guo Y, Dng H, Song F, Zhu C, Liu J (2018) Breast Cancer histology image classification based on deep neural networks. In: International conference image analysis and recognition. Springer, Cham, pp 827–836
13. Gupta V, Bhavsar A (2017) Breast cancer histopathological image classification: is magnification important?. In: proceedings of the IEEE conference on computer vision and pattern recognition workshops - CVPRW'17, pp 17–24. doi: <https://doi.org/10.1109/CVPRW.2017.107>.
14. Kassani SH, Kassani PH, Wesolowski MJ, Schneider KA, Deters R (2019) Breast cancer diagnosis with transfer learning and global pooling. In: Proceedings of the international conference on information and communication technology convergence –ICTC'19, pp 519–524. <https://doi.org/10.1109/ICTC46691.2019.8939878>
15. Malon CD, Cosatto E (2013) Classification of mitotic figures with convolutional neural networks and seeded blob features Journal of Pathology Informatics 4.
16. Mayo Clinic (2019) Breast Cancer. <https://www.mayoclinic.org/diseases-conditions/breastcancer/diagnosis-treatment/drc-20352475>. Accessed 19 July 2020.
17. Migowski A (2015) Early detection of breast cancer and the interpretation of results of survival studies. *Ciencia & saude coletiva* 20(4):1309–1310. <https://doi.org/10.1590/1413-81232015204.17772014>. (Portugese)
18. Milosevic M, Jankovic D, Milenkovic A, Stojanov D (2018) Early diagnosis and detection of breast cancer. *Technol Health Care* 26(4):729–759. <https://doi.org/10.3233/THC-181277>
19. Murtaza G, Shuib L, Wahab AWA, Mujtaba G, Raza G (2020) Ensembled deep convolution neural network-based breast cancer classification with misclassification reduction algorithms. *Multimed Tools Appl* 79:18447–18479. <https://doi.org/10.1007/s11042-020-08692-1>
20. Nahid AA, Mehrabi MA, Kong Y (2018) Histopathological breast cancer image classification by deep neural network techniques guided by local clustering, BioMed Research International
21. Nawaz W, Ahmed S, Tahir A, Khan HA (2018) Classification of breast cancer histology images using AlexNet. In: International conference image analysis and recognition. In: international conference image analysis and recognition. Springer, Cham, pp 869–876
22. Rouhi R, Jafari M, Kasaee M, Keshavarzian P (2015) Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Syst Appl* 42(3):990–1002
23. Samala RK, Chan HP, Hadjiiski LM, Helvie MA, Richter C, Cha K (2018) Evolutionary pruning of transfer learned deep convolutional neural network for breast cancer diagnosis in digital breast tomosynthesis. *Phys Med Biol* 63(9):095005. <https://doi.org/10.1088/1361-6560/aabb5b>
24. Sarker MI, Kim H, Tarasov D, Akhmetzanov D (2019) Inception architecture and residual connections in classification of breast Cancer histology images. arXiv 2019 arXiv preprint arXiv:1912.04619
25. Sommer C, Fiaschi L, Hamprecht FA, Gerlich DW (2012) Learning-based mitotic cell detection in histopathological images. In: Proceedings of the 21st international conference on pattern recognition – ICPR'12, pp 2306–2309
26. Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2016) A dataset for breast cancer histopathological image classification. *IEEE Trans Biomed Eng* 63(7):1455–1462
27. US Cancer Statistics Working Group (2017) United States cancer statistics: 1999–2014 incidence and mortality web-based report. US Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute, Atlanta
28. Vang YS, Chen Z, Xie X (2018) Deep learning framework for multi-class breast Cancer histology image classification. In: International conference image analysis and recognition. Springer, Cham, pp 914–922
29. Wang Z, Dong N, Dai W, Rosario SD, Xing EP (2018) Classification of breast cancer histopathological images using convolutional neural networks with hierarchical loss and global pooling. In: International conference image analysis and recognition. Springer, Cham, pp 745–753
30. WHO. Breast cancer. <https://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/>. Accessed 18 July 2020.
31. Zhang L, Zhang B, Coenen F, Lu W (2013) Breast cancer diagnosis from biopsy images with highly reliable random subspace classifier ensembles. *Mach Vis Appl* 24(7):1405–1420. <https://doi.org/10.1007/s00138-012-0459-8>
32. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition – CVPR'17, pp 1800–1807. <https://doi.org/10.1109/CVPR.2017.195>
33. Xie S, Girshick R, Dollár P, Tu Z, He K (2017) Aggregated residual transformations for deep neural networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition – CVPR'17, pp 1492–1500. <https://doi.org/10.1109/CVPR.2017.634>
34. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-first AAAI conference on artificial intelligence – AAAI'17, pp 4278–4284. <https://dl.acm.org/doi/10.5555/3298023.329818838>

35. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition – CVPR’17, pp 4700–4708. <https://doi.org/10.1109/CVPR.2017.243>
36. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition – CVPR’16, pp 2818–2826. <https://doi.org/10.1109/CVPR.2016.308>
37. Ragab DA, Sharkas M, Marshall S, Ren J (2019) Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ* 7:e6201
38. Hadush S, Girmay Y, Sinamo A, Hagos G (2020) Breast Cancer detection using convolutional neural networks. arXiv preprint arXiv:2003.07911
39. Keleş MK (2019) Breast cancer prediction and detection using data mining classification algorithms: a comparative study. *Tehnički vjesnik* 26(1):149–155
40. Zheng J, Lin D, Gao Z, Wang S, He M, Fan J (2020) Deep learning assisted efficient AdaBoost algorithm for breast cancer detection and early diagnosis. *IEEE Access* 8:96946–96954. <https://doi.org/10.1109/ACCESS.2020.2993536>
41. Wang Y, Lei B, Elazab A, Tan EL, Wang W, Huang F, Wang T (2020) Breast cancer image classification via multi-network features and dual-network orthogonal low-rank learning. *IEEE Access* 8:27779–27792
42. Hamed G, Marey MAER, Amin SES, Tolba MF (2020) Deep learning in breast Cancer detection and classification. In: Joint European-US workshop on applications of invariance in computer vision. Springer, Cham, pp 322–333
43. Ak MF (2020) A comparative analysis of breast Cancer detection and diagnosis using data visualization and machine learning applications. *Healthcare* 8(2):111 Multidisciplinary digital publishing institute
44. Vrigazova BP (2020) Detection of malignant and benign breast Cancer using the ANOVA-BOOTSTRAP-SVM. *J Data Inform Sci* 5(2):62–75
45. Chan A, Tuszynski JA (2016) Automatic prediction of tumour malignancy in breast cancer with fractal dimension. *R Soc Open Sci* 3(12):160558
46. Bardou D, Zhang K, Ahmad SM (2018) Classification of breast Cancer based on histology images using convolutional neural networks. *IEEE Access* 6:24680–24693
47. Kahya MA, Al-Hayani W, Algamal ZY (2017) Classification of breast cancer histopathology images based on adaptive sparse support vector machine. *J Appl Mathematics Bioinform* 7(1):49
48. Alom MZ, Yakopcic C, Nasrin MS, Taha TM, Asari VK (2019) Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. *J Digit Imaging* 32(4):605–617
49. Arslan AK, Yaşar Ş, Çolak C (2019) Breast cancer classification using a constructed convolutional neural network on the basis of the histopathological images by an interactive web-based interface. In: Proceedings of the 3rd international symposium on multidisciplinary studies and innovative technologies – ISMSIT’19, pp 1–5. <https://doi.org/10.1109/ISMSIT.2019.8932942>
50. BBayramoglu N, Kannala J, Heikkilä J (2016) Deep learning for magnification independent breast cancer histopathology image classification. In: Proceedings of the 23rd international conference on pattern recognition –ICPR’16, pp 2440–2445. <https://doi.org/10.1109/ICPR.2016.7900002>
51. Spanhol FA, Oliveira LS, Cavalin PR, Petitjean C, Heutte L (2017) Deep features for breast cancer histopathological image classification. In: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics –SMC’17, pp 1868–1873. <https://doi.org/10.1109/SMC.2017.8122889>
52. Araújo T, Aresta G, Castro E, Rouco J, Aguiar P, Eloy C, Campilho A (2017) Classification of breast cancer histology images using convolutional neural networks. *PLoS One* 12(6):e0177544
53. Jiang Y, Chen L, Zhang H, Xiao X (2019) Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. *PLoS One* 14(3):e0214587. <https://doi.org/10.1371/journal.pone.0214587>
54. Rajarapolu PR, Mankar VR (2017) Bicubic interpolation algorithm implementation for image appearance enhancement. *Int J Comput Sci Inf Technol* 8(2):23–26 <http://www.ijest.com/vol18/8.2/4-prachi-r-rajapollu.pdf>
55. López-García G, Jerez JM, Franco L, Veredas FJ (2019) A transfer-learning approach to feature extraction from Cancer transcriptomes with deep autoencoders. In: Advances in computational intelligence, IWANN’19. Lecture notes in computer science, vol 11506. Springer, Cham. https://doi.org/10.1007/978-3-030-20521-8_74
56. Tkachenko R, Izonin I (2019) model and principles for the implementation of neural-like structures based on geometric data transformations. In: advances in computer science for engineering and education, ICCSEE’18. Advances in intelligent systems and computing, vol 754. Springer, Cham. https://doi.org/10.1007/978-3-319-91008-6_58

57. Wenzhong L, Huanlan L, Caijian H, Liangjun Z (2020) Classifications of breast cancer images by deep learning. medRxiv. <https://doi.org/10.1101/2020.06.13.20130633>
58. Gupta K, Chawla N (2020) Analysis of histopathological images for prediction of breast Cancer using traditional classifiers with pre-trained CNN. *Procedia Comput Sci* 167:878–889
59. Nawaz M, Sewissy AA, Soliman THA (2018) Multi-class breast Cancer classification using deep learning convolutional neural network. *Int J Advanced Comput Sci Appl (IJACSA)* 9:6. <https://doi.org/10.14569/IJACSA.2018.090645>
60. Patil A, Tamboli D, Meena S, Anand D, Sethi A (2019) Breast Cancer histopathology image classification and localization using multiple instance learning. In: *Proceedings of the IEEE international WIE conference on electrical and computer engineering -WIECON-ECE'19*, pp 1–4. <https://doi.org/10.1109/WIECON-ECE48653.2019.9019916>
61. Hameed Z, Zahia S, Garcia-Zapirain B, Javier Aguirre J, María Vanegas A (2020) Breast cancer histopathology image classification using an ensemble of deep learning models. *Sensors* 20(16):4373
62. Ali A, Zhu Y, Zakarya M (2021) A data aggregation based approach to exploit dynamic spatio-temporal correlations for citywide crowd flows prediction in fog computing. *Multimed Tools Appl*:1–33. <https://doi.org/10.1007/s11042-020-10486-4>
63. Ali A, Zhu Y, Chen Q, Yu J, Cai H (2019) Leveraging spatio-temporal patterns for predicting citywide traffic crowd flows using deep hybrid neural networks. In: *Proceedings of the IEEE 25th international conference on parallel and distributed systems – ICPADS'19*, pp 125–132. <https://doi.org/10.1109/ICPADS47876.2019.00025>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Jitendra V. Tembhurne¹ · Anupama Hazarika² · Tausif Diwan¹

Anupama Hazarika
anupama_csb17@agnee.tezu.ernet.in

Tausif Diwan
tdiwan@iiitn.ac.in

¹ Department of Computer Science & Engineering, Indian Institute of Information Technology, Nagpur, India

² Department of Computer Science & Engineering, Tezpur University, Assam, India